December 2018

State Innovation Models (SIM) Initiative Evaluation

Model Test Year Five Annual Report
Methods Appendix

Prepared for

Jennifer Lloyd, PhD

Centers for Medicare
& Medicaid Services

Center for Medicare and Medicaid Innovation

7500 Security Boulevard
Mail Stop WB-06-05
Baltimore, MD 21244-1850

Prepared by

RTI International
3040 E. Cornwallis Road
Research Triangle Park, NC 27709

RTI Project Number: 0212790.007

Contract Number:
HHSM-500-2010-00021i

Task Order Number:
HHSM-500-T0007

RTI INTERNATIONAL

URBAN INSTITUTE

NASHP
NATIONAL ACADEMY FOR STATE HEALTH POLICY

CMS
CENTERS FOR MEDICARE & MEDICAID SERVICES

[this page intentionally left blank]

State Innovation Models (SIM) Initiative Evaluation

Model Test Year Five Annual Report
Methods Appendix

RTI International

Stephanie Kissam, Project Director

Heather Beil, Associate Project Director

Leslie Greenwald, Associate Project Director

CMS Contract No. HHSM-500-2010-00021i

Task Order Number: HHSM-500-T0007

December 2018

This project was funded by the Centers for Medicare & Medicaid Services under contract no. HHSM-500-2010-00021i. The statements contained in this report are solely those of the authors and do not necessarily reflect the views or policies of the Centers for Medicare & Medicaid Services. RTI International assumes responsibility for the accuracy and completeness of the information contained in this report.

[this page intentionally left blank]

SIM Initiative Evaluation Model Test Year 5 Annual Report – Methods Appendix

by

RTI International

Stephanie Kissam, Project Director
Heather Beil, Associate Project Director
Leslie Greenwald, Associate Project Director

Richard Alterbaum
Peter Amico
Samuel Arbes
Sarah Arnold
Rebecca Brophy
Judy Cannada
Denise Clayton
Aubrey Collins
Julia Kish Doto
Patrick Edwards
Rose Feinberg
Sara Freeman
Emily Graf
Catherine Hersey
Jesse Hinde
Elise Hooper
Sara Jacobs
Ravikumar Jeyaraman
Madeleine Jones
Monika Juzumaite
Douglas Kendrick
Vince Keyes
Konny Kim
Marianne Kluckman
Paul Moore
Marisa Morrison
Michelle Myers
Tim O'Brien
William Parish
Sheila Patel
Melissa Romaire
Regina Rutledge
Noelle Richa Siegfried
Holly Stockdale
Elizabeth Tant
Matthew Toth
Carol Urato
Matthew Urato
Nathan West
Suzanne West
Nathaniel Wiecha
Alton Wright

National Academy for State Health Policy

Christina Cousart
Rachel Donlon
Hannah Dorr
Maureen Hensley-Quinn
Scott Holladay
Diane Justice
Jennifer Reck

The Urban Institute

Eva Allen
Kyle Caswell
Teresa Coughlin
Sharon Long
Timothy Waidmann

Federal Project Officer: Jennifer Lloyd

December 2018

[this page intentionally left blank]

Table of Contents

The SIM Initiative Model Test Year Five Annual Report is available in a separate file, available at the following location: https://downloads.cms.gov/files/cmmi/sim-rd1-mt-fifthannrpt.pdf.

Appendices

List of Figures

Number **Page**

List of Tables

Number **Page**

List of Acronyms

Acronym	Definition
AC	Accountable Communities
ACH	Accountable Community for Health
ACO	accountable care organization
AHCPII	Arkansas Health Care Payment Improvement Initiative
AHRQ	Agency for Healthcare Research and Quality
APCD	all-payer claims database
BCBS	Blue Cross Blue Shield
BHH	behavioral health home
CCM	Coordinated Care Model
CCO	Coordinated Care Organizations
CFR	Code of Federal Regulations
CHIP	Children's Health Insurance Program
CMS	Centers for Medicare & Medicaid Services
DD	developmental disabilities
DHS	Department of Human Services
D-in-D	difference-in-differences
D-SNPs	Dual Eligible Special Needs Plans
DSRIP	Delivery System Reform Incentive Payment
EDIE	Emergency Department Information Exchange
EHR	electronic health record
ENS	event notification system or service
EOC	episode of care
ER	emergency room
ERISA	Employee Retirement Income Security Act
FFS	fee for service or fee-for-service (*adj.*)
FQHC	federally qualified health centers
GMCB	Green Mountain Care Board
HCBS	home and community-based services
HCH	health care home
health IT	health information technology
HIE	health information exchange
HPC	Health Policy Commission
I/DD	intellectual or developmental disabilities
IHP	Integrated Health Partnership
LTSS	long-term services and supports

Acronym	Definition
MCO	managed care organization
MCPAP	Massachusetts Child Access and Psychiatry Project
MMIS	Medicaid Management Information System
NCQA	National Committee for Quality Assurance
OEBB	Oregon Educators Benefit Board
PAP	principal accountable provider
PCC	Primary Care Clinicians
PCMH	patient-centered medical home
PCPCH	patient-centered primary care home
PCPRI	Primary Care Payment Reform Initiative
PEBB	Public Employees Benefit Board
PMPM	per member per month
QHP	qualified health plan
SIM	State Innovation Models
SPA	state plan amendment
SSP	Shared Savings Program
VHIE	Vermont Health Information Exchange
VPM	value-based payment model

Sub-appendix A-2. Methods for Arkansas Analyses

A-2.1 Methods for the Arkansas URI Episode Impact Analysis

URIs are one of the most common sources of illness in Arkansas (Arkansas Center for Health Improvement [ACHI], 2015). Inappropriate use of antibiotics for those with URI is a significant problem for Medicaid programs nationally and in Arkansas (Brown, Taylor, Rogers, Weiser, & Kelley, 2003; Li, Metlay, Marcus, & Doshi, 2014; Zuckerman, Perencevich, & Harris, 2007). To reduce the inappropriate use of antibiotics for URI episodes among Medicaid beneficiaries in Arkansas, the state introduced a retrospective EOC that sets expected expenditure thresholds holding the principal accountable provider (PAP) at risk for gain or loss. The model encourages and incentivizes the PAP to coordinate with a patient's patient-centered medical home, Arkansas's complimentary and concurrent SIM reform, during the 21-day episode.

Profile of participating providers. The PAP for an episode is defined as the first Arkansas Medicaid-enrolled and qualified provider who diagnoses a beneficiary with an acute ambulatory URI during an in-person visit. All providers who treat any Medicaid beneficiaries are required to participate in the EOC model. By 2015, 421 to 601 PAPs, depending on the type of URI, that saw five or more of the three types of URI episodes during the year. As described in the introduction, 20 to 50 percent of PAPs had what was considered commendable average URI expenditures in 2012. By 2014, the range of PAPs' performance remained similar across the three types of URIs (Arkansas Department of Human Services [ADHS], 2016).

Intervention group. Arkansas introduced URI EOCs in July 2012, but providers were not subject to financial risk until October 2012. The intervention group for this analysis comprises Medicaid beneficiaries who were diagnosed with a URI in an office, outpatient clinic, or emergency department (ED) in Arkansas from October 2012 to September 2014 and who met exclusion criteria, as described ***Section A-2.1.5*** below. The baseline period spans October 2011 to September 2012, and the post period was October 2012 to September 2014.

Comparison group. Our comparison group consisted of Medicaid beneficiaries who were diagnosed with a URI in an office, outpatient clinic, or emergency department in Mississippi or Missouri. We selected these two states as comparisons states based upon data availability, calculated Euclidean distance scores, and similar income thresholds for Medicaid eligibility (Kaiser Family Foundation [KFF], 2018). The Medicaid income eligibility limit for parents of children (other than pregnant women) in Missouri was 23–37 percent below the federal poverty level (FPL), followed by Mississippi's 28–44 percent below the FPL; in comparison, Arkansas's income limit was 16–17 percent below FPL from 2010 to 2013 but reached 138 percent below FPL in 2014 when Arkansas expanded Medicaid under the private

option. Even so, the beneficiaries covered by the private option plans are not included in our data as their claims are not in the MAX files.

Study sample. Our sample included Medicaid-covered URI episodes with a live birth that was diagnosed in an office, clinic, or emergency department from 2011 through 2014. We excluded beneficiaries with different types of coverage during the episode, including Children's Health Insurance Program, supplemental private insurance coverage, or dual Medicare-Medicaid enrollment. To mimic the Arkansas episode criteria, we removed beneficiaries with certain comorbidities and those with any hospitalization or observation care during the episode. We also removed overlapping episodes, and episodes with restricted Medicaid benefits.

Balancing URI EOC and comparison groups. Following comparison group selection, we constructed an episode-level propensity score weight to balance the Arkansas URI EOCs and comparison group EOCs on select observed individual characteristics. Weights were created for the combined subtypes of all URI episodes. We decided to apply propensity score weights to the analysis instead of other methods, such as propensity score matching, to retain sample size and produce less-biased estimates for binary outcomes. After propensity score weighting, the standardized differences between the weighted comparison group means and Arkansas URI EOC group means were under the standard 10 percent threshold, except for some area-level characteristics. More information on propensity score weighting is available in ***Section A-2.1.6***.

Study design. We used a D-in-D design, comparing changes in the outcome variables before and during the implementation of the demonstration period for Arkansas's URI EOC group with changes in the outcomes before and during the demonstration period for the comparison group. We conducted a pre-post multivariate regression analysis to examine URI related expenditures. We used retrospective annual cross-sections of URI EOCs from 2011 to 2014.

Statistical approach. The analyses used ordinary least squares for expenditure outcomes and logistic regression for utilization and quality outcomes. All regression analyses used clustered standard errors at the individual level to account for the clustering of multiple episodes for a person. The outcome models controlled for age, gender, race, disability status, health status, urban/rural area of residence, and county-level characteristics. More information on the study outcomes is available in ***Section A-2.1.3*****,** and more information on the regression model is available in ***Section A-2.1.9***.

A-2.1.1 Data sources

MAX data. The RTI evaluation team used Medicaid data from the CMS MAX and Alpha-MAX research files made available through the Chronic Conditions Data Warehouse enclave. Each state's Medicaid Statistical Information System (MSIS) data is the source of the MAX and Alpha-MAX files. The MAX processing adds enhancements, such as claims

adjustments, the creation of a national type of service field, and state-specific quality issue corrections; Alpha-MAX provides fewer enhancements. The MAX and Alpha-MAX files include a person summary (PS) file, with all enrollment information and summary claims information and four claims files: inpatient hospital (IP), long-term care (LT), prescription drugs (RX), and other (OT) claims. The quarterly Alpha-MAX files are generated for a state once all five MSIS file types for a single quarter are approved. The quarterly files are overwritten and updated each time a new quarter of run-out data is added. Quarterly versions of Alpha-MAX are being produced for each state through seven quarters of run-out data; therefore, the quarterly files are based on zero to seven quarters of run-out time. Annual calendar year MAX files are prepared from data with seven quarters of run-out time. For simplicity, the MAX and Alpha-MAX data are simply referred to as MAX data for the remainder of this appendix

Area Health and Resource File (AHRF). The AHRF comprises data collected by the Health Resources and Services Administration from more than 50 sources containing more than 6,000 variables related to health care access at the county level. We used information on health professions supply, poverty at the county level, age, rural/urban status, and uninsured rates from 2010–2015 to select the comparison group and to use as covariates in the analysis.

A-2.1.2 Outcome measures

Utilization

Utilization measures are reported as a probability of any utilization within the episode. These measures include whether the 21-day episode included any of the following visits or use: any physician visits (overall and URI specific), ED visit (overall and URI specific), and antibiotic use.

Any physician visit (overall). We used OT claims to identify whether any of the following procedure codes were on a claim where the service end date falls between episode begin date + 1 and the episode end date. The procedure codes checked were 99201, 99202, 99203, 99204, 99205, 99211, 99212, 99213, 99214, 99215, 99241, 99242, 99243, 99244, 99245, 99304, 99305, 99306, 99308, 99309, 99310, 99315, 99316, 99318, 99324, 99325, 99326, 99327, 99328, 99334, 99335, 99336, 99337, 99339, 99340, 99341, 99342, 99343, 99344, 99345, 99347, 99348, 99349, 99350, 99358, 99359, 99366, 99368, 99374, 99375, 99376, 99377, 99378, 99379, 99380, 99385, 99386, 99387, 99396, 99397, 99401, 99402, 99403, 99404, 99405, 99406, 99407, 99408, 99409, 99410, 99411, 99412, 99420, 99495, 99496, G0402, G0438, G0439, and G0463.

Any physician visit (URI-specific). This indicator is defined above but includes an additional check for the primary diagnosis code. The procedure code on the claim must also have an accompanying diagnosis code of 034.0x, 460.xx 461.0x–461.3x, 461.8x, 461.9x, 462.xx, 463.xx, 464.0x, 464.00, 464.10, 464.20, 465.0x, 465.8x, 465.9x, 460.xx, 464.0x, 464.00, 464.10,

464.20, 465.0x, 465.8x or 465.9x, 462.xx, 463.xx, 034.0x, 461.0x, 461.1x, 461.2x, 461.3x, 461.8x, 461.9x, or 786.2x

Any ED visit (overall). This indicator is used to identify unique values of MSIS_ID and service end dates in the OT claim where the revenue code equals 0450–0459, 0981, or 0762. We also identify whether the place of service code equals 23 and at least one procedure code includes 99281, 99282, 99283, 99284, or 99285. Claims are excluded if the procedure code always equals 70000 through 79999 or 80000 through 89999 (it can equal one of these values as long as it also equals another value) and if the revenue code is never equal to 0762.

Any ED visit (URI-specific). This indicator is defined above but includes the additional condition that the identifying claim must have a diagnosis code of 034.0x, 460.xx, 461.0x–461.3x, 461.8x, 461.9x, 462.xx, 463.xx, 464.0x, 464.00, 464.10, 464.20, 465.0x, 465.8x, 465.9x, 460.xx, 464.0x, 464.00, 464.10, 464.20, 465.0x, 465.8x or 465.9x, 462.xx, 463.xx, 034.0x, 461.0x, 461.1x, 461.2x, 461.3x, 461.8x, 461.9x.

Any antibiotic use. This is an indicator for whether there was a prescription fill date during the 21-day URI episode where the National Drug Code (NDC) identifies the following antibiotics: amoxicillin, sulfamethoxazole, clindamycin, doxycycline, erythromycin, cephalexin, minocycline, penicillin, azithromycin, moxifloxacin, ofloxacin, clarithromycin, cefuroxime axetil, ciprofloxacin, levofloxacin, cefdinir, and cefpodoxime proxetil. These antibiotics are identified using the NDC codes in Healthcare Effectiveness Data and Information Set (HEDIS) Table CWP-C (2015).

Quality of care

Quality measures are reported as the probability that the episode meets the quality measure numerator criteria, conditional on meeting the denominator criteria. To evaluate the impact of the URI EOC on quality of care, we report the following measures:

Appropriate treatment for children with pharyngitis (National Committee for Quality Assurance). This is an indicator for children 3–18 years of age who were diagnosed with pharyngitis, ordered an antibiotic, and received a group A streptococcus (strep) test for the episode (Agency for Healthcare Research and Quality [AHRQ], 2015).

Denominator. Children 3–18 years of age, with a Negative Medication History, who had an outpatient visit, an observation visit or an ED visit with only a diagnosis of pharyngitis and a dispensed antibiotic for that EOC.

Negative Medication History. There were no antibiotic prescriptions filled between the URI begin date and 30 days prior to the begin date. NDC codes were identified using HEDIS Table CWP-C (2015).

Dispensed Antibiotic. Any NDC code from Table CWP-C (2015) on or 3 days after the beginning of the episode.

Numerator. An indicator for beneficiaries who met the denominator condition and received a group A streptococcus test in the 7-day period from 3 days prior to the episode start date through 3 days after the episode start date. The numerator would equal one if any procedure code in the OT claims were 87070, 87071, 87081, 87430, 87650–87652, or 87880.

Appropriate treatment for children with URI (National Quality Forum #0069). This outcome identifies children between the ages of 3 months and 18 years with a URI diagnosis who were NOT prescribed an antibiotic within the first 3 days of the URI diagnosis. The URI EOC excludes those under 1 year; thus, for the purposes of our analysis, we selected ages 1 to 18.

Denominator. Children aged 1 year to 18 years and who had a negative history of antibiotic use in 30 days prior to the start of the episode. A negative history was identified using the NDC codes from Table CWP-C (2015) where the prescription fill date fell between the URI begin data and 30-days prior to the began date on the RX claim. Pharyngitis URI episodes were also excluded.

Numerator. The numerator was set to one if there were no prescription fill dates between the URI begin date and 3-days after the begin date. We used NDC codes from Table CWP-C (2015) to identify any antibiotic prescriptions were filled.

Medicaid Payments. This study does not evaluate the effect of the URI EOC model on Medicaid payments. However, to understand trends in URI Medicaid payments over the study period in Arkansas, we considered the following types of payments:

Total other services payments. Sum all OT payments from the beginning of the episode until the end of the episode. MAX_TOS = 1–19, 23–54, and 99, and TYPE_CLM_CD=1 OR 5.

Total URI-related other services payments. Sum all OT payments from the beginning of the episode until the end of the episode. MAX_TOS = 1–19, 23–54 and 99 and TYPE_CLM_CD=1 OR 5, and DIAG_CD = 034.0x, 460.xx 461.0x–461.3x, 461.8x, 461.9x, 462.xx, 463.xx, 464.0x, 464.00, 464.10, 464.20, 465.0x, 465.8x, 465.9x, 460.xx, 464.0x, 464.00, 464.10, 464.20, 465.0x, 465.8x or 465.9x, 462.xx, 463.xx, 034.0x, 461.0x, 461.1x, 461.2x, 461.3x, 461.8x, 461.9x, or 786.2x

Total prescription payments. Sum of all RX payments from the beginning to the end of the episode. MAX_TOS = 1–19, 23–54, and 99, and TYPE_CLM_CD=1 OR 5

Total antibiotic payments. Sum of all RX payments from the beginning to the end of the episode. MAX_TOS = 1–19, 23–54, and 99, and TYPE_CLM_CD=1 OR 5, and NDC is found in Table CWP-C (2015).

A-2.1.3 Population studied

Intervention (Arkansas URI EOC) group. The URI episode is triggered by an office, clinic, or ED visit with a primary diagnosis of acute ambulatory URI that does not fall within the 21-day window as a previous diagnosis.

Identifying acute ambulatory URIs. To capture these episodes, we used OT claims and identified all claims where there was a primary diagnosis of 034.0x, 460.xx, 461.0x–461.3x, 461.8x, 461.9x, 462.xx, 463.xx, 464.0x, 464.00, 464.10, 464.20, 465.0x, 465.8x, or 465.9x AND a corresponding procedure code equal to 99201–99205, 99211–99215, 99241–99245, 99281–99285, or T1015, T1015 U1-U3, or G0463 AND the place of service code equaled 11, 20, 22, 23, 49, 50, 71, or 72. From the service begin date of the URI claim, 20 days were added to identify the episode end date, where the service begin date of the triggering claims counts as day 1, creating a 21-day window. Subsequent URI claims that fall between the begin and end dates do not start a new episode. A new episode only begins when the begin date of the subsequent claims is greater than the end date of the index claim. Therefore, if the begin date of the subsequent claim is less than or equal to the end date of the previous claim of the same MSIS_ID, reset the subsequent begin and end dates to equal those of the earliest claim associated with that episode to indicate that they are a part of the same episode.

Arkansas defined three types of URI episodes: non-specific URI, sinusitis, and pharyngitis. To create mutually exclusive subtypes, we identified the last claim in the URI episode where the diagnosis code equaled 460.xx, 464.0x, 464.00, 464.10, 464.20, 465.0x, 465.8x or 465.9x, 462.xx, 463.xx, 034.0x, 461.0x, 461.1x–461.3x, 461.8x, or 461.9x. *Non-specific URI* was identified if the diagnosis code equaled 460.xx, 464.0x, 464.00, 464.10, 464.20, 465.0x, 465.8x or 465.9x. *Sinusitis* was identified if the diagnosis code equaled 461.0x, 461.1x–461.3x, 461.8x, or 461.9x. Pharyngitis was identified if the diagnosis code equaled 462.xx, 463.xx, or 034.0x.

Applying exclusion criteria. We attempted to create URI episodes consistent with the exclusion and inclusion criteria rules applied in Arkansas. Initially, 3,271,939 total episodes among Arkansas and the comparison group between 2011 and 2014 were included. Episodes where the beneficiary had restricted benefits (N = 121,166), was not continuously enrolled in full Medicaid coverage (N = 574,195) for the 21 day episode period, or was younger than 1 year (N = 367,507) were excluded. To be consistent with Arkansas's episode creation logic, we excluded episodes in which a comorbid diagnosis (e.g., croup, epiglottitis, URI with obstruction, pneumonia, influenza, and otitis media) occurred (N = 285,357). Episodes with the following comorbidities that were diagnosed at least twice during the year prior to the episode start date were further excluded: asthma, cancer, chronic URI, end-stage renal disease, HIV and other immuno-compromised conditions, post-procedure state for transplants, pulmonary disorders, rare genetic diseases, and sickle cell anemia (N = 477,521). We excluded beneficiaries with tonsillectomy or adenoidectomy (n = 7,469) and those with any inpatient stays (N = 24,976) or observation stays (N = 195,755) during the episode. After the exclusion criteria were applied, our sample included 1,681,962 total URI episodes among 802,357 unique Medicaid beneficiaries.

Our exclusion criteria removed episodes that may have been more complex or high cost and, therefore, are not subject to the episode payment criteria established by Arkansas. As a

sensitivity analysis to determine if our results are robust to the inclusion of all episodes, we estimated its outcome models using all URI episodes represented within the claims.

Finally, the overall URI sample was stratified by children aged 1 through 18 and adults aged 19 and over. The prevalence of URI diagnosis among Medicaid beneficiaries was concentrated among children; therefore, we would expect the URI EOC to have more pronounced effects on children than on the adult population.

A-2.1.4 Comparison group and propensity score weighting

For the impact analysis, a pre-post comparison group design was used, in which the comparison group provides an estimate of what would have happened among Arkansas URI EOCs absent the introduction of the EOC model. The difference in the changes over time from the pre period to the intervention period between Arkansas URI EOCs and their comparison group provides an estimate of the impact of the Arkansas URI EOC model. The comparison group should be similar to the intervention group on all relevant dimensions (e.g., demographic, socioeconomic, political, regulatory, and health and health systems) except for the policy change being tested.

The following section details the procedures used to select the comparison group for the Arkansas URI EOC population.

Selection of comparison group

Our comparison group consisted of Medicaid beneficiaries who were diagnosed with a URI in an office, outpatient clinic, or in emergency department in Mississippi or Missouri. We selected these two states as comparison states based upon data availability, calculated Euclidean distance scores, and similar income thresholds for Medicaid eligibility (Kaiser Family Foundation [KFF], 2018). The Medicaid income eligibility limit for parents of children (other than pregnant women) in Missouri was 23–37 percent below the federal poverty level (FPL), followed by Mississippi's 28–44 percent below the FPL; in comparison, Arkansas's income limit was 16–17 percent below FPL from 2010 to 2013 but reached 138 percent below FPL in 2014 when Arkansas expanded Medicaid under the private option. Even so, the beneficiaries covered by the private option plans are not included in our data as their claims are not in the MAX files.

The availability of MAX claims was also a key determinant in our selection of comparison states. Other fee-for-service Medicaid states did not have enough post-years of MAX claims to compare with Arkansas. However, each state manages and pays for Medicaid services differently. Missouri primarily relies on comprehensive managed care plans to serve their Medicaid beneficiaries. Similarly, starting in 2014, the population enrolled in Medicaid managed care plans in Mississippi expanded greatly. As a result, this analysis relies on encounter data to observe utilization, and we did not include expenditures for the comparison group in the analysis. This analysis also depends on the completeness of inpatient and professional encounter records

for these EOCs. A relatively recent encounter analysis by Mathematica Policy Research indicated that in both Mississippi and Missouri, 2010 and 2011 inpatient and professional claims were complete and usable for health services research (Byrd & Dodd, 2015). No more recent analysis is available; however, episodes had similar inpatient, professional, and drug encounter utilization within each state for each year, indicating that the data were reasonably complete for this analysis.

Calculation of person-level weights

To balance the population characteristics for the claims-based analyses, we estimated propensity scores for all episodes from the comparison group. A propensity score is the probability that an episode will be in the intervention group rather than the comparison group.

The objective of propensity score modeling is to create a weighted comparison group with characteristics equivalent to those of the URI EOC episodes. To the extent that these characteristics are correlated with utilization and quality outcomes, propensity weighting will also help balance pre-intervention levels of the outcomes.

Person-level characteristics

The initial step in the process was to select the person-level characteristics to be used in each propensity score model. ***Table A-2-1*** shows the characteristics used grouped by whether they control for demographics, enrollment, attribution, or beneficiary health status. Because limited information is available in claims data, we considered also including county-level characteristics to control for geographic characteristics, such as physician supply and median income, to account for potential differences in access to care or other geographic differences. However, little variation in county-level characteristics was found, which made balancing on these variables difficult. Therefore, to optimize the balance and avoid extreme weights, county-level covariates were excluded from the propensity score model. However, we did control for county-level characteristics in the outcome model.

Estimation and weighting procedures

Using the characteristics listed in ***Table A-2-1***, we estimated propensity models by logistic regression, in which the outcome was one for a URI episode among Arkansas Medicaid beneficiaries and zero for the comparison group. Separate models were estimated for 2011, 2012, 2013, and 2014 data. To achieve suitable balance, age and disability were interacted with race and health status characteristics.

Table A-2-1. Covariates for propensity score logistic regressions

Covariates
Demographic characteristics
Male (Ref: Female)
Age and age squared
Black
Hispanic
Other
White (Ref)
Disabled
Trigger location was ED
Health status measures
Chronic Illness and Disability Payment score (count of major comorbidities)
ED use during baseline years
Inpatient admission during baseline years
Area-level characteristics (outcome model only)
Metropolitan status of county of residence
Percent of population at FPL, 2012
Hospital beds per 1,000, 2010
Median age, 2010
Percent uninsured, ages <65, 2012

ED = emergency department; FPL = federal poverty level.

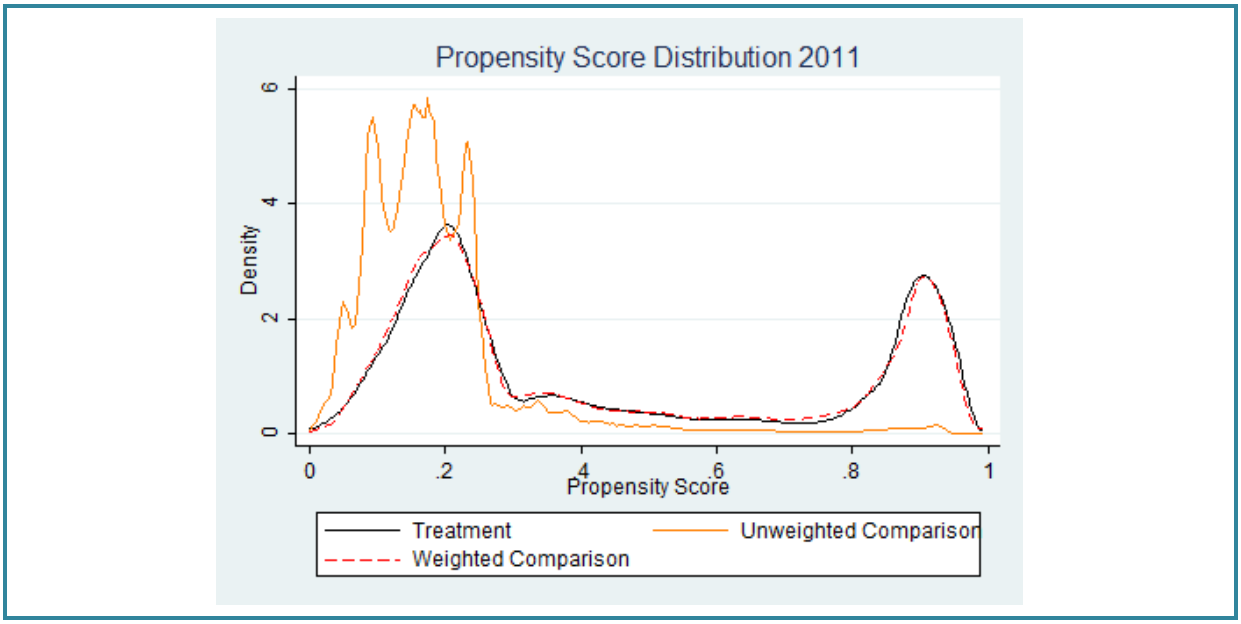
Propensity weights were set to one for all individuals in the intervention group. The propensity weight for a comparison individual was a function of his or her predicted propensity score, where weight = p/(1-p), and p is the predicted propensity. Our procedure typically includes trimming weights that are either less than 0.05 or greater than 20, although in this analysis, no weights needed trimming.

A-2.1.5 Propensity model evaluation

We evaluated several aspects of the propensity score models. First, plots of predicted probabilities were examined to ensure sufficient overlap in the distributions of the intervention and comparison groups. This feature, known as common support, is critical because it provides the basis for inferring effects from group comparisons (***Figures A-2-1*** to ***A-2-4***).

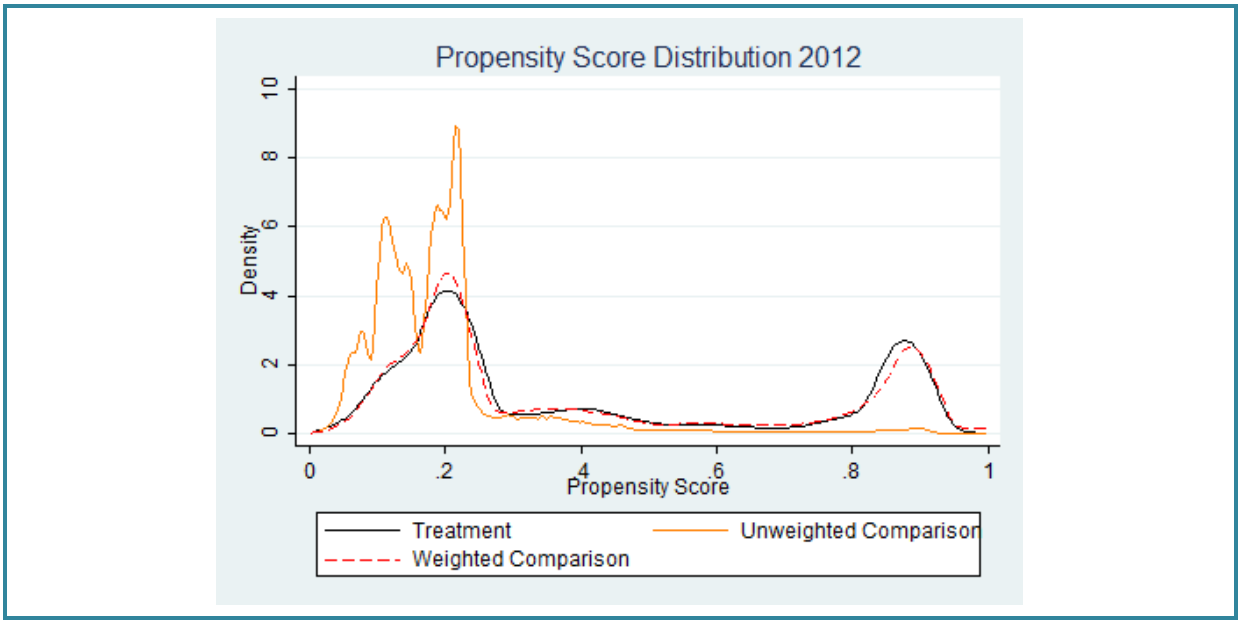
In all years, that the comparison group passed the common support assumption ($P(D = 1|X)>0$) for almost the entire range of the intervention group's propensity scores. The only exceptions were in the uppermost percentiles of the intervention group's distribution (above the 99th percentile). These plots provide ample evidence that the common support assumption is upheld.

Figure A-2-1. Weighted and unweighted propensity score density plots for the Arkansas URI Episode of Care and comparison groups, 2011[1]



URI = upper respiratory infection.

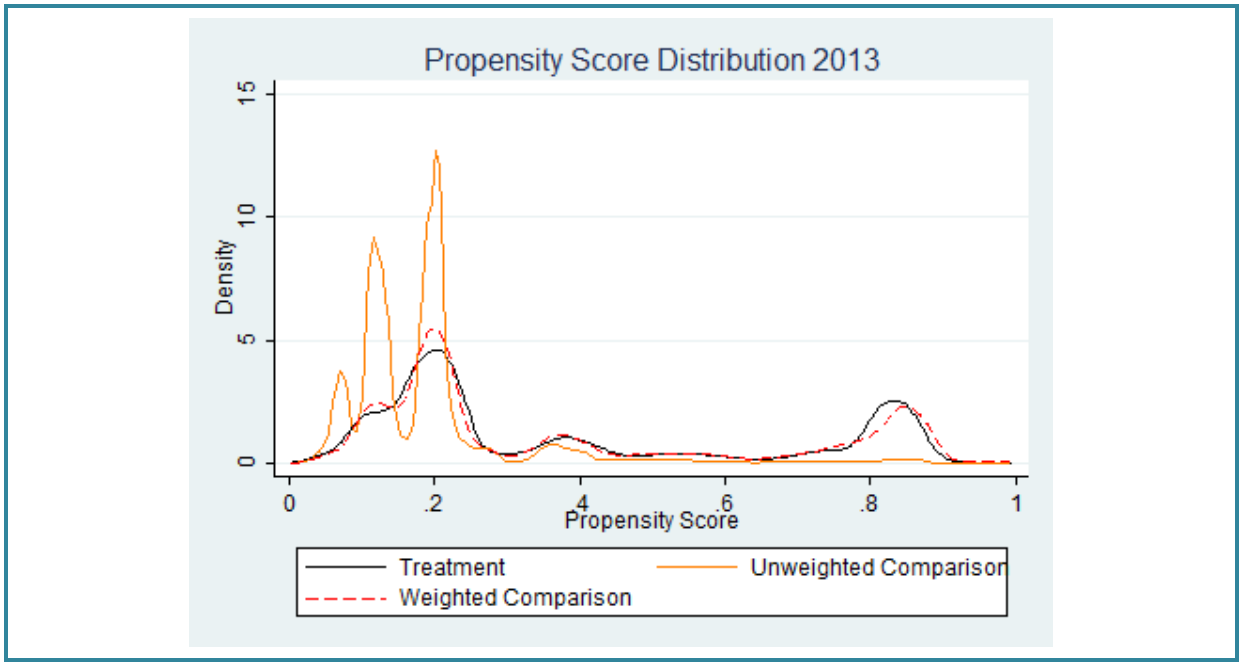
Figure A-2-2. Weighted and unweighted propensity score density plots for the Arkansas URI Episode of Care and comparison groups, 2012



URI = upper respiratory infection.

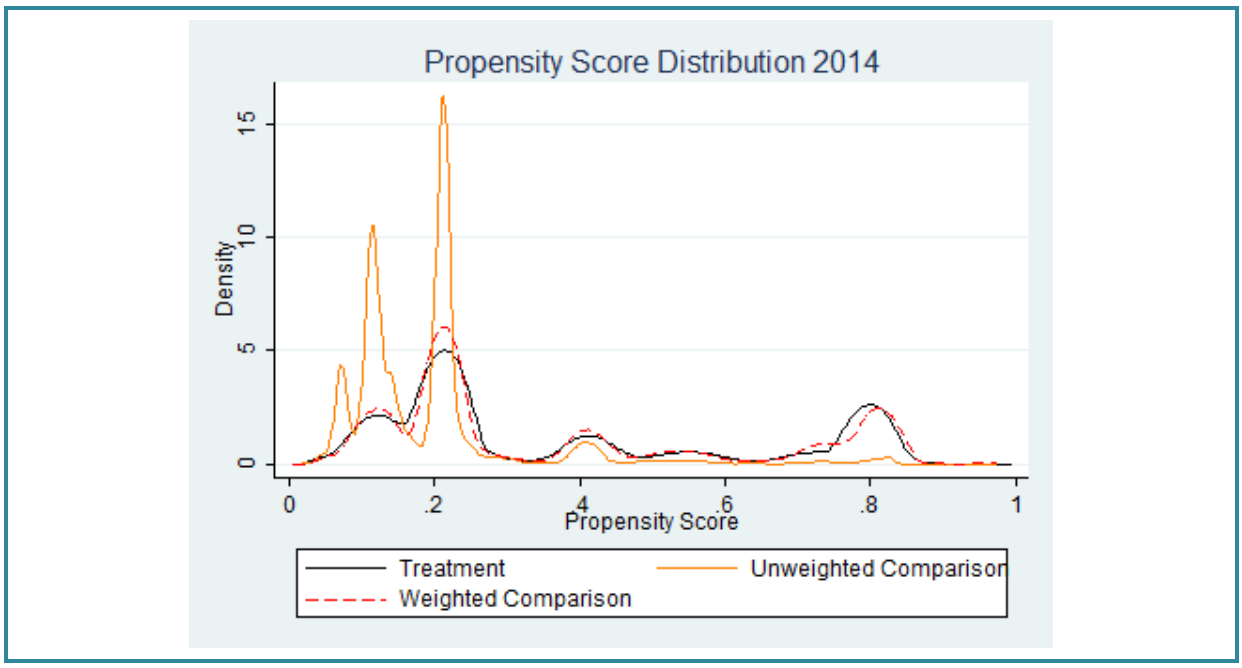
[1] In ***Figures A-2-1*** through ***A-2-4***, the Treatment lines represent those in the Arkansas URI EOC group.

Figure A-2-3. Weighted and unweighted propensity score density plots for the Arkansas URI Episode of Care and comparison groups, 2013



URI = upper respiratory infection.

Figure A-2-4. Weighted and unweighted propensity score density plots for the Arkansas URI Episode of Care and comparison groups, 2014



URI = upper respiratory infection.

Second, we compared the logistic results of the models to see which variables had the greatest impact on the propensity score weights. The major differences between the groups were race, whether the episode was triggered at an ED, and age. Beneficiaries in Arkansas with a URI episode were younger, less likely to be black, more likely to be Hispanic, and less likely to have the episode triggered at an ED.

Finally, unweighted and propensity-weighted means for the characteristics in the model were compared. As expected, the comparison group means were within a few percentage points of the values for the intervention group after applying the propensity weights.

Tables A-2-2 to ***A-2-5*** show the unweighted and propensity score-weighted means/proportions for 2011–2014. The notable group differences in the unweighted samples—age, attribution, and socioeconomic factors—are substantially mitigated post-weighting, as evidenced by the minimized standardized differences.

A-2.1.6 Propensity model evaluation for subpopulation

In addition to the overall model, common support graphs and standardized differences in the propensity score models for the subpopulation analyses were also compared. Among the mental health and behavioral health subpopulations, we found similar differences across racial categories, baseline utilization, and location of the triggering event. After weighting the comparison group, these differences were mitigated and similar to the overall standardized differences. Similar findings were observed among children and adult subpopulations.

A-2.1.7 Statistical analysis

Regression model

The underlying assumption in the D-in-D models used to estimate the impact of the Arkansas URI EOC model is that trends in the Arkansas URI EOC group would be similar to those in the comparison group in the absence of the initiative (i.e., that the two were on "parallel paths" prior to the start of the URI EOC).

To assess the parallel assumption's validity more empirically, we modeled core utilization outcomes during the baseline period with a linear time trend interacted with a dichotomous variable indicating that the URI episode occurred in Arkansas. The following section describes the baseline analysis conducted to inform the D-in-D model.

Table A-2-2. Unweighted and weighted means and standardized differences, Arkansas URI Episode of Care and comparison groups, 2011

Characteristic	Unweighted			Weighted			
	Arkansas URI EOC group	Comparison group	Standardized difference[a]	Arkansas URI EOC group	Comparison group	Standardized difference[a]	p-value
N	112,776	346,109		112,776	111,360		
Individual-level sociodemographic characteristics							
Age	8.8	11.0	20.6	8.8	9.2	4.0	<0.001
Male (%)	45.7	42.2	7.2	45.7	46.0	0.5	0.22
Black (%)	13.5	31.9	45.1	13.5	13.9	1.1	0.01
Hispanic (%)	38.5	3.9	93.3	38.5	37.6	1.9	<0.001
White (%)	39.8	59.0	39.1	39.8	40.4	1.2	0.01
Other (%)	8.2	5.2	12.1	8.2	8.2	0.0	0.93
Health status characteristics							
CDPS score	1.0	0.9	4.2	1.0	1.1	4.2	<0.001
Medicaid eligibility: Disability (%)	11.5	8.7	9.3	11.5	12.3	2.6	<0.001
ED as triggering location (%)	8.8	16.4	23.1	8.8	9.2	1.5	<0.001
ED visit in previous year (%)	5.4	7.6	8.8	5.4	4.6	3.7	<0.001
Inpatient admission in previous year (%)	0.4	0.6	3.4	0.4	0.3	0.9	0.03
County-level characteristics							
Metropolitan status of county of residence (%)	55.3	44.7	21.3	55.3	46.4	17.8	<0.001
Percent of population at FPL, 2012	20.2	21.9	28.0	20.2	21.4	19.9	<0.001
Hospital beds per 1,000, 2010	3.7	4.2	18.2	3.7	4.2	17.6	<0.001
Median age, 2010	37.9	37.4	11.1	37.9	37.3	12.9	<0.001
Percent uninsured among under 65 years old, 2012	19.6	18.7	31.6	19.6	18.7	31.9	<0.001

CDPS = Chronic Illness and Disability Payment System; EOC = Episode of Care; ED = emergency department; FPL = federal poverty level; URI = upper respiratory infection.

[a] Absolute standardized differences are expressed as percentages. Differences below 10% are considered acceptable.

Table A-2-3. Unweighted and weighted means and standardized differences, Arkansas URI Episode of Care and comparison groups, 2012

Characteristic	Unweighted			Weighted			
	Arkansas URI EOC group	Comparison group	Standardized difference[a]	Arkansas URI EOC group	Comparison group	Standardized difference[a]	p-value
N	103,815	323,102		103,815	103,126		
Individual-level sociodemographic characteristics							
Age	8.9	11.3	22.4	8.9	9.0	0.4	0.40
Male	45.6	42.0	7.4	45.6	45.8	0.3	0.56
Black	13.7	32.9	46.6	13.7	13.8	0.5	0.30
Hispanic	35.6	4.2	85.6	35.6	35.1	1.1	0.01
White	41.4	57.6	32.9	41.4	41.8	0.8	0.07
Other	9.3	5.4	15.2	9.3	9.3	0.1	0.84
Health status characteristics							
CDPS score	1.0	0.9	6.3	1.0	1.2	5.1	<0.001
Medicaid eligibility: Disability	11.8	9.3	7.9	11.8	12.1	1.1	0.01
ED as triggering location	8.8	16.1	22.2	8.8	9.6	2.6	<0.001
ED visit, 2011	5.7	7.5	7.6	5.7	5.8	0.6	0.20
Inpatient admission, 2012	0.4	0.6	3.7	0.4	0.4	0.1	0.91
County-level characteristics							<0.001
Metropolitan status of county of residence	56.2	42.5	27.7	56.2	45.6	21.3	<0.001
Percent of population at FPL, 2012	20.1	22.2	33.6	20.1	21.3	19.7	<0.001
Hospital beds per 1,000, 2010	3.7	4.3	19.9	3.7	4.2	16.2	<0.001
Median age, 2010	37.9	37.4	12.2	37.9	37.4	12.8	<0.001
Percent uninsured among under 65 years old, 2012	19.6	18.7	32.3	19.6	18.6	36.6	<0.001

CDPS = Chronic Illness and Disability Payment System; ED = emergency department; FPL = federal poverty level; URI = upper respiratory infection.

[a] Absolute standardized differences are expressed as percentages. Differences below 10% are considered acceptable.

Table A-2-4. Unweighted and weighted means and standardized differences, Arkansas URI Episode of Care and comparison groups, 2013

Characteristic	Unweighted			Weighted			
	Arkansas URI EOC group	Comparison group	Standardized difference[a]	Arkansas URI EOC	Comparison group	Standardized difference[a]	p-value
N	98,868	326,356		98,868	98,862		
Individual-level sociodemographic characteristics							
Age	9.3	11.5	21.0	9.3	9.1	1.3	<0.001
Male	45.3	42.2	6.3	45.3	45.4	0.2	0.69
Black	13.8	33.1	46.7	13.8	13.9	0.2	0.66
White	33.5	4.5	79.7	33.5	33.4	0.3	0.53
Hispanic	42.3	56.7	29.1	42.3	42.4	0.2	0.66
Other	10.4	5.7	17.1	10.4	10.3	0.1	0.80
Health status characteristics							
CDPS score	0.8	0.8	6.6	0.8	1.0	6.7	<0.001
Medicaid eligibility: Disability	12.2	9.4	8.8	12.2	12.3	0.3	0.46
ED as triggering location	9.8	16.7	20.6	9.8	10.9	3.7	<0.001
ED visit, 2011	4.6	6.7	9.1	4.6	5.0	1.7	<0.001
Inpatient admission, 2012	0.3	0.5	3.1	0.3	0.4	1.2	0.01
County-level characteristics							
Metropolitan status of county of residence	56.4	43.1	26.8	56.4	46.6	19.6	<0.001
Percent of population at FPL, 2012	20.0	22.2	36.1	20.0	21.3	21.4	<0.001
Hospital beds per 1,000, 2010	3.7	4.4	21.2	3.7	4.2	17.7	<0.001
Median age, 2010	37.8	37.3	12.0	37.8	37.3	11.9	<0.001
Percent uninsured among under 65 years old, 2012	19.7	18.7	33.3	19.7	18.5	38.4	<0.001

CDPS = Chronic Illness and Disability Payment System; ED = emergency department; FPL = federal poverty level; URI = upper respiratory infection.

[a] Absolute standardized differences are expressed as percentages. Differences below 10% are considered acceptable.

Table A-2-5. Unweighted and weighted means and standardized differences, Arkansas URI Episode of Care and comparison groups, 2014

Characteristic	Unweighted			Weighted			
	Arkansas URI EOC group	Comparison group	Standardized difference[a]	Arkansas URI EOC group	Comparison group	Standardized difference[a]	p-value
N	87,942	282,995		87,942	87,810		
Individual-level sociodemographic characteristics							
Age	9.5	11.6	19.6	9.5	9.3	1.6	<0.001
Male	44.7	41.9	5.7	44.7	45.2	0.9	0.06
Black	13.2	34.9	52.4	13.2	13.3	0.3	0.51
Hispanic	31.7	4.6	75.3	31.7	31.4	0.7	0.16
White	43.9	55.1	22.6	43.9	44.1	0.4	0.43
Other	11.2	5.5	20.8	11.2	11.2	0.1	0.89
Health status characteristics							
CDPS score	0.8	0.7	6.5	0.8	1.0	5.3	<0.001
Medicaid eligibility: Disability	11.8	9.6	7.0	11.8	11.8	0.1	0.90
ED as triggering location	10.4	16.7	18.6	10.4	11.6	4.0	<0.001
ED visit, 2011	4.1	6.1	9.1	4.1	4.1	0.3	0.50
Inpatient admission, 2012	0.3	0.5	3.9	0.3	0.3	0.5	0.27
County-level characteristics							
Metropolitan status of county of residence	56.7	41.5	30.9	56.7	45.6	22.3	<0.001
Percent of population at FPL, 2012	19.8	22.6	44.2	19.8	21.4	25.8	<0.001
Hospital beds per 1,000, 2010	3.6	4.5	29.2	3.6	4.3	24.1	<0.001
Median age, 2010	37.7	37.2	11.1	37.7	37.3	8.7	<0.001
Percent uninsured among under 65 years old, 2012	19.6	18.9	28.1	19.6	18.6	35.7	<0.001

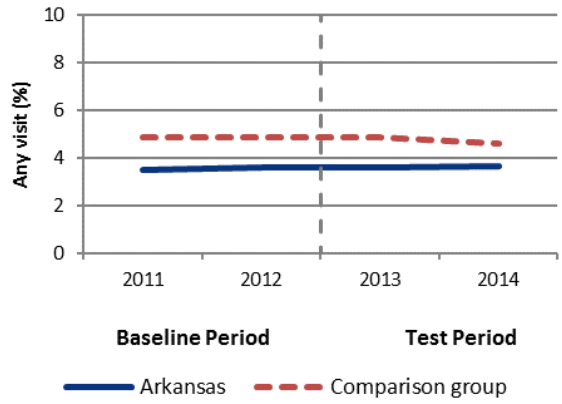
CDPS = Chronic Illness and Disability Payment System; ED = emergency department; FPL = federal poverty level; URI = upper respiratory infection.

[a] Absolute standardized differences are expressed as percentages. Differences below 10% are considered acceptable.

To examine descriptively whether the trends in the Arkansas URI EOC and comparison groups are parallel, graphs of annual unadjusted averages for the Arkansas URI EOC and comparison groups for the baseline period (2011–2012) and the first 2 years of implementation (2013–2014) are presented.

Figures A-2-5 through ***A-2-9*** provide the unadjusted averages of the utilization measures by year.

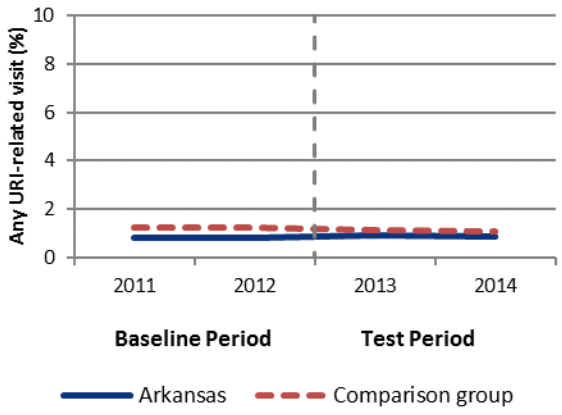
Figure A-2-5. Percentage of Medicaid URI episodes with an ED visit, FY 2011–FY 2014, Arkansas URI Episode of Care and comparison groups



The percentage of EOCs with an **ED visit** was consistently higher among the comparison group than among Arkansas EOCs across the study period. There was an observable decline in the percentage among the comparison group in 2014 (***Figure A-2-5***).

EOC = Episode of Care; ED = emergency department; FY = fiscal year; URI = upper respiratory Infection.

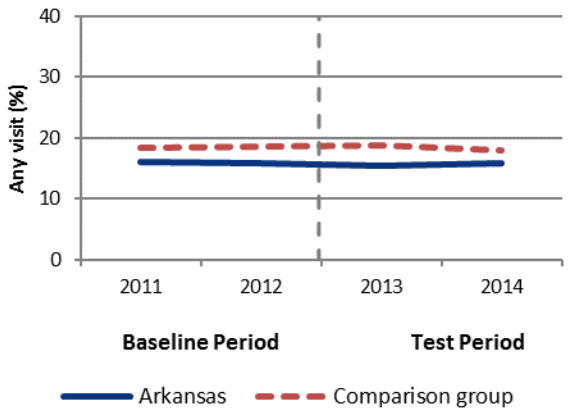
Figure A-2-6. Percentage of Medicaid URI episodes with a URI-related ED visit, FY 2011–FY 2014, Arkansas URI Episode of Care and comparison groups



The percentage of EOCs with a **URI-related ED visit** was consistently higher among the comparison group than among Arkansas EOCs across the study period, with an observable convergence between the two groups in the post period (***Figure A-2-6***).

EOC = Episode of Care; ED = emergency department; FY = fiscal year; URI = upper respiratory Infection.

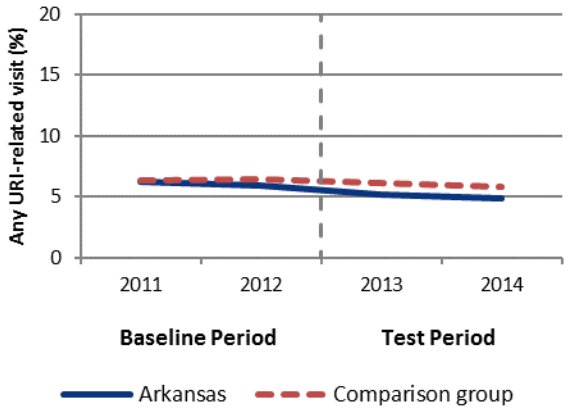
Figure A-2-7. Percentage of Medicaid URI episodes with a physician visit, FY 2011–FY 2014, Arkansas URI Episode of Care and comparison groups



The percentage of EOC with a **physician visit** was consistently higher among the comparison group than among Arkansas EOCs across the study period. The percentage remained mostly unchanged across the study period for both groups (***Figure A-2-7***).

EOC = Episode of Care; FY = fiscal year; URI = upper respiratory Infection.

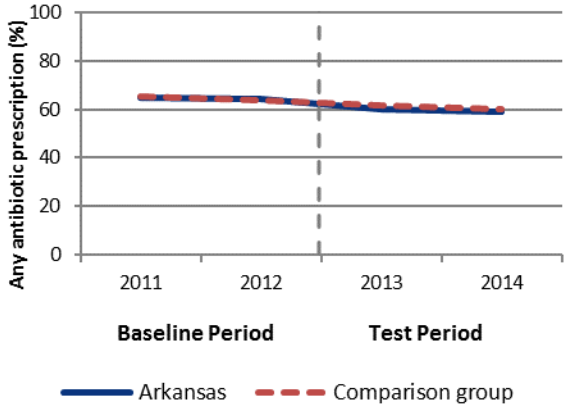
Figure A-2-8. Percentage of Medicaid URI episodes with a URI-related physician visit, FY 2011–FY 2014, Arkansas URI Episode of Care and comparison groups



The percentages of EOCs with a **URI-related physician visit** were similar among Arkansas EOCs and the comparison group in the pre period and declined for both groups across the study period, with an observably larger decline among Arkansas EOCs in the post period (***Figure A-2-8***).

EOC = Episode of Care; FY = fiscal year; URI = upper respiratory Infection.

Figure A-2-9. Percentage of Medicaid URI episodes with an antibiotic prescription, FY 2011–FY 2014, Arkansas URI Episode of Care and comparison groups

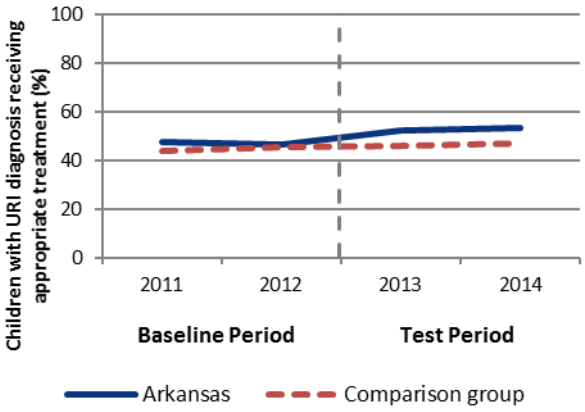


The percentages of EOC with an **antibiotic prescription** were similar among Arkansas EOCs and the comparison group, with both groups exhibiting similar declines throughout the study period (***Figure A-2-9***).

EOC = Episode of Care; FY = fiscal year; URI = upper respiratory Infection.

Figures A-2-10 and ***A-2-11*** provide the unadjusted averages of the quality of care measures by year.

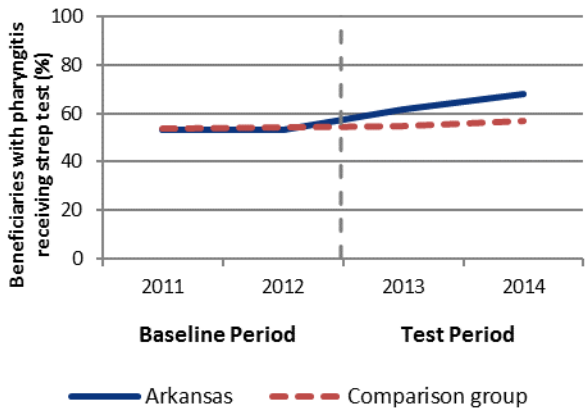
Figure A-2-10. Percentage of children with URI episodes receiving appropriate treatment, FY 2011–FY 2014, Arkansas URI Episode of Care and comparison groups



The percentages of children with URI EOCs receiving **appropriate treatment** were similar among Arkansas EOC and the comparison group in the pre-period. While the percentage remained mostly unchanged in the comparison group, it increased noticeably among Arkansas EOCs during the post period (***Figure A-2-10***).

EOC = Episode of Care; FY = fiscal year; URI = upper respiratory Infection.

Figure A-2-11. Percentage URI episodes with pharyngitis receiving a strep test, FY 2011–FY 2014, Arkansas URI Episode of Care and comparison groups



The percentages of URI EOCs with pharyngitis receiving a **strep test** were similar among Arkansas EOC and the comparison group in the pre period. Although both groups' percentages increased during the post period, the increase was more pronounced among Arkansas EOCs (***Figure A-2-11***).

EOC = Episode of Care; FY = fiscal year; URI = upper respiratory Infection.

An annual fixed-effects model considered for the evaluation is shown in ***Equation A-2.1***:

$$\gamma = \alpha_0 + \alpha_1 I + \sum \beta_n Q_{n,b} + \sum \phi_t Q_{t,p} \bullet I + \delta X + \mu \qquad \text{(A-2.1)}$$

where

y = a performance measure (e.g., any antibiotic use per episode per year) for the i-th beneficiary in the j-th group (Arkansas URI EOC or comparison group), in period t (i,j,t subscripts suppressed).

I = a 0,1 indicator (0 = comparison group, 1 = Arkansas URI EOC group).

X = a vector of patient and demographic characteristics.

$Q_{n,b}$, $Q_{t,d}$ = 0,1 indicator of the n-th or t-th calendar year in the base (b) or post (p) period (n starts counting at first baseline period, whereas t starts with first EOC model year).

μ = error term.

The model in ***Equation A-2.1*** assumes that, except for an intercept difference α_1, the outcomes for episodes in the Arkansas URI EOC group and those in the comparison group followed a similar growth trend during the baseline period. We investigated whether the baseline period before the start of the Arkansas URI EOC satisfied the baseline trend assumptions of the D-in-D model in ***Equation A-2.1***—that is, whether the outcome trends for episodes in the Arkansas URI EOC and comparison groups were similar during this period.

To test the similarity of the baseline trends, a model with a linear trend during the baseline period was used. We tested whether this trend differed for Arkansas URI episodes

relative to comparison group episodes. Specifically, the model for the outcomes may be written as follows:

$$y = \alpha_0 + \alpha_1 I + \theta * t + \lambda I * t + \delta X + \mu. \quad \text{(A-2.2)}$$

In ***Equation A-2.2***, *y, I,* X, and μ are defined as in ***Equation A-2.1***. The variable *t* is linear time ranging from 1 to 2. The linear time trend in the comparison group is θ•*t*, whereas for Arkansas URI EOC group beneficiaries (*I*=1), it is $(\theta + \lambda) * t$. Hence, λ measures the difference in linear trends, and the *t*-statistic for this coefficient can be used to test the null hypothesis of equal trends (λ = 0). In other words, rejecting the null hypothesis would suggest that the assumption of equal trends underlying our outcome models is not met.

The parameters of ***Equation A-2.2*** were estimated using weighted least squares regression models for three key outcomes. The weights are the propensity score. For each outcome, we calculate the estimate and standard error of the difference between the baseline trend in the Arkansas URI EOC group and comparison group (λ).

Table A-2-6 shows estimates of the baseline trend differences for the following outcomes:

- Probability of an antibiotic prescription fill
- Probability of a physician visit
- Probability of an ED visit.

Table A-2-6. Differences in the average percent probability of utilization outcomes during the baseline period, treatment group beneficiaries and comparison group beneficiaries

Parameter estimate	Any ED use	Any physician visit	Any antibiotic use
Arkansas URI EOC–comparison group trend difference	0.2	−0.3	0.7*
	(0.2)	(0.3)	(0.4)

EOC = Episode of Care; ED = emergency department; URI = upper respiratory Infection.

Note: Probability estimates multiplied by 100 to represent percent probability.

Baseline is the period October 2011–September 2012. The trend (slope) is the year-to-year change in the outcome variable. Standard errors are given in parentheses. *p < 0.10; **p < 0.05; ***p < 0.01.

Relative to the comparison group, there was no statistically significant difference in the baseline trend for any ED use and any physician visit. There was a slight difference in the baseline trend of the probability of any antibiotic use at the 0.10 level. However, the magnitude of the estimate was very small (0.007), and the difference was statistically significant at a p < 0.10 level of significance only. Moreover, the descriptive graph in Figure A-2-9 shows that

the baseline trends for antibiotic use in Arkansas and the comparison group were almost identical. As such, we concluded that the URI EOC group in Arkansas had a trajectory similar to that of the comparison group prior to the URI EOC model, and thus, the parallel trend assumption of the D-in-D model was satisfied.

D-in-D regression model. The D-in-D model is shown in ***Equation A-2.3***. This model is an annual fixed effects model, as shown in ***Equation A-2.1***. As in ***Equation A-2.1***, Y_{ijt} is the outcome for episode *i* (Arkansas URI EOC or comparison group) in state *j* in year *t*, I_{ij} (=0,1) is an indicator equal to 1 if the episode is in the Arkansas URI EOC group and 0 if the episode is in its comparison group, *Qn* is a series of yearly dummies for the baseline period (Years 1 to 2), and Q_t is a series of yearly dummies for the post years (2013–2014). The interaction of the Arkansas URI EOC group indicator and Q_t (I_{ij}*Q_t) measures the difference in the pre-post change between the Arkansas URI EOC group and the comparison group.

$$Y_{ijt} = \alpha_0 + \beta_1 I_{ij} + \sum \beta_2 Q_n + \sum \alpha_2 Q_t + \sum \gamma I_{ij} * Q_t + \lambda X_{ijt} + \varepsilon_{ijt} \qquad \text{(A-2.3)}$$

Table A-2-7 illustrates the interpretation of the D-in-D estimate from this model. The coefficient β_1 in ***Equation A-2.3*** is the difference in the measure between Arkansas URI EOC episodes and comparison episodes at the start of the baseline period, holding constant other variables in the equation. The β_2 and α_2 coefficients are for the annual fixed effects and capture differences over time for each baseline and post year, respectively. The coefficient of the interaction term between Q_t and URI EOC (*I*) measures any differences for the Arkansas URI EOC group relative to the comparison group in the post years relative to the baseline years. Thus, in the post period, the comparison group mean is captured by $\alpha_0 + \alpha_2$, whereas the Arkansas URI EOC group mean is captured by $(\alpha_0 + \beta_1) + (\alpha_2 + \gamma)$. In other words, the between-group difference changes from β_1 during the baseline years to $\beta_1 + \gamma$ during the post period. The D-in-D parameter, γ, shows whether the between-group difference increased ($\gamma>0$) or decreased ($\gamma<0$) after the Arkansas URI EOC was implemented. Using the annual fixed effects model, overall estimates were calculated by taking linear combinations of the yearly estimates.

Table A-2-7. Difference-in-differences estimate

Group	Pre period	Post period	Pre-post difference
Arkansas URI EOC group	$\alpha_0 + \beta_1 + \beta_2$	$(\alpha_0 + \beta_1) + (\alpha_2 + \gamma)$	$\alpha_2 + \gamma$
Comparison group	$\alpha_0 + \beta_2$	$\alpha_0 + \alpha_2$	α_2
Between group	β_1	$\beta_1 + \gamma$	Γ

EOC = Episode of Care; URI = upper respiratory Infection.

All outcomes models were estimated with the beneficiary-episode as the level of analysis. We used weighted logistic regression models to report the estimated likelihood of the dependent variable and the marginal effect as a percent probability.

The models for overall ED visits, overall physician visits, and antibiotic use were ran for children and adults and for those with behavioral health conditions.

Pre-post regression model. We also examined URI-related Medicaid payments in Arkansas from 2011 to 2014. The comparison states have a high managed care penetration so we were not able to measure expenditures for comparison group episodes. Instead, we used the pre-post model shown in ***Equation A-2.4***. Y_{it} is the outcome for episode i in year t, Q_t is a series of yearly dummies for the post years (Years 4 to 5), and the baseline period is the reference category. The post years indicator Q_t measures the difference in the pre-post change among Arkansas URI episodes.

$$Y_{it} = a0 \quad + \beta_1 2013_i + \beta_2 2014_i + \mu_3 X_i + \varepsilon \qquad \text{(A-2.4)}$$

where β_{1-2} is the average marginal effect of each post-demonstration year on Y_i, relative to the baseline years (2011 and 2012); X represents person- and area-level characteristics; μ_3 is the coefficient associated with each of those characteristics; and ε is the error term.

Table A-2-8 illustrates the interpretation of the pre-post estimate from this model. The coefficient β1 in ***Equation A-2.4*** is the difference in the measure before and after URI EOC implementation in Arkansas, holding constant other variables in the equation. Using the annual fixed effects model, overall estimates were calculated by taking linear combinations of the yearly estimates.

Table A-2-8. Pre-post estimate

Group	Pre period	Post period	Pre-post difference
Arkansas URI EOC	$\alpha_0 + \beta_2$	$(\alpha_0 + \beta_2) + \alpha_2$	α_2

EOC = Episode of Care; URI = upper respiratory Infection.

Control variables. In all models we controlled for the following variables:

- Male (Ref: Female)
- Age (age and age squared)
- Black
- Hispanic
- Other
- White (Ref)
- Medicaid eligibility: Disabled
- ED as triggering location

- Chronic Illness and Disability Payment (CDPS) score (count of major comorbidities)
- Metropolitan status of county of residence
- Percent of population at FPL, 2012
- Number hospital beds per 1,000, 2010
- Median age, 2010
- Percent uninsured, ages <65, 2012.

Weighting and clustering. All the regression models were estimated weighted by the propensity score. In addition, standard errors were clustered at the individual episode level to account for correlation in the error term between multiple episodes among the same person.

Adjusted means. The regression-adjusted D-in-D estimate and the D-in-D calculated from regression-adjusted means will differ for one of two reasons. First, in nonlinear specifications the D-in-D calculated from the regression-adjusted means is known to be a biased estimator for the treatment effect. To address this bias, we use the nonlinear D-in-D approach described in Puhani (2012). In some cases the bias may be extreme, leading to substantial differences between the regression-adjusted D-in-D estimates versus the D-in-D calculated from regression-adjusted means.

Second, in linear specifications the D-in-D calculated from the overall regression-adjusted means may be substantially different than the overall regression-adjusted D-in-D estimate because we use different weights to obtain the overall figures. Specifically, the regression-adjusted D-in-D estimates are weighted using the number of treatment beneficiaries observed in each year relative to the total number of treatment beneficiaries ever observed during the test period. This is mathematically equivalent to weighting the test-period adjusted means for both groups with the same weights that are applied to the treatment group. However, the test-period adjusted means that are presented for the comparison group are weighted using the number of comparison beneficiaries observed in each year relative to the total number of comparison beneficiaries ever observed during the test period. The implication of this is that in cases where there are large differences in the rates of rolling entry or exit across the two groups, we may observe large differences in the D-in-D calculated from the overall regression-adjusted means versus the overall regression-adjusted D-in-D estimate.

A-2.2 Methods for the Arkansas Perinatal Episode Impact Analysis

To estimate the impact of the perinatal episode of care (EOC) model in Arkansas, we conducted an alternative specification of a difference-in-differences (D-in-D) regression analyses that allows for differences in baseline trends using Medicaid Analytic eXtract (MAX) and Alpha-MAX Medicaid claims from 2010 through 2014. In ***Appendix A, Section A.3***, we present D-in-D analyses for outcomes across three evaluation domains: (1) care coordination,

(2) service utilization, and (3) quality of care. Additionally, we conducted a pre-post trend analysis of Medicaid expenditures in Arkansas. This sub-appendix details the methods we used for this analysis.

The perinatal EOC model in the context of Arkansas Medicaid. Arkansas began introducing retrospective EOCs of several conditions, including perinatal episodes, to encourage better management of care and promotion of quality during the perinatal period. An EOC is a collection of care provided to treat an illness or condition over a fixed time window. Providers submit claims and pay for the service provided, but after each performance period the provider is assessed for their average costs, including all pregnancy-related costs, and performance on process measure (ADHS, 2016). The goal of the perinatal episode is to reduce complicated pregnancies and lower Arkansas Medicaid spending.

Arkansas Medicaid covered pregnant women up to 200 percent of the Federal Poverty Level (FPL) from 2009 through 2013, with a small uptick to 208 percent FPL in 2014 resulting from the adoption of MAGI (modified adjusted gross income) to determine Medicaid eligibility nationwide under the Affordable Care Act. Post 2014, pregnant women must meet the 208 percent income limit and not qualify for other pregnancy related coverage. These women are eligible for prenatal, delivery, and postpartum care and for other care for conditions that would complicate the pregnancy. Pregnant women can also be eligible for full benefits through Arkansas's low income pregnant women coverage. The income cut off for this group was 16–17 percent of the FPL from 2009 to 2013 (KFF, 2018).

Starting in 2014 under the private option, Arkansas adults with incomes up to 133 percent of the FPL and pregnant women with incomes up to 209 percent of the FPL have been able to receive premium assistance to purchase private coverage from the health insurance marketplace. Therefore, in 2014 and 2015, many women who would have received maternity benefits covered under traditional Medicaid would instead receive them under the private option. Compared to the 2011–2013 period, we would expect private option women to have higher income on average than those in traditional Medicaid. We would also expect the percentage of all births financed by Medicaid to decline as a result. Further, the composition of women with traditional Medicaid maternity benefits would change—we would expect these women to have less income on average (because the higher income women are now enrolled in the private option) and, consequently, to be in poorer health and have increased use of services.

Profile of participating providers. Participation in the Perinatal EOC was compulsory for physicians performing deliveries in Arkansas starting in FY 2013. The principal accountable provider (PAP) is the physician or physician group that performs the delivery. By 2015, there were 139 PAPs that performed five or more deliveries during the year and were subject to the requirements of the episode payment model. As described in the introduction, 23 percent of

PAPs had what was considered commendable average perinatal expenditures in 2012. By 2014, 45.3 percent of PAPs had commendable performance (ADHS, 2016).

Intervention group. Arkansas introduced perinatal EOCs in July 2012, but providers were not subject to up- or downside risk until October 2012. The intervention group for this analysis are Medicaid-financed deliveries in Arkansas from October 2010 to September 2014 that met exclusion criteria and resulted in a live single birth. The baseline period consists of October 2010 to September 2012, and the post-period was October 2012 to September 2014.

Comparison group. Our comparison group consisted of Medicaid covered deliveries with a live birth that occurred in an inpatient setting in Mississippi and Missouri. We selected these two states as comparison states based on availability of data, calculated Euclidean distance scores, similarity in the percentage of pregnancies covered by Medicaid, availability of MAX claims, and similar income thresholds for Medicaid eligibility (KFF, 2018).

Study sample. Our sample included Medicaid-covered deliveries with a live birth that occurred in an inpatient setting from fiscal year 2011 through 2014. We exclude beneficiaries with different types of coverage during the episode, including Children's Health Insurance Program, supplemental private insurance coverage, or dual Medicare-Medicaid enrollment. To mimic the Arkansas episode criteria, we removed beneficiaries with certain comorbidities and those with pregnancy-related conditions. We also excluded those with limited enrollment, overlapping episodes, no claim for any prenatal care, episodes without full benefits during the delivery month and for at least 6 months prior to delivery, and episodes among adolescents 15 years and younger. We further restrict the sample to those with full Medicaid eligibility up to 60 days post-delivery for postpartum outcome measures.

Balancing perinatal EOC and comparison group. Following comparison group selection, we constructed an episode-level propensity score weight to balance the Arkansas perinatal EOC and comparison group on select observed individual characteristics. We used weighting to apply propensity scores to the analysis, as opposed to other methods like matching, to retain sample size and produce less biased estimates for binary outcomes. After propensity score weighting, the standardized differences between the weighted comparison group means and the Arkansas perinatal EOC group means were under the standard 10 percent threshold, with exception to some area-level characteristics. More information on propensity score weighting is available in ***Section A-2.2.4***.

Study design. We used an alternative D-in-D design, comparing changes in the outcome variables before and during the implementation of the perinatal EOC model in the Arkansas perinatal EOC group with changes in the outcomes before and during the implementation of the perinatal EOC model for the comparison group allowing the baseline trends to differ between the

two groups. We used a pre-post multivariate regression analysis to examine perinatal related expenditures. We used retrospective annual cross-sections of perinatal EOCs from 2011 to 2014.

Statistical approach. Analyses used ordinary least squares (OLS) for expenditure outcomes and logistic and negative binomial regression for utilization and quality outcomes. All regression analyses used clustered standard errors at the individual level to account for clustering of multiple episodes for a person. The outcome models controlled for age, gender, race, disability-related Medicaid eligibility, poverty-related Medicaid eligibility, health status, urban/rural area of residence, and county-level characteristics. More information on the study outcomes is available in ***Section A-2.2.2***, and more information on the regression model is available in ***Section A-2.2.7***.

A-2.2.1 Data sources

Medicaid MAX data. The RTI evaluation team used Medicaid data from the CMS MAX and Alpha-MAX research files made available through the Chronic Condition Warehouse (CCW) enclave. Each state's Medicaid Statistical Information System (MSIS) data are the source of the MAX and Alpha-MAX files. The MAX processing adds enhancements such as claims adjustments, creation of a national type of service field, and state-specific quality issues corrections; Alpha-MAX provides fewer enhancements. The MAX and Alpha-MAX files include a person summary (PS) file, with all enrollment information and summary claims information and four claims files: inpatient hospital (IP), long-term care (LT), prescription drugs (RX), and other (OT) claims. The quarterly Alpha-MAX files are generated for a state once all five MSIS file types for a single quarter are approved. The quarterly files are overwritten and updated each time a new quarter of run-out data is added. Quarterly versions of Alpha-MAX are being produced for each state through 7 quarters of run-out data; therefore, the quarterly files are based on 0 to 7 quarters of run-out time. Annual calendar-year MAX files are prepared from data with 7 quarters of run-out time. For simplicity, we refer to the MAX and Alpha-MAX data as simply MAX data for the remainder of this appendix

Area Health and Resource File (AHRF). The AHRF comprises data collected by the Health Resources and Services Administration from more than 50 sources containing more than 6,000 variables related to health care access at the county level. We used information on health professions supply, poverty at the county level, age, rural/urban status, and uninsured rates from 2010–2015 to select the comparison group and to use as covariates in the analysis.

A-2.2.2 Outcome measures

Utilization

Utilization measures are reported as a probability of any utilization within the episode, with exception to total episode emergency department (ED) visits, prenatal ED visits, and length

of stay. For count measures we estimated count models to capture the estimated number of events during the episode, and for length of stay we estimated the expected average using OLS.

- **Number of ED visits during prenatal period:** This is a count of the number of visits to the ED that occurred during the prenatal period divided by the number of beneficiaries in the same period. ED visits are identified as any OT claims with a revenue code equal to 0450–0459, 0981, or 0762 or a claim where the place of service code equals 24 and procedure codes equal 99281, 99282, 99283, 99284, or 99285. We exclude claims where the procedure code always equals 70000 through 79999 or 80000 through 89999 and the revenue code never equals 762.
- **Probability of any ED visits during postpartum period**: An indicator for any ED visit, defined above, within the postnatal period: date of delivery plus 60 days.
- **Number of total ED visits during the perinatal episode**: This is a count of the number of ED visits, defined above, during the entire episode.
- **Probability of having any inpatient use during prenatal period**: This is an indicator of whether the beneficiary had at least one admission during the prenatal period (episode start date to the date of admission for the delivery).
- **Probability of having any 30-day readmission.** This is an indicator for whether there was any readmission during the postnatal period, 30 days post-delivery.
- **Probability of having any 60-day readmission**. This is an indicator for whether there was any readmission during the postnatal period, 60 days post-delivery.
- **Number of days inpatient for the delivery admission**. This is a continuous number of days in which the beneficiary resided in the inpatient facility during the delivery. This was calculated as the service end date on an IP claim minus the admission date on the same line among claims where the delivery date was greater than or equal to the admission date and less than or equal to the service end date.

Quality of care

To evaluate the impact on quality of care, we report the following quality measures. The measures were calculated as binary indicators for having any procedures during the episode. We consider these quality of care of measures because they are routine screening measures to prevent pregnancy-related complications during the prenatal period.

- **Probability of any HIV screening**. This is an indicator for whether the beneficiary was screened for HIV during the prenatal period. Procedure codes from OT claims included 80055, 84181, 84182, 86701, 86702, 86703, 87300, 87390, 87391, 87534, 87535, 87536, 87537, 87538, and 87539. Diagnosis codes include 042.
- **Probability of any chlamydia screening**. This is an indicator for whether the beneficiary had any screening for chlamydia during the prenatal period. Procedure codes from OT claims included 3511F, 87110, 87270, 87320, 87451, 87490, 87491, 87492, 87797, 87798, 87799, 87800, 87801, and 87810.

- **Probability of group B streptococcus screening**. This is an indicator for whether the beneficiary had any screening for group B strep during the prenatal period. Procedure codes from OT claims included 86403, 87070, 87071, 87075, 87077, 87081, 87147, 87149, 87449, 87653, 87797, 87798, 87799, 87800, 87801, and 87802.
- **Probability of any screening for gestational diabetes**. This is an indicator for whether the beneficiary had any screening for gestational diabetes during the prenatal period. Diagnosis codes from OT claims were 250.xx and procedure codes were 82947, 82950, 82951, and 82952.
- **Probability of any screening for asymptomatic bacteriuria (urinary tract infection; UTI).** This is an indicator for whether the beneficiary was screened for a UTI during the prenatal period. Procedure codes from OT claims included 81000, 81001, 81002, 81003, 81005, 81007, 81015, 87077, 87081, 87086, 87088, 87149, 87152, and P7001.
- **Probability of any hepatitis B screening.** This is an indicator for whether the beneficiary was screened for hepatitis B during the prenatal period. Procedure codes from the OT claims included 80055, 80074, 86704, 86705, 86706, 86707, 87340, 87341, 87350, 87515, 87516, and 87517.
- **Probability of a caesarian section.** This is an indicator for whether the beneficiary had a Cesarean section. This indicator was flagged if any OT record or at least one IP record within the episode window had a CPT procedure code of 59510, 59514, 59515 59618, 59620, or 59622 or an ICD-9 procedure code of 74, 740, 741, 742, 744, or 7499. It is important to note that it is impossible to measure necessary versus elective caesarian sections from claims administrative data.

Medicaid payments

This study does not evaluate the effect of the perinatal EOC model on Medicaid payments between Arkansas and the comparison group because of incomparability across datasets. However, to understand trends within the Arkansas Medicaid program, we examine changes in payment over time.

Total perinatal-related inpatient and other services payments. Sum all OT and IP payments from the beginning until the end of the episode. MAX_TOS = 1–19, 23–54, and 99 and TYPE_CLM_CD = 1 OR 5 and DIAG_CD = any of the ICD-9 codes identified in Arkansas's original EOC specification (Arkansas Medicaid, n.d.). This includes the professional services and inpatient services delivered during the inpatient stay of the delivery.

A-2.2.3 Population studied

Our sample included Medicaid-covered deliveries with a live birth that occurred in an inpatient setting from fiscal year 2011 through 2014 (N = 272,879). We exclude beneficiaries with different types of coverage during the episode, including Children's Health Insurance Program, supplemental private insurance coverage, or dual Medicare-Medicaid enrollment

(N = 39,289). To more closely resemble the Arkansas episode criteria, we removed beneficiaries with claims indicating one or more of the following comorbidities within 280 days (episode start date) prior to the delivery date: cancer, cystic fibrosis, congenital cardiovascular disorders, DVT/pulmonary embolism, other phlebitis and thrombosis, end-stage renal disease, sickle cell anemia, or type 1 diabetes (n = 12,039). Additionally, we excluded those with pregnancy-related conditions such as amniotic fluid embolism, obstetric blood clot embolism, placenta previa, severe preeclampsia, multiple gestation >= 3, late-effect complications of pregnancy/childbirth, puerperal sepsis, suspected damage to fetus from viral disease in mother, or cerebrovascular disorder (n = 24,799). Finally, we excluded those with limited enrollment (n = 41,707), overlapping episodes (n = 2,913), no claim for any prenatal care (n = 18,325), episodes without full benefits during the delivery month (n = 3,183) and for at least 6 months prior to delivery (n = 83,560), and episodes among adolescents 15 years and younger (n = 2,436). Our final sample size consisted of 148,872 episodes. We further restrict the sample to those with full Medicaid eligibility up to 60 days post-delivery for postpartum outcome measures, which reduced the sample to 141,645 deliveries.

Our exclusion criteria remove deliveries that may be more complex or high cost and deliveries where no prenatal visit was observed, indicating either limited Medicaid eligibility or poor access to care. These deliveries are not subject to the episode payment criteria; however, it is worth examining whether our model results are robust to the inclusion of all deliveries in Arkansas. As a sensitivity analysis, we estimated our outcome models on the entire population of deliveries.

Intervention group

Perinatal episodes are triggered by a delivery procedure code and span the time period from the date of 280 days before delivery to 60 days following delivery. Thus, identification of the delivery date is a key factor in identifying claims for the perinatal episodes. The logic and code to be used for identifying delivery dates in the claims files can be found below.

The description of the perinatal algorithm suggests that CPT delivery procedure codes on any claim record and MS-DRGs for delivery in inpatient records trigger an episode. However, Arkansas MAX files do not include DRGs and some hospitals in Arkansas and comparison states may use ICD-9 procedure codes and CPT delivery procedure codes. Therefore, we also search for ICD-9 procedure codes in the IP claims file.

Identification of perinatal episode

We use both the inpatient and outpatient files to search for delivery procedure codes. A claim that indicates either vaginal or cesarean deliveries identifies a potential episode. The procedure codes for vaginal delivery include 59400, 59409, 59410, 59610, 59612, and 59614 (CPT) or 72, 720, 721, 722, 723, 7221, 7229, 7231, 7239, 724, 725, 7251–7254, 726, 727, 7271,

7279, 728, 729, 7322, 735, 7351, or 7359 (ICD-9). Caesarian section codes include 59510, 59514, 59515, 59618, 59620, and 59622 (CPT) or 74, 740, 741, 742, 744, or 7499 (ICD-9).

We create a single record for each unique Medicaid beneficiary and delivery date. The delivery date is identified with the following logic: If there is only one record for a unique beneficiary then the delivery date is the service end date or principle procedure date on the claim. The admission date is used if the principle procedure date is missing.

If there are multiple records for a unique beneficiary and they all have the same procedure service date then the delivery date is the service begin date on the claim.

If there are multiple records for a unique beneficiary and the service begin date is less than 6 months from the previous service end date, then we apply the following logic:

- Delivery date = principal procedure date for the latest IP record with any of the nine diagnosis codes = V270, V272, V273, V275, V276, 650, or 640.xx–6799x with a fifth digit of 1 or 2;
- Else delivery date = service end date for the latest OT record where the procedure modifier = AA (obstetrician or surgeon performing the delivery) and any of the nine diagnosis codes = V270, V272, V273, V275, V276, 650, or 640xx–679xx with a fifth digit of 1 or 2;
- Else delivery date = service end date for the latest OT record where any of the two diagnosis codes = V270, V272, V273, V275, V276, 650, or 640xx–6799x with a fifth digit of 1 or 2;
- Else delivery date = principal procedure date for the latest IP record;
- Else delivery date = service end date for the latest OT record.

Among members where the delivery date was identified using professional claims, we assigned an **admission date** and **discharge date** to the delivery based on a corresponding claim in the inpatient file where the admission date on the facility claim was less than or equal to the service end date assigned on the OT claims, and the discharge date was greater than or equal to the service end date.

The episode begin date was defined as the delivery date minus 280, and the episode end date was defined as the delivery date plus 60 days.

A-2.2.4 Comparison group and propensity score weighting

For the impact analysis, we are using a pre-post comparison group design, in which the comparison group provides an estimate of what would have happened among Arkansas perinatal episodes absent the EOC payment model. The difference in the changes over time from the pre-period to the intervention period between Arkansas episodes and their comparison group

provides an estimate of the impact of the perinatal EOC model. The comparison group should be similar to the intervention group on all relevant dimensions (e.g., demographic, socioeconomic, political, regulatory, and health and health systems) except for the policy change being tested.

In the following section, we detail the procedures we used to select the comparison group for Arkansas perinatal episodes.

Selection of comparison group

Our comparison group consisted of Medicaid covered deliveries with a live birth that occurred in an inpatient setting in Mississippi and Missouri. We selected these two states as comparisons states based on calculated Euclidean distance scores based on 25 state characteristics, similarity in the percent of pregnancies covered by Medicaid, availability of MAX claims, and similar income thresholds for Medicaid eligibility (KFF, 2018).

Availability of MAX claims was a key determinant in our selection of these states. Other Fee-for-services Medicaid states did not have enough post-years of MAX claims to conduct this analysis. Mississippi and Missouri both have MAX claims extending beyond the availability of claims in Arkansas, allowing this analysis to have 2-years of post-episode of care implementation. A limitation to this selection, however, is that Missouri primarily relies on comprehensive managed care plans to serve their Medicaid beneficiaries. Similarly, starting in 2013, Medicaid financed deliveries were primarily financed through managed care plans in Mississippi. As such, we could not compare Medicaid FFS payments during the episodes of care in Arkansas relative to the comparison group.

This analysis relies on the completeness of inpatient and professional encounters for these episodes of care. A relatively recent encounter analysis by Mathematica Policy Research indicated that in both states, 2010 and 2011 inpatient and professional claims were complete and usable for health services research. More recent analysis is not available; however, we found that episodes had similar inpatient, professional, and drug encounter utilization within each state for each year. This indicates that encounter submission for MO and MS from 2012 to 2014 is reasonably complete for this analysis.

These states have Medicaid programs that have similar income cut-off criteria. From 2009 through 2013, Mississippi and Missouri, all had an income cutoff limits for pregnant women of 185 percent of FPL but had slightly higher cutoffs because of MAGI adoption in 2014 and 2015, close to Arkansas's cut off of 200 percent. The Medicaid income eligibility limits for other categories such as parents for Missouri was 23–37 percent of FPL, followed by Mississippi with 28–44 percent of FPL, as compared to 16–17 percent of FPL in Arkansas from 2010 to 2013, and 138 percent of FPL in 2014.

Calculation of person-level weights

To balance the population characteristics for the claims-based analyses, we estimated propensity scores for all individuals from the comparison group. A propensity score is the probability that an individual is in the intervention group rather than the comparison group.

The objective of propensity score modeling is to create a weighted comparison group with characteristics equivalent to those for the Arkansas perinatal EOC population. To the extent that these characteristics are correlated with expenditure, utilization, and quality outcomes, propensity weighting will help balance pre-intervention levels of the outcomes as well.

There are other methods to apply propensity scores to an analysis. Aside from weighting, one frequently used method is matching, whereby an intervention beneficiary is matched to a comparison group beneficiary who has a similar propensity score. Although we considered this method, we decided not to pursue matching for several reasons. First, propensity score weighting has been shown to produce less biased estimates, less modeling error (e.g., mean squared error, type 1 error), and more accurate variance estimation and confidence intervals when modeling dichotomous outcomes; and this analysis includes many dichotomous utilization and quality of care outcomes. Second, matching may exclude many comparison group beneficiaries from the analysis if a good match cannot be found. Weighting has the advantage of preserving sample size.

Person-level characteristics

The initial step in the process was to select person-level characteristics to be used in each propensity score model. ***Table A-2-9*** shows the characteristics we used grouped by whether they control for demographics, enrollment, attribution, or beneficiary health status. Because there is limited information available in claims data, we considered also including county-level characteristics to control for geographic characteristics such as physician supply and median income to account for potential differences in access to care or other geographic differences. However, we found that there was little variation in county-level characteristics, which made it difficult to balance on these variables. To optimize the balance and to avoid extreme weights, we therefore excluded county-level covariates from the propensity score model. However, we do control for county-level characteristics in the outcome model.

Table A-2-9. Covariates for propensity score logistic regressions

Covariate	Definition
Demographic characteristics	
Age indicator: 16 to 19	Dichotomous
Age indicator: 20 to 24	Dichotomous
Age indicator: 25 to 34	Dichotomous
Age indicator: 35 and older (Referent)	Dichotomous
Black	Dichotomous
Hispanic	Dichotomous
Other	Dichotomous
White (Referent)	Dichotomous
Enrollment	
Poverty-related eligibility	Dichotomous indicator for Medicaid eligibility during the month of the delivery.
Months of full-Medicaid enrollment during prenatal period	Continuous count of the number months the beneficiary had full Medicaid benefits during the prenatal period.
Health status measures	
Chronic Illness and Disability Payment score	Continuous
Disability	Dichotomous indicator for having any month during episode where beneficiary was eligibility for disability-related Medicaid.
Diabetes	Dichotomous indicator for having any diabetes diagnosis in the year prior to the delivery date.
Asthma	Dichotomous indicator for having any asthma diagnosis in the year prior to the delivery date.
Hypertension	Dichotomous indicator for having any hypertension diagnosis in the year prior to the delivery date.
Inpatient admission during previous year	Dichotomous for having any inpatient admission during year prior to the delivery date.
Emergency department visit during previous year	Dichotomous for having any emergency department visit during year prior to the delivery date.

Estimation and weighting procedures

Using the characteristics listed in ***Table A-2-9***, we estimated propensity models by logistic regression, in which the outcome was 1 for beneficiaries attributed to a Medicaid ACO provider and 0 for the comparison group. Separate models were estimated for 2011, 2012, 2013, 2014, and 2015 data.

We set propensity weights to 1 for all individuals in the intervention group. The propensity weight for a comparison individual was a function of his or her predicted propensity score—where weight = p/(1-p), with p the predicted propensity. Our procedure typically includes

trimming weights that are either less than 0.05 or greater than 20, although in this analysis no weights needed trimming.

A-2.2.5 Propensity model evaluation

We evaluated several aspects of the propensity score models. First, we examined plots of predicted probabilities to ensure sufficient overlap in the distributions of the intervention and comparison groups. This feature, known as common support, is critical because it provides the basis for inferring effects from group comparisons (***Figure A-2-12*** to ***A-2-15***).

In all years, we found the comparison group passed the common support assumption ($P(D = 1|X)>0$) for almost the entire range of the intervention group's propensity scores. The only exceptions were in the uppermost percentiles of the intervention group's distribution (above the 99th percentile). These plots provide ample evidence that the common support assumption is upheld.

Second, we compared the logistic results of the models to see which variables had the greatest impact on the propensity score weights. We found that the major differences between the groups were individual characteristics such as race, disability, utilization of inpatient and ED visits during the previous year, CDPS scores, and poverty-related income eligibility for Medicaid. Overall, we found that pregnant women in Arkansas were more likely to be disabled, Hispanic, have income-related poverty Medicaid eligibility, lower CDPS scores, and less utilization during the baseline years. We found these differences to be fairly consistent and stable over time.

Finally, we compared unweighted and propensity-weighted means for the characteristics in the model. As expected, we found that, after weighting, the comparison group means were within a few percentage points of the values for the intervention group.

Figure A-2-12. Weighted and unweighted propensity score density plots for the Arkansas perinatal Episode of Care and comparison groups, 2011[2]

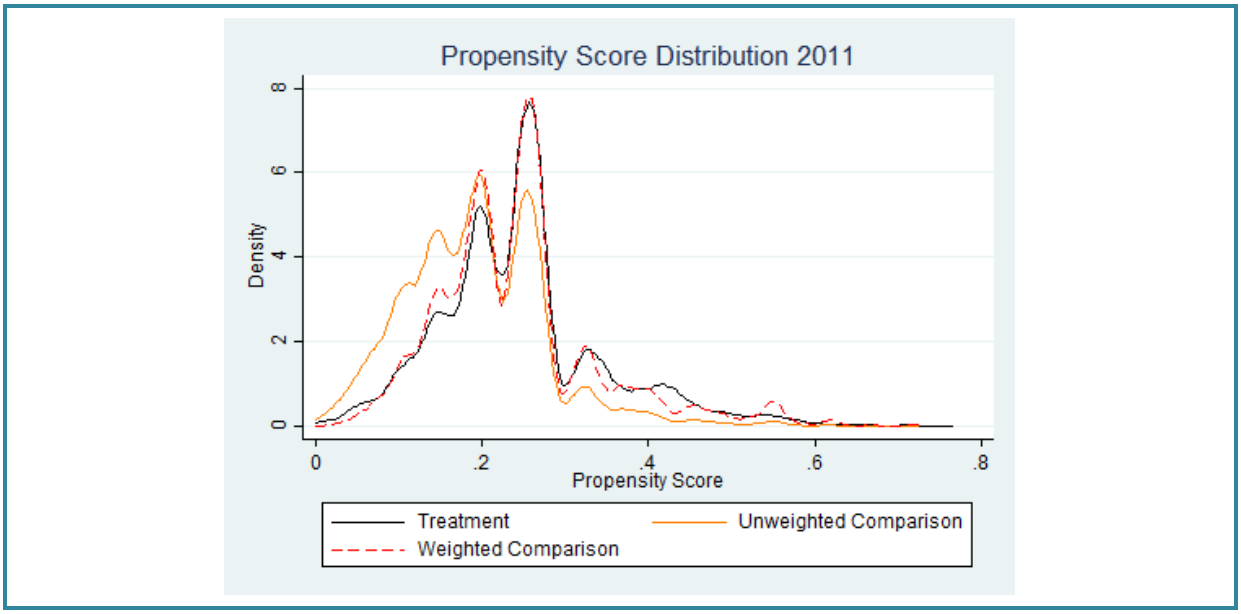
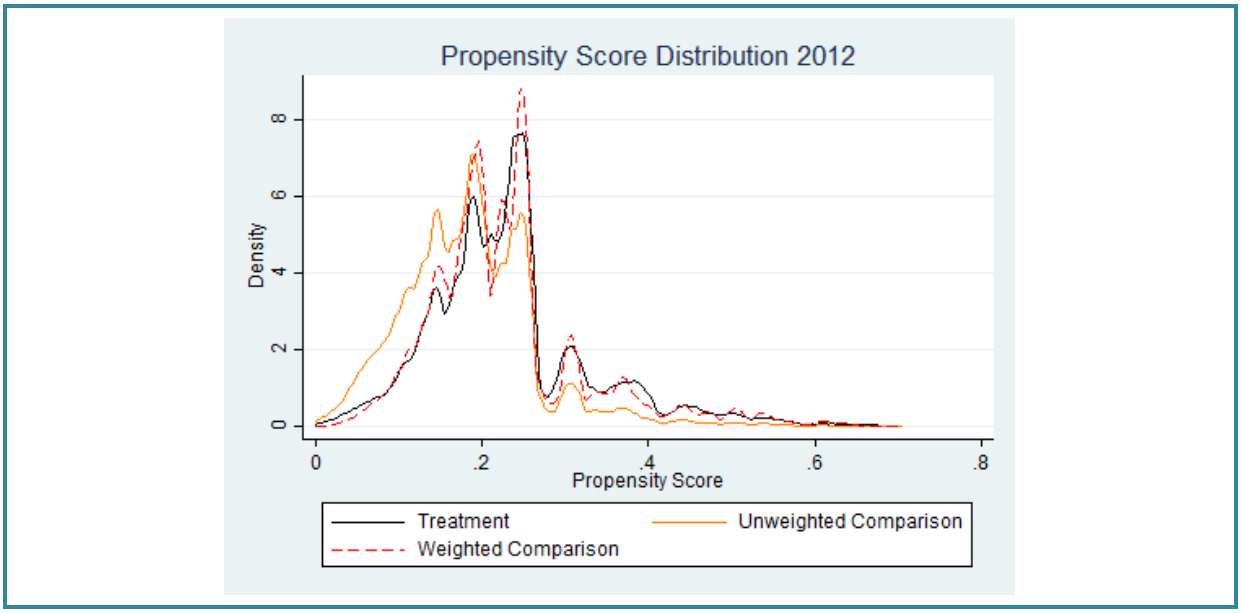


Figure A-2-13. Weighted and unweighted propensity score density plots for the Arkansas perinatal Episode of Care and comparison groups, 2012



[2] In ***Figures A-2-12*** through ***A-2-15***, the Treatment lines represent those in the Arkansas perinatal EOC group.

Figure A-2-14. Weighted and unweighted propensity score density plots for the Arkansas perinatal Episode of Care and comparison groups, 2013

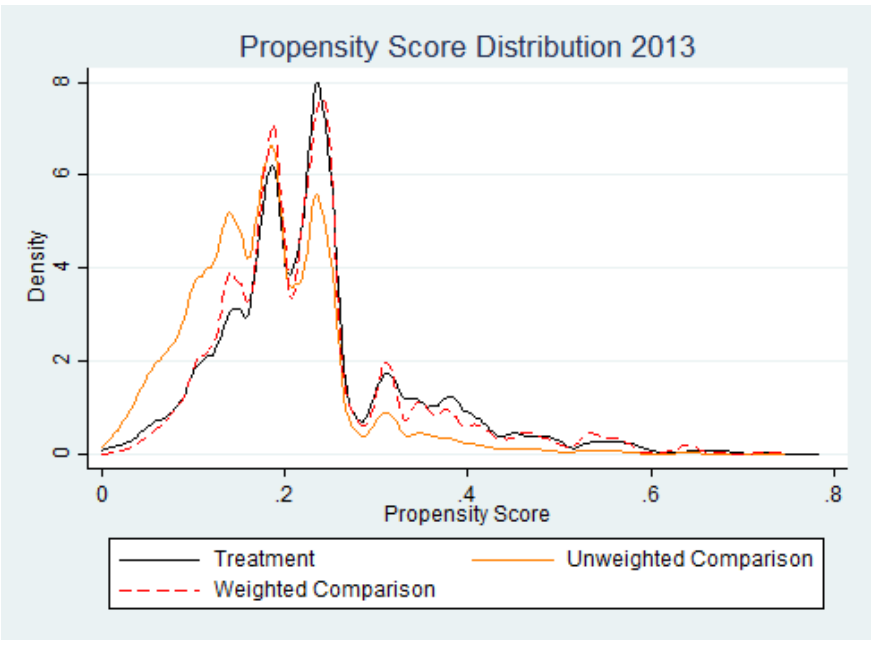
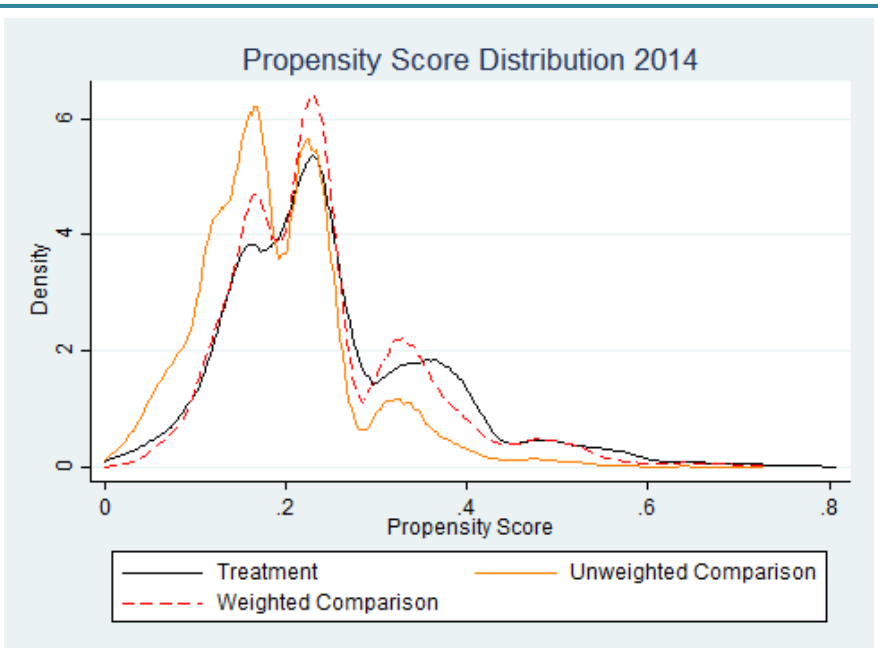


Figure A-2-15. Weighted and unweighted propensity score density plots for the Arkansas perinatal Episode of Care and comparison groups, 2014



Tables A-2-10 to ***A-2-13*** show unweighted and (propensity score) weighted means/proportions for 2011–2014. The notable group differences in the unweighted samples—health, race, and Medicaid eligibility—are substantially mitigated post-weighting as evidenced by the minimized standardized differences. Area-level characteristics have standardized differences greater than 10 percent; yet this is in part the result of limited variation in county-level variables across episodes. Furthermore, the absolute differences in the averages between the treatment and comparison group are small.

Table A-2-10. Unweighted and weighted means and standardized differences, Arkansas perinatal Episode of Care and comparison groups, 2011

	Unweighted			Weighted			
Characteristic	Arkansas perinatal EOC group	Comparison group	Standardized difference[a]	Arkansas perinatal EOC group	Comparison group	Standardized difference[a]	p-value
N	8,580	33,701		8,580	8,560		
Age indicator: 16 to 19 (%)	19.6	17.3	6.0	19.6	19.9	0.7	0.64
Age indicator: 20 to 24 (%)	43.2	43.2	0.0	43.2	42.8	0.8	0.60
Age indicator: 25 to 34 (%)	33.7	35.4	3.5	33.7	33.8	0.1	0.92
Age indicator: 35 and older (Referent)	3.5	4.2	3.5	3.5	3.5	0.3	0.87
Black (%)	26.8	35.2	18.3	26.8	26.9	0.4	0.81
Hispanic(%)	5.7	2.6	15.8	5.7	5.9	0.9	0.55
Other(%)	4.5	3.4	5.8	4.5	4.5	0.1	0.93
White (Referent)	63.0	58.9	8.6	63.0	62.7	0.7	0.63
Disability(%)	5.3	2.1	17.3	5.3	5.5	0.8	0.61
Concurrent Chronic Illness and Disability Payment System Score	1.6	1.9	34.7	1.6	1.6	3.5	0.02
Poverty-related eligibility (%)	78.8	70.9	18.2	78.8	78.5	0.7	0.67
Months of full-Medicaid enrollment during prenatal period	9.0	9.1	5.6	9.0	9.0	0.4	0.79
Diabetes (%)	4.1	2.4	9.4	4.1	4.1	0.1	0.96
Asthma (%)	2.6	5.1	13.4	2.6	2.6	0.5	0.75
Hypertension (%)	1.6	2.6	6.7	1.6	1.7	0.1	0.95
Metropolitan status of county of residence (%)	54.0	56.0	4.1	54.0	55.0	1.9	0.20
Percent of population at federal poverty level, 2012	20.7	21.1	6.5	20.7	20.8	1.6	0.30
Hospital beds per 1,000, 2010	3.8	4.6	25.0	3.8	4.4	20.2	<0.001
Median age, 2010	37.9	37.2	18.1	37.9	37.3	15.2	<0.001
Percent uninsured, ages <65, 2012	19.4	18.1	45.8	19.4	18.2	44.7	<0.001

EOC = episode of care.

[a] Absolute standardized differences are expressed as percentages.

Table A-2-11. Unweighted and weighted means and standardized differences, Arkansas perinatal Episode of Care and comparison groups, 2012

Characteristic	Unweighted			Weighted			
	Arkansas perinatal EOC group	Comparison group	Standardized difference[a]	Arkansas perinatal EOC group	Comparison group	Standardized difference[a]	p-value
N	7,438	31,163		7,438	7,423		
Age indicator: 16 to 19 (%)	18.1	15.3	7.5	18.1	18.6	1.2	0.48
Age indicator: 20 to 24 (%)	43.3	42.7	1.1	43.3	43.0	0.5	0.76
Age indicator: 25 to 34 (%)	35.0	37.4	5.0	35.0	34.8	0.4	0.80
Age indicator: 35 and older (Referent)	3.6	4.5	4.8	3.6	3.6	0.0	1.00
Black (%)	27.0	35.4	18.2	27.0	27.1	0.4	0.83
Hispanic(%)	5.7	2.6	15.7	5.7	5.9	0.8	0.64
Other(%)	4.8	3.4	6.8	4.8	4.6	0.8	0.64
White (Referent)	62.6	58.7	8.0	62.6	62.4	0.4	0.83
Disability(%)	5.8	2.2	18.8	5.8	5.9	0.4	0.80
Concurrent Chronic Illness and Disability Payment System Score	1.6	1.9	33.6	1.6	1.6	3.4	0.04
Poverty-related eligibility (%)	76.5	70.6	13.4	76.5	76.2	0.7	0.66
Months of full-Medicaid enrollment during prenatal period	9.1	9.2	6.2	9.1	9.1	0.3	0.87
Diabetes (%)	3.9	2.5	8.0	3.9	3.9	0.2	0.92
Asthma (%)	2.3	5.6	17.2	2.3	2.4	0.5	0.76
Hypertension (%)	1.7	2.8	8.0	1.7	1.7	0.6	0.72
Metropolitan status of county of residence (%)	53.9	56.2	4.5	53.9	55.2	2.4	0.14
Percent of population at federal poverty level, 2012	20.7	21.0	4.9	20.7	20.7	0.6	0.71
Hospital beds per 1,000, 2010	3.8	4.6	23.6	3.8	4.5	19.6	<0.001
Median age, 2010	37.9	37.2	20.2	37.9	37.3	17.6	<0.001
Percent uninsured, ages <65, 2012	19.5	18.1	48.3	19.5	18.2	46.5	<0.001

EOC = episode of care.

[a] Absolute standardized differences are expressed as percentages.

Table A-2-12. Unweighted and weighted means and standardized differences, Arkansas perinatal Episode of Care and comparison groups, 2013

Characteristic	Unweighted			Weighted			
	Arkansas perinatal EOC group	Comparison group	Standardized difference[a]	Arkansas perinatal EOC group	Comparison group	Standardized difference[a]	p-value
N	6,851	29,580		6,851	6,843		
Age indicator: 16 to 19 (%)	17.6	14.7	7.9	17.6	18.2	1.6	0.35
Age indicator: 20 to 24 (%)	41.8	41.6	0.4	41.8	41.2	1.1	0.50
Age indicator: 25 to 34 (%)	36.6	38.9	4.7	36.6	36.5	0.2	0.90
Age indicator: 35 and older (Referent)	4.0	4.8	4.1	4.0	4.1	0.3	0.86
Black (%)	28.0	35.8	16.9	28.0	28.3	0.7	0.69
Hispanic(%)	5.7	2.6	15.8	5.7	5.9	0.9	0.60
Other(%)	5.1	3.6	7.5	5.1	4.7	1.7	0.31
White (Referent)	61.2	58.0	6.5	61.2	61.1	0.3	0.87
Disability(%)	6.2	2.2	20.2	6.2	6.2	0.1	0.96
Concurrent Chronic Illness and Disability Payment System Score	1.6	1.9	37.0	1.6	1.6	3.6	0.04
Poverty-related eligibility (%)	75.5	70.1	12.2	75.5	75.3	0.5	0.78
Months of full-Medicaid enrollment during prenatal period	9.1	9.2	9.5	9.1	9.1	0.4	0.80
Diabetes (%)	4.5	2.4	11.6	4.5	4.5	0.3	0.88
Asthma (%)	2.4	5.9	17.6	2.4	2.5	0.4	0.80
Hypertension (%)	1.7	2.9	7.9	1.7	1.7	0.1	0.97
Metropolitan status of county of residence (%)	54.7	56.1	2.8	54.7	55.0	0.5	0.76
Percent of population at federal poverty level, 2012	20.7	21.0	5.3	20.7	20.8	2.0	0.25
Hospital beds per 1,000, 2010	3.9	4.6	21.6	3.9	4.5	17.6	<0.001
Median age, 2010	37.8	37.2	15.0	37.8	37.3	12.7	<0.001
Percent uninsured, ages <65, 2012	19.4	18.1	46.5	19.4	18.2	44.1	<0.001

EOC = episode of care.

[a] Absolute standardized differences are expressed as percentages.

Table A-2-13. Unweighted and weighted means and standardized differences, Arkansas perinatal Episode of Care and comparison groups, 2014

Characteristic	Unweighted			Weighted			
	Arkansas perinatal EOC group	Comparison group	Standardized difference[a]	Arkansas perinatal EOC group	Comparison group	Standardized difference[a]	p-value
N	6,373	25,186		6,373	6,313		
Age indicator: 16 to 19 (%)	17.4	13.5	10.7	17.4	18.0	1.6	0.37
Age indicator: 20 to 24 (%)	40.6	41.1	1.0	40.6	39.9	1.6	0.38
Age indicator: 25 to 34 (%)	37.7	40.5	5.8	37.7	37.8	0.1	0.95
Age indicator: 35 and older (Referent)	4.2	4.8	2.8	4.2	4.3	0.4	0.80
Black (%)	26.0	38.7	27.3	26.0	26.4	0.8	0.67
Hispanic (%)	6.0	2.4	17.9	6.0	5.9	0.3	0.87
Other (%)	8.6	4.1	18.5	8.6	8.3	1.2	0.50
White (Referent)	59.4	54.8	9.3	59.4	59.5	0.2	0.93
Disability (%)	5.8	2.4	17.1	5.8	5.6	0.9	0.62
Concurrent Chronic Illness and Disability Payment System Score	1.6	1.9	33.0	1.6	1.6	6.2	<0.001
Poverty-related eligibility (%)	72.1	70.0	4.8	72.1	71.8	0.8	0.65
Months of full-Medicaid enrollment during prenatal period	9.0	9.2	12.0	9.0	9.1	1.0	0.57
Diabetes (%)	3.9	2.4	8.5	3.9	3.9	0.1	0.94
Asthma (%)	2.6	6.0	16.6	2.6	2.7	0.6	0.72
Hypertension (%)	1.9	2.7	5.6	1.9	1.9	0.2	0.91
Metropolitan status of county of residence (%)	54.4	54.4	0.0	54.4	54.4	0.1	0.98
Percent of population at federal poverty level, 2012	20.6	21.5	13.0	20.6	20.9	3.3	0.06
Hospital beds per 1,000, 2010	3.9	4.7	26.6	3.9	4.5	20.2	<0.001
Median age, 2010	37.9	37.1	21.9	37.9	37.2	17.6	<0.001
Percent uninsured, ages <65, 2012	19.5	18.3	39.0	19.5	18.3	39.9	<0.001

EOC = episode of care.

[a] Absolute standardized differences are expressed as percentages.

As a way to observe potential changes in episode characteristics in Arkansas at the beginning of Arkansas's Medicaid expenditure, ***Table A-2-14*** shows the means/proportions of Arkansas perinatal episode characteristics, 2011–2013 compared to 2014. The notable group differences were observed in eligibility category, race, and baseline inpatient admissions.

Table A-2-14. Unweighted perinatal episode characteristics in Arkansas, post-Medicaid private option and pre-Medicaid private option, N = 29,610

Characteristic	2014 episodes	2011–2013 episodes	Standardized difference[a]	p-value
N	6,449	23,161		
Age indicator: 16 to 19 (%)	17.2	18.3	2.9	0.04
Age indicator: 20 to 24 (%)	40.2	42.3	4.3	<0.001
Age indicator: 25 to 34 (%)	37.3	34.6	5.7	<0.001
Age indicator: 35 and older (Referent)	4.2	3.6	2.9	0.04
Black (%)	26.3	27.4	2.6	0.06
Hispanic (%)	6.0	5.7	1.3	0.37
Other (%)	8.6	4.8	15.2	<0.001
White (Referent)	59.1	62.1	6.0	<0.001
Disability (%)	5.8	5.8	0.0	0.99
Concurrent Chronic Illness and Disability Payment System Score	1.6	1.6	0.7	0.63
Poverty-related eligibility (%)	72.3	77.1	11.0	<0.001
Months of full-Medicaid enrollment during prenatal period (%)	9.1	9.1	3.0	0.03
Diabetes (%)	3.9	4.1	1.5	0.27
Asthma (%)	2.6	2.4	1.1	0.43
Hypertension (%)	1.9	1.7	1.7	0.23
Metropolitan status of county of residence (%)	54.4	54.2	0.4	0.76
Percent of population at federal poverty level, 2012	20.7	20.7	0.3	0.86
Hospital beds per 1,000, 2010	3.9	3.9	0.4	0.81
Median age, 2010	37.9	37.9	0.1	0.95
Percent uninsured, ages <65, 2012	19.4	19.4	0.9	0.52

A-2.2.6 Propensity model evaluation for subpopulation

In addition to the overall model, we evaluated common support graphs and standardize differences of the propensity score models for the subpopulation analyses. We found that among pregnant women with a mental health or behavioral health disorder there was reasonable overlap in the distribution of the propensity score between Arkansas and the comparison group in the mental health and behavioral health subpopulation, but less substantial than in the total population. In each year there was a high proportion of comparison group episodes with scores between 0 and .2, relative to perinatal episodes in Arkansas. However, weighted standardized differences between the treatment and comparison groups indicate reasonable balance across all covariates.

A-2.2.7 Statistical analysis

Regression model

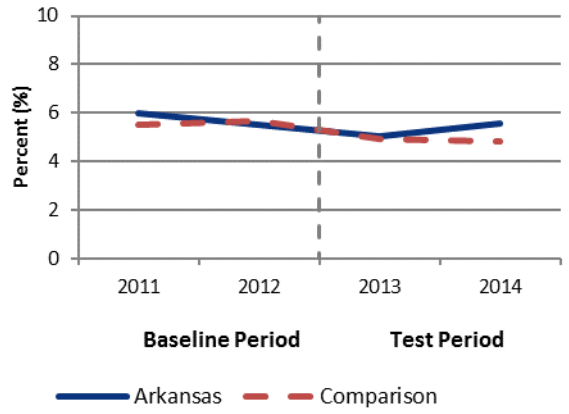
The underlying assumption in D-in-D models estimating the impact of the perinatal EOC model is that trends in the Arkansas perinatal EOC group would be similar to that of the comparison group in the absence of the EOC model (i.e., that the two were on "parallel paths" prior to the start of the perinatal EOC).

To assess the parallel assumption's validity more empirically, we modeled core utilization outcomes during the baseline period with a linear time trend interacted with a dichotomous variable indicating the perinatal episode occurred in Arkansas (i.e., the "test" group). The following section describes the baseline analysis we conducted to inform the D-in-D model.

To examine descriptively whether the trends in the perinatal EOC and comparison groups are parallel, we present graphs of annual, unadjusted averages for Arkansas perinatal episodes and the comparison group for the baseline period (2011–2012) and the first 2 years of the implementation (2013–2014).

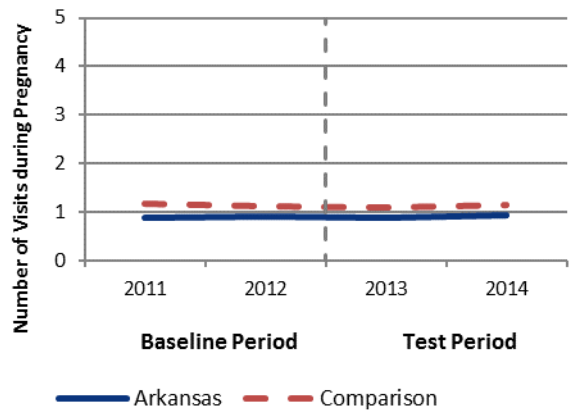
Figures A-2-16 to ***A-2-23*** provide unadjusted annual percent of inpatient admissions during pregnancy, outpatient ED visits during pregnancy, ultrasounds, 30-day readmissions post-delivery, 60-day readmission post-delivery, ED visits post-delivery, and total ED visits during the episode. With exception to the average number of ultrasounds, visual observation of unadjusted trends appear similar between the Arkansas perinatal episodes and the comparison group.

Figure A-2-16. Percentage of all-cause acute inpatient admissions during pregnancy, FY 2011–FY 2014, Arkansas perinatal Episodes of Care and comparison groups



The percentage of **acute inpatient admissions** among episodes in Arkansas declined during the baseline period, while inpatient admissions increased among the comparison group from 2011 to 2012, but declined from 2012 through 2014. In contrast, the percentage of episodes with an inpatient admission during pregnancy increased in Arkansas during the test period (***Figure A-2-16***).

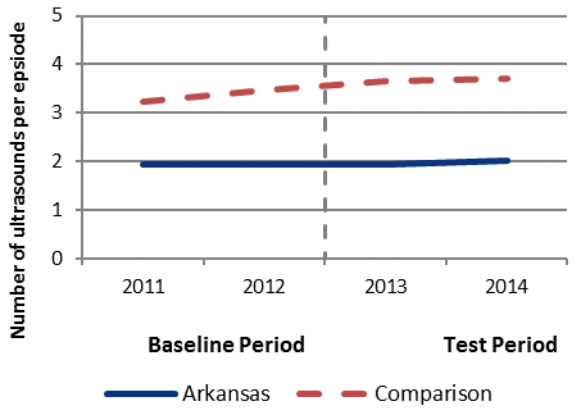
Figure A-2-17. Average number of ED visits that did not lead to a hospitalization during pregnancy, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The average number of **outpatient ED visits during pregnancy** among perinatal episodes in both Arkansas and the comparison group were similar and mostly unchanged from the baseline through the test period. The average number of visits in the comparison group were consistently higher than in Arkansas (***Figure A-2-17***).

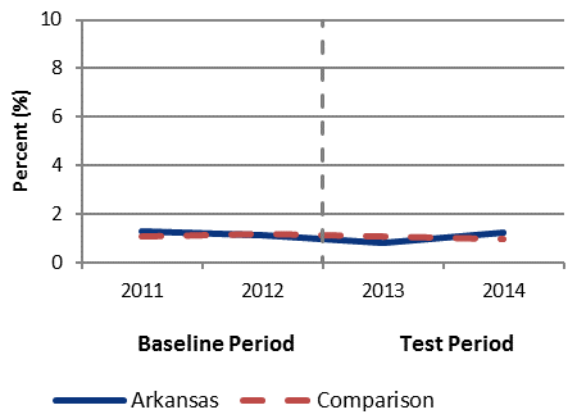
ED = emergency department.

Figure A-2-18. Average number of ultrasounds during pregnancy, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The average number of **ultrasounds during pregnancy** among perinatal episodes in Arkansas was consistent from 2011 to 2014. The average number of ultrasounds during pregnancy increased during the baseline period and leveled off during the test period for episodes in the comparison group (**Figure A-2-18).**

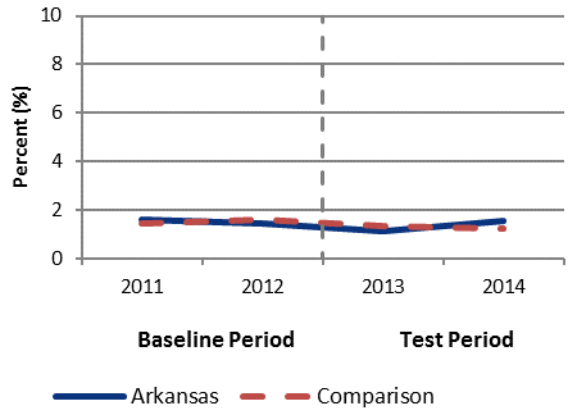
Figure A-2-19. Percentage of perinatal episodes with 30-day readmission post-delivery, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The percentage of episodes with a 30-day readmission post-delivery was generally unchanged during the baseline period for both Arkansas and the comparison group. The Arkansas perinatal EOC group had a small increase readmissions from 2013 to 2014, whereas there was no change in the comparison group (**Figure A-2-19).**

EOC = episode of care.

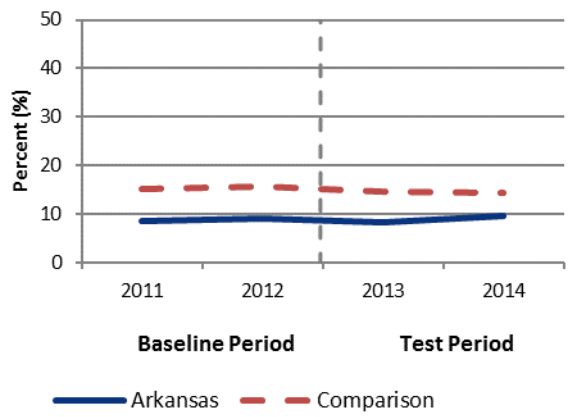
Figure A-2-20. Percentage of perinatal episodes with 60-day readmission post-delivery, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The percentage of episodes with a 60-day readmission post-delivery was generally unchanged during the baseline period for both Arkansas and the comparison group. The Arkansas perinatal EOC group had a small increase readmissions from 2013 to 2014, whereas there was no change in the comparison group (**Figure A-2-20).**

EOC = episode of care.

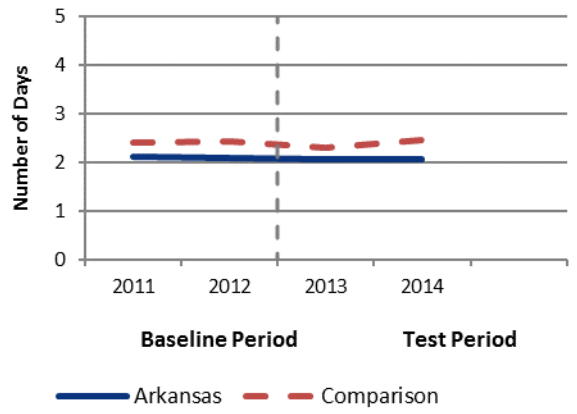
Figure A-2-21. Percentage of perinatal episodes with an ED visit post-delivery, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The percentage of episodes with an emergency department visit post-delivery was generally unchanged during the baseline period for both Arkansas and the comparison group. The percentage of visits was consistently higher in the comparison group, compared to the Arkansas perinatal EOC group (**Figure A-2-21).**

EOC = episode of care.

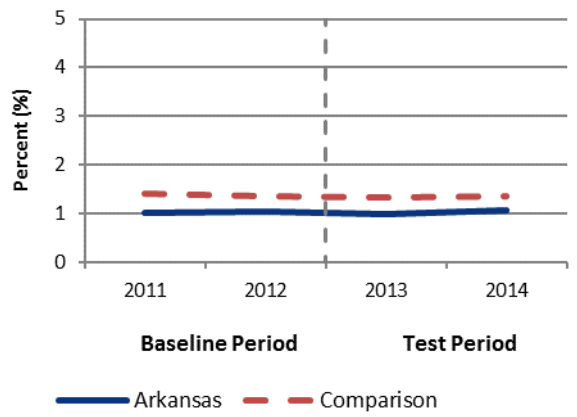
Figure A-2-22. Average length of stay during delivery-related hospitalizations, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The average **length of stay** for perinatal episodes was consistently higher in the comparison group, relative to Arkansas EOC group from 2011 to 2014. There was a small uptick in average length of stay during the test period for the comparison group (**Figure A-2-22).**

EOC = episode of care.

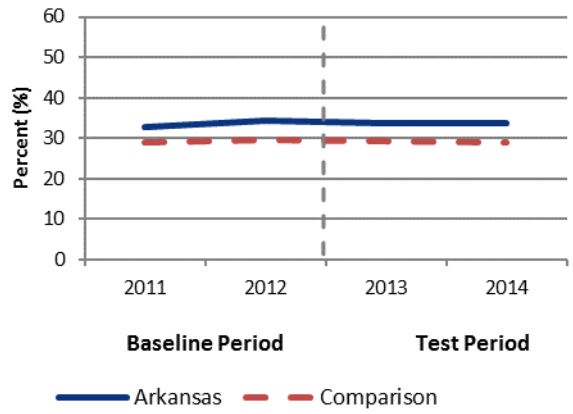
Figure A-2-23. Average number of emergency department visits during the entire perinatal episode, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The average number of emergency department visits during the entire episode was generally unchanged in both Arkansas and the comparison group from 2011 to 2014. The comparison group consistently had a higher number of visits than Arkansas over the study period (**Figure A-2-23).**

Figures A-2-24 to ***A-2-30*** provide the unadjusted averages of the quality of care measures by year.

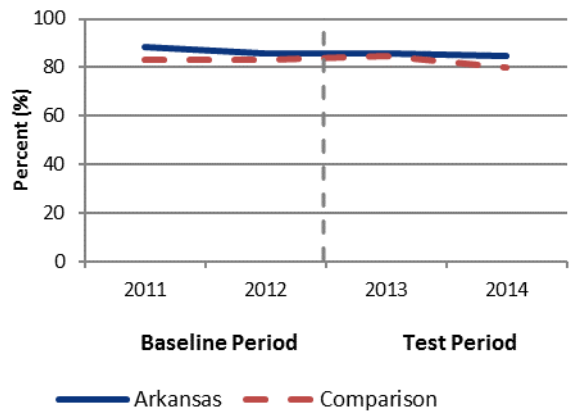
Figure A-2-24. Percentage of caesarian sections as the mode of delivery, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The percentage of **caesarian sections** among Arkansas perinatal EOC group was consistently higher than the comparison group from 2011 to 2014. There was generally observable change in the percentage between the baseline and test periods.

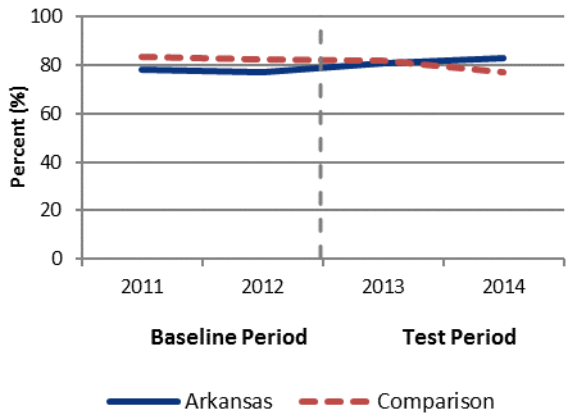
EOC = episode of care.

Figure A-2-25. Percentage of episodes with an HIV screening, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The percentage of **episodes with an HIV screening** declined slightly in Arkansas and in the comparison group from the baseline to the test period (**Figure A-2-25**). There were very few differences between Arkansas and the comparison group from 2011 to 2014.

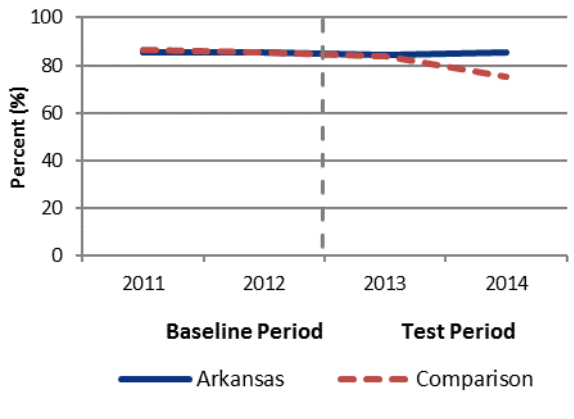
Figure A-2-26. Percentage of episodes with a chlamydia screening, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The percentage of EOCs with a **chlamydia screening** was stable from 2011 to 2012 for both Arkansas and the comparison group. The percentage of episodes with a chlamydia screening increased from 2012 to the test period, whereas it declined for EOCs in the comparison group (**Figure A-2-26**).

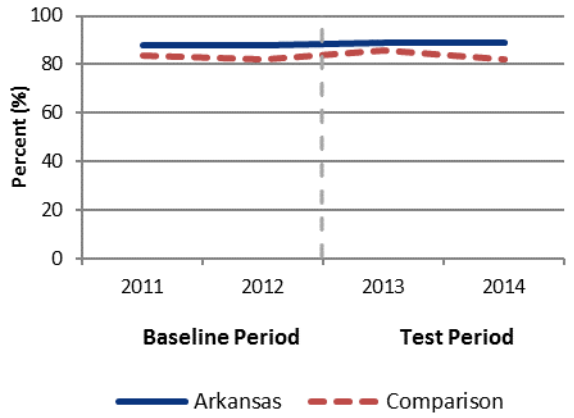
EOC = episode of care.

Figure A-2-27. Percentage of episodes with group B streptococcus screening, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The percentage of episodes with a **group B strep test** were stable during the baseline period for both Arkansas and the comparison group. There as a noticeable decline from 2013 to 2014 for the comparison group, whereas Arkansas remained stable (**Figure A-2-27**).

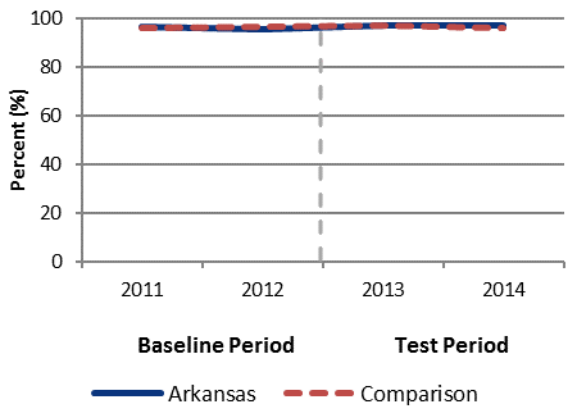
Figure A-2-28. Percentage of episodes with gestational diabetes screening, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The percentage of EOCs with a **gestational diabetes screening** was stable during the baseline period for both Arkansas and the comparison group. There was a small increase from 2012 to 2013, but then decline from 2013 to 2014 for the comparison group, whereas Arkansas remained stable (**Figure A-2-28**).

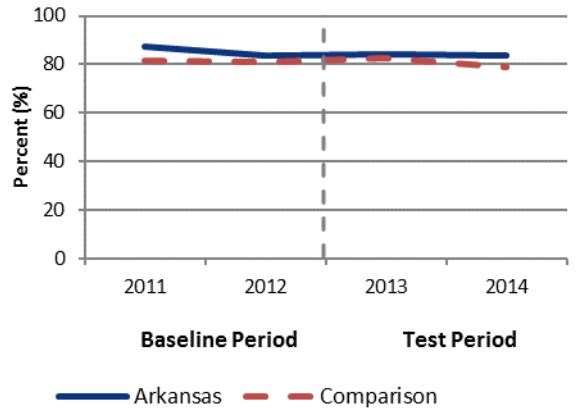
EOC = episode of care.

Figure A-2-29. Percentage of episodes with asymptomatic bacterium screening, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The percentage of episodes with a **asymptomatic bacterium screening** was stable during the baseline period for both Arkansas and the comparison group. There were no noticeable differences between the two groups (**Figure A-2-29**).

Figure A-2-30. Percentage of episodes with a hepatitis B screening, FY 2011–FY 2014, Arkansas perinatal Episode of Care and comparison groups



The percentage of episodes with a **hepatitis B screening** declined from the baseline to test period for perinatal EOCs in Arkansas, whereas the comparison group was unchanged. Both groups remained stable during the test period, with Arkansas having a slightly higher percentage (**Figure A-2-30**).

EOC = episode of care.

An annual fixed-effects model considered for the evaluation is shown in ***Equation A-2.5***:

$$\gamma = \alpha_0 + \alpha_1 I + \sum \beta_n Q_{n,b} + \sum \phi_t Q_{t,p} \bullet I + \delta X + \mu \qquad \text{(A-2.5)}$$

where

y = a performance measure (e.g., any inpatient admissions during the year) for the i-th beneficiary in the j-th group (test or comparison), in period t (i,j,t subscripts suppressed).

I = a 0,1 indicator (0 = comparison group, 1 = test group).

X = a vector of patient and demographic characteristics.

$Q_{n,b}$, $Q_{t,d}$ = 0,1 indicator of the n-th or t-th calendar year in the base (b) or post (p) period (n starts counting at first baseline period, whereas t starts with first SIM Model year).

μ = error term.

The model in ***Equation A-2.5*** assumes that, except for an intercept difference α_1, the outcomes for beneficiaries in the intervention group and beneficiaries in the comparison group followed a similar growth trend during the baseline period. We investigated whether the baseline period before the start of perinatal EOC model satisfied the baseline trend assumptions of the D-in-D model in ***Equation A-2.5***—that is, whether the outcome trends for beneficiaries in intervention and in the comparison group were similar during this period.

To test the similarity of baseline trends, we used a model with a linear trend during the baseline period. We tested whether this trend differed for perinatal EOC participants relative to

comparison group participants. Specifically, the model for the outcomes may be written as follows.

$$y = \alpha_0 + \alpha_1 I + \theta \bullet t + \lambda I \bullet t + \delta X + \mu. \quad \text{(A-2.6)}$$

In ***Equation A-2.6***, *y, I,* X, and μ are defined as in ***Equation A-2.5***. The variable *t* is linear time ranging from 1 to 2. The linear time trend in the comparison group is θ•*t*, whereas for test group beneficiaries (*I* = 1) it is $(\theta + \lambda) * t$. Hence, λ measures the difference in linear trends and the *t*-statistic for this coefficient can be used to test the null hypothesis of equal trends (λ = 0). In other words, rejecting the null hypothesis would suggest that the assumption of equal trends underlying our outcome models is not met.

The parameters of ***Equation A-2.6*** were estimated using weighted least-squares regression models for five key outcomes. These models were weighted by the propensity score. For each outcome we report estimates and standard errors of the difference between the baseline trend in the test and the comparison groups (λ).

Table A-2-15 shows estimates of the baseline trend differences for the following outcomes:

- Probability of a caesarian section
- Probability of an acute inpatient stay
- Probability of an outpatient ED visit
- Number of ultrasounds
- Length of delivery-related hospital stay

Table A-2-15. Differences in average expenditure and utilization outcomes during the baseline period, treatment and comparison group beneficiaries

Parameter estimate	C-section delivery	Any inpatient	Any prenatal outpatient ED visit	Ultrasounds	Length of stay
Test–CG trend difference	0.01	−0.01	0.02*	−0.24***	−0.04**
	(0.0)	(0.004)	(.0009)	(0.027)	(0.017)

CG = comparison group; ED = outpatient emergency department; C-section: caesarian section.

Baseline is the period January 2011–December 2013. The trend (slope) is the year-to -ear change in the outcome variable. Standard errors are given in parentheses. *p < 0.10; **p < 0.05; ***p < 0.01.

Relative to the comparison group, there was no statistically significant difference in the baseline trend for any inpatient visits and C-sections. However, there were statistical significance differences in the baseline trend for ED visits, ultrasounds, and length of stay. Based on the

overall results, we concluded that, in general, perinatal episodes in Arkansas appeared to have dissimilar trends in baseline utilization. We opted to take a conservative approach that allows us to generate effect estimates that net out the potential baseline differences between Arkansas and the comparison group. To do this, we included an interaction term between the Arkansas EOC indicator and a linear time trend in the final model. This alternative D-in-D model is described in detail below. The linear time trend controls for differences between Arkansas and the comparison group over time. As such, the D-in-D interaction term measures the deviation of the difference between Arkansas and the comparison group in the post period from the trend line. This model specification allows for differences in estimates in Arkansas and the comparison group during the baseline period, and it allows for a straightforward interpretation of the D-in-D coefficient. In this way, the alternative D-in-D model can be thought of as an interrupted time series or structural break equation that captures whether trends in Arkansas changed relative to the comparison group after the Arkansas perinatal EOC implementation.

Alternative D-in-D regression model. The alternative D-in-D model is shown in ***Equation A-2.7***. The model is an annual fixed-effects model as shown in ***Equation A-2.5.*** As in ***Equation A-2.5***, Y_{ijt} is the outcome for individual i (test or comparison group) in state j in year t; I_{ij} (=0,1) is an indicator equal to 1 if the individual is in the test group and 0 if the individual is in its comparison group; Qn is a series of yearly dummies for the baseline period (years 1 to 2); and Q_t is a series of yearly dummies for the post years (Years Three to Four). The term that interacts the Arkansas indicator and time (I_{ij}*Time) measures differences in trends between Arkansas and the comparison group over the entire period. The interaction of the test group indicator and Q_t ($I_{ij}* Q_t$) measures the difference in the pre-post change between the test group and its comparison group. Y_t is a series of yearly dummies for the post years. With this model specification, the post year*Arkansas interactions measure any deviation from the trend line in the post period.

$$Y_{ijt} = \alpha_0 + \beta_1 I_{ij} + I_{ij} * t + \sum \beta_2 Q_n + \sum \alpha_2 Q_t + \sum \gamma I_{ijt} * Q_t + \lambda X_{ijt} + \varepsilon_{ijt} \qquad \text{(A-2.7)}$$

Models for all outcomes were estimated at the episode level. Post-delivery outcomes, such as 30- and 60-day readmission and ED visits, were estimated in a subsample of episodes that had full-Medicaid benefits for at least 60 days post-delivery.

The outcome models for number of ultrasounds and length of stay, were estimated using OLS. To show the adjusted means in the pre- and post-periods for the intervention and comparison groups, we used a linear model that allows for the calculation of means that will sum to the D-in-D estimate. Although this model has strong assumptions of normality of the outcome, the OLS model still produces unbiased estimates even when the normality assumptions is violated as long as errors are uncorrelated and have a constant variance (Gauss-Markov Theorem). However, we can and do control for the correlation and variance in errors with clustered standard errors. Additionally, the model yields estimates that are readily interpretable as absolute differences and do not require additional transformation.

For quality of care outcomes, inpatient utilization during the prenatal period, readmission, and post-delivery ED use, we converted utilization counts into binary outcomes (1 = any use) and used weighted logistic regression models to estimate the probability of having any event.

We estimated count models for ED visits during the prenatal period and total ED visits during the perinatal episode. These outcomes have distributions where 21 percent of episodes had at least two visits. To account for the effects of outliers on estimated averages, we capped these outcomes at the 99.5 percentile of the distribution by state and year.

The models for inpatient admissions, caesarian sections, and total ED visits were run separately for those with mental health and behavioral health conditions.

Pre-post regression model—For the expenditure outcomes, we did not have comparison group. As such, we used the annual pre-post fixed-effects model shown in ***Equation A-2.8***:

$$\gamma = \alpha_0 + \sum \phi_t Q_{t,p} + \delta X + \mu \qquad \text{(A-2.8)}$$

where

y = a performance measure (e.g., total expenditures) for the i-th, in period t (i,t subscripts suppressed).

X = a vector of patient and demographic characteristics.

$Q_{t,d}$ = 0,1 indicator of the n-th or t-th calendar year in the base (b) or post (p) period (n starts counting at first baseline period, whereas t starts with first model year).

μ = error term.

Table A-2-16 illustrates the interpretation of the pre-post estimate from this model. The coefficient β1 in ***Equation A-2.8*** is the difference in the measure before and after perinatal EOC implementation, holding constant other variables in the equation. Using the annual fixed-effects model, we calculated overall estimates by taking linear combinations of the yearly estimates.

Table A-2-16. Pre-post estimate

Group	Pre period	Post period	Pre-post difference
Test	$\alpha_0 + \beta_2$	$(\alpha_0 + \beta_2) + \alpha_2$	α_2

The outcome model for total perinatal related payments was estimated using OLS. To show the adjusted means in the pre- and post-periods for the intervention group we used a linear model that allows for the calculation of means that will sum to the pre-post estimate. Although

this model has strong assumptions of normality of the outcome, the OLS model still produces unbiased estimates even when the normality assumptions is violated as long as errors are uncorrelated and have a constant variance (Gauss-Markov Theorem). However, we can and do control for the correlation and variance in errors with clustered standard errors. Additionally, the model yields estimates that are readily interpretable in dollars and do not require additional transformation.

Control variables. In all models we controlled for the following variables:

- Age indicator: 16 to 19 (%)
- Age indicator: 20 to 24 (%)
- Age indicator: 25 to 34 (%)
- Age indicator: 35 and older (Referent)
- Black (%)
- Hispanic (%)
- Other (%)
- White (Referent)
- Disability (%)
- Concurrent Chronic Illness and Disability Payment System Score
- Poverty-related eligibility (%)
- Months of full-Medicaid enrollment during prenatal period
- Diabetes (%)
- Asthma (%)
- Hypertension (%)
- Metropolitan status of county of residence (%)
- Percent of population at federal poverty level, 2012
- Hospital beds per 1,000, 2010
- Median age, 2010
- Percent uninsured, ages <65, 2012

Weighting and clustering. All of the regression models were estimated using weighted regressions, with exception to payments where the regression was not weighted by a propensity score. In addition, standard errors were clustered at the individual level to account for error correlation within beneficiaries with multiple perinatal episodes.

To estimate the impact of the perinatal Episode of Care (EOC) model in Arkansas, we conducted difference-in-differences (D-in-D) regression analyses using the Medicaid Analytic eXtract (MAX) files and Alpha-MAX claims from 2011 to 2014. In ***Appendix A, Section A.3***, D-in-D analyses for outcomes across two evaluation domains are presented: (1) service utilization and (2) quality of care. This sub-appendix details the methods used for these analyses.

Adjusted means. The regression-adjusted D-in-D estimate and the D-in-D calculated from regression-adjusted means will differ for one of two reasons. First, in nonlinear specifications the D-in-D calculated from the regression-adjusted means is known to be a biased estimator for the treatment effect. To address this bias, we use the nonlinear D-in-D approach described in Puhani (2012). In some cases the bias may be extreme, leading to substantial differences between the regression-adjusted D-in-D estimates versus the D-in-D calculated from regression-adjusted means.

Second, in linear specifications the D-in-D calculated from the overall regression-adjusted means may be substantially different than the overall regression-adjusted D-in-D estimate because we use different weights to obtain the overall figures. Specifically, the regression-adjusted D-in-D estimates are weighted using the number of treatment beneficiaries observed in each year relative to the total number of treatment beneficiaries ever observed during the test period. This is mathematically equivalent to weighting the test-period adjusted means for both groups with the same weights that are applied to the treatment group. However, the test-period adjusted means that are presented for the comparison group are weighted using the number of comparison beneficiaries observed in each year relative to the total number of comparison beneficiaries ever observed during the test period. The implication of this is that in cases where there are large differences in the rates of rolling entry or exit across the two groups, we may observe large differences in the D-in-D calculated from the overall regression-adjusted means versus the overall regression-adjusted D-in-D estimate.

A-2.3 Methods for the Impact Analysis of Spillover Effects of the Medicaid PCMH Model Using the Arkansas All-Payer Claims Database[3]

To estimate the spillover impact of the Arkansas Medicaid PCMH model on commercial beneficiaries, we conducted a difference-in-differences (D-in-D) quasi-experimental design that takes advantage of the rolling adoption of PCMH across practices using Arkansas All-Payer Claims Database (APCD). We present results of both descriptive trends and pre-post analyses for

[3] This report and its findings are independent research conducted by RTI International. The Arkansas Insurance Department and the Arkansas All-Payer Claims Database have not evaluated the content of the report or its findings beyond determining compliance with minimum cell size and complimentary cell suppression rules; incorporation of appropriate protections to prevent inferential identification; consistency with the initial project description. The said report or findings do not represent the positions or opinions of the Arkansas Insurance Department or the Arkansas Healthcare Transparency Initiative Board.

outcomes across two evaluation domains: (1) service utilization and (2) expenditures. This appendix details the methods we used for this analysis.

PCMH in the context of Arkansas Medicaid. Arkansas's Medicaid program, representing about 22 percent of the state's population, has undergone several reforms since 2012, in part because of SIM funding, to move beyond a fee-for-service model. The PCMH model that began in 2014 was built off these reforms by replacing primary care management fees with a risk-based per member per month (PMPM) fee to cover care coordination and more intensive case management. The state also extended the PCMH model requirements of Arkansas's Comprehensive Primary Care initiative for Medicare beneficiaries to Medicaid beneficiaries. The PCMH model in Arkansas was a delivery model led by a primary care provider (PCP) who coordinates access to patient care. Providers enrolled in Medicaid's PCMH model received a PMPM fee to cover care coordination and ongoing transformation costs such as meeting criteria to become a medical home in addition to fee-for-service payments. Our analysis focused on members enrolled in commercial plans who were attributed to PCMH practices, and whether the Medicaid PCMH initiative had any spillover impacts on the commercial population.

Profiles of PCMH participating providers. Roughly 47.5 percent of eligible provider groups in the state were enrolled in the PCMH program by the initial start date of January 2014. These practices served up to 72 percent of eligible Medicaid beneficiaries in 2014. By 2016, 71.6 percent of practices and 86.9 percent of eligible providers in Arkansas were enrolled in the Medicaid PCMH model. Participating PCMHs were mostly family care practices, and about a third were pediatric practices.

Attributing members to PCMH practices. Because PCMH is a practice-level intervention, we retained only individuals who can be attributed to a PCMH practice in a given year for the analytic sample. We did not have information on whether individuals were assigned to a PCP by their insurer and instead use an algorithm to assign individuals to a practice based on the pattern of PCP visits during the calendar year. We attributed individuals each calendar year to a single PCMH group, however, practice attribution could vary across years as utilization patterns changed. To be attributed to a PCMH practice in a calendar year, we required that an individual have at least three PCP visits associated with a specific PCMH practice and for that practice to have the plurality of total PCP visits. We excluded an individual-year observation if there were no PCMH PCP visits within that year. We excluded members who do not have at least 6 months continuous enrollment in a commercial insurance plan or 9 months noncontinuous enrollment in a commercial insurance plan during the calendar year.

We tracked PCMH practices using the billing National Provider Indicator (NPI) listed on the claim. For each member, we counted the number of PCP visits to each billing NPI in a year. We then merged on billing NPIs from the enrollment file by year and identified PCP visits to PCMH and non-PCMH practices. To be attributed, a member must have had most of their

primary care at a single practice and have had at least three PCP visits per to that practice per year. If a member had equal number of visits to multiple practices they were attributed to the provider with the most recent date of service in a given calendar year. There was a small sample of members who had multiple most recent visits occurring on the same date. These individuals were dropped. The final analytic sample included all claims for members attributed to a PCMH practice in a given year.

Regression-adjusted analysis. The quasi-experimental study design used a difference-in-differences (D-in-D) approach to compare PCMH practices based on the year they become a PCMH between 2014–2016. Each provider was part of the comparison group until they begin receiving PMPM payments, after which they become part of the treatment group. The later adopters provided a comparison set of observations for the early adopters. ***Table A-2-17*** shows how practices transitioned from the comparison to the treatment group over time.

Table A-2-17. PCMH implementation between 2014 and 2016

Analytic year	Treatment group	Comparison group
2013	*No* practices	*All* practices
2014	Practices that became a PCMH in early and mid-2014	Practices that became a PCMH in 2015 or 2016
2015	Practices that became a PCMH in 2014 or 2015	Practices that became a PCMH practice is 2016
2016	*All* practices	*No* practices

PCMH = patient-centered medical home.

In our design we assigned treatment status based on PCMH certification as reported by the Arkansas Department of Human Services. We did not observe the uptake or maintenance of all PCMH components and were not able to measure variation in implementation. Our approach exploits the rolling adoption of PCMHs across providers to compare outcomes for patients who received care at a PCMH clinic (treatment group) to those who received care from clinic who had not yet become a PCMH (comparison group). We limited the sample only to practices that became a PCMH during the study period. This helped control for practice level characteristics that we could not measure or control for with the available data such as selection bias (some practices were more eager to take on practice transformation activities than others). Although this improved the internal validity and comparability between our treatment and comparison groups, it limited generalizability to all practices in the state.

A second strength of the model was that the natural variation in adoption allowed us to control for secular changes outside of PCMH adoption that occurred in both the treatment and comparison group. For example, the 2014 ACA implementation presented a large challenge in estimating effects of PCMH adoption given the amount of change occurring. By observing

repeated adoptions of PCMH across time, we may be able to detect effects separate from the ACA that are attributable to PCMH.

D-in-D and CITS models are subject to a critical assumption that there are similar pre-treatment trends, or the parallel trends assumption. We tested this assumption empirically using the common baseline year of 2013 and find little evidence of a differences in pre-PCMH trends—this information is presented in ***Table A-2-18***. There were some slight differences in inpatient expenditures and visits between the 2015 and 2016 PCMH groups, however we do not expect these to impact our estimates.

More information on the study outcomes is available in ***Section A-2.3.2,*** and more information on the regression model is available in ***Section A-2.3.4***.

A-2.3.1 Data sources

Arkansas All-Payers Claims Database (APCD). Commercial claims were extracted from the Arkansas APCD for calendar years 2013–2016. The APCD includes claims from individual market, small employer, large employer and state/Federal health plans. Claims from self-insured employers are largely not included and we excluded any out of state claims.

Members were assigned an identifier based upon last name and date of birth. However, this methodology did not allow us to always identify unique members and produced duplicate identifiers for approximately 20% of the sample. For example, twins or individuals with common last names and the same birthdate would be assigned the same unique identifier. We investigated additional covariates that could be used to determine unique members such as gender and county of residence, however, these also could not always identify unique members. Given this limitation of the data and to avoid introducing potential bias, we dropped all duplicate member numbers.

The claims file was structured in separate files for medical claims, pharmacy claims, and eligibility information. We used these files to create member-quarter inpatient, outpatients and pharmacy analytic files merged with eligibility information for calendar years for claims with a date of service in 2013–2016.

Table A-2-18. Test of parallel assumptions for utilization and expenditures in 2013, Arkansas commercial beneficiaries not enrolled through the Marketplace, 2013–2016

Outcome	Early 2014 PCMH change in slope estimate (90% confidence interval)	p-value	Mid 2014 PCMH change in slope estimate (90% confidence interval)	p-value	2015 PCMH change in slope estimate (90% confidence interval)	p-value	Weighted N
Total expenditures	−0.2 (−16.02, 15.68)	0.99	−1.4 (−21.38, 18.68)	0.91	97.3 (39.06, 58.20)	0.01	121,516
Professional expenditures	0.9 (−5.05, 6.88)	0.80	4.3 (−1.42, 9.94)	0.22	9.2 (0.94, 8.22)	0.07	121,516
Pharmaceutical expenditures	−1.1 (−4.24, 1.95)	0.54	−1.5 (−4.81, 1.84)	0.46	6.1 (−4.69, 10.79)	0.35	121,516
Inpatient facility expenditures	2.1 (−5.78, 10.01)	0.66	4.9 (−6.79, 16.55)	0.49	76.8 (15.31, 61.45)	0.04	121,516
Outpatient facility expenditures	−2.1 (−8.17, 4.06)	0.58	−9.0 (−23.69, 5.68)	0.31	5.2 (−5.12, 10.37)	0.41	121,516
Inpatient stays per 1,000 member-quarters	−0.3 (−1.29, 0.79)	0.69	0.5 (−0.74, 1.68)	0.52	2.7 (0.64, 2.01)	0.03	115,597
Primary care visits per 100 member-quarters	0.6 (−0.16, 1.42)	0.19	0.8 (−0.11, 1.75)	0.15	0.5 (−0.83, 1.35)	0.53	121,516
Specialist visits per 100 member-quarters	0.1 (−0.40, 0.63)	0.70	−0.2 (−0.69, 0.34)	0.58	0.3 (−0.25, 0.60)	0.34	121,488
Emergency department visits per 1,000 member-quarters	−1.9 (−4.95, 1.22)	0.32	0.4 (−3.23, 3.97)	0.87	−3.1 (−6.94, 3.83)	0.18	121,404

PCMH = patient-centered medical home.

Note: Comparative interrupted time series regression models were estimated using Ordinary Least Squares for the expenditure outcomes and maximum likelihood logit for the service use outcomes. The change in slope estimates represents the difference in the linear trend relative to the 2016 PCMH group. Standard statistical practice is to use confidence intervals of 90% or higher.

Data source: RTI analysis of AR APC data, 2013.

PCMH practice enrollment. Arkansas provided RTI with a list of practices meeting the PCMH certification process between 2014 and 2016. There were 4 different points at which practices were certified as being enrolled in PCMH: early 2014; mid-2014; 2015; and 2016. We track practices across years using a billing NPI. For each year, we considered a practice to be enrolled in PCMH if they are certified for the full calendar year. If a practice's PCMH certification was suspended or terminated early during a calendar year, they were not considered enrolled for that year. If a practice lost certification across analytic years it was removed from the subsequent years of analysis to ensure that former PCMH practices were not in the comparison group.

Area Health and Resource File (AHRF). The AHRF comprises data collected by the Health Resources and Services Administration from more than 50 sources containing more than 6,000 variables related to health care access at the county level. We used information on health system capacity, poverty at the county level, age, rural/urban status, and uninsured rates from 2013–2015 as covariates in the analysis. Values for 2015 are also used in 2016; values in the relevant variables did not update between 2015 and 2016. In instance of missing county information, we assigned the median value for the state for that year.

A-2.3.2 Outcome measures

Utilization

Utilization measures were reported as rates per 1,000 or 100 covered lives. For each measure, we first calculated the probability of any use. To calculate the probability, the numerator was an indicator of having had at least one event (inpatient admission or emergency department [ED] visit that did not lead to a hospitalization), and the denominator was the number of eligible members (or discharges) during the quarter. We multiplied the probability of use by 1,000 or 100 to obtain approximate rates of utilization per 1,000 members. Multiplying the probability by 1,000 did not produce an exact rate of utilization per 1,000 members because it assumed no person has more than one visit or admission per quarter. However, we concluded that this is a reasonable approximation because the majority of the population had zero or one ED visit or admission per quarter. Events were included in a period's total if discharge or service date on the claim was during the period.

- **Rate of all-cause inpatient hospitalizations (per 1,000 covered members):** This is an indicator of whether the member had at least one admission to an acute-care hospital reported in the medical claims file for the quarter, divided by the number of members in the same quarter. Inpatient admissions were defined as claims with a Bill Type equal to 11 or 12. Some records in the inpatient claims files appeared to be multiple admissions but were in fact transfers between facilities; these records were counted as a single admission. To roll up transfers into one acute admission, we first identified claims that had no more than 1 elapsed day between discharge date of the index claim and admission date of the subsequent claim. We then combined these

claims into one record by taking the earliest admission date and latest discharge date and summing all payment amounts. This same roll-up procedure was applied to claims with overlapping or identical admission and discharge dates (i.e., claims associated with the same visit).

- **Rate of ED visits that did not result in an inpatient hospital admission (per 1,000 covered members):** This is an indicator of whether the member had at least one visit to the ED that did not result in an inpatient hospital admission divided by the number of members in the same quarter. ED visits (including observation stays) were identified in the outpatient file (Bill Type ≠ 11 or 12) as visits with a revenue code equal to 0450–0459, 0981, or 0762 *or* facility type equal to 23 with procedure code equal to 99281, 99282, 99283, 99284, or 99285. If the procedure code on every line item of the ER claim equaled 70000–89999 and no line items had a revenue center code equal to 0762, that claim was excluded (thus excluding claims for which only radiology or pathology/laboratory services were provided unless they were observation stays). Multiple ED visits on a single day were counted as a single visit.
- **Rate of PCP visits (per 100 covered members):** This is an indicator for whether the member had at least one visit to a primary care provider reported in the medical claims file for the quarter, divided by the number of members in the same quarter. Primary care physicians were identified using their primary taxonomy code, which was obtained from the National Plan and Provider Enumeration System (NPPES) file. A taxonomy code was considered primary where it was denoted in the NPPES file with a Y or an X. When searching for primary care visits, claims were restricted to those with Healthcare Common Procedure Coding System (HCPCS)/Current Procedural Terminology (CPT) codes indicating evaluation and management (E&M) visits associated with planned physician care (i.e., office visits). Both inpatient and outpatient files were included, although E&M codes used to identify physician visits should occur only in the outpatient file.
- **Rate of Specialist visits (per 1000 covered members):** This is an indicator of whether the member had at least one visit to a specialty provider reported in the medical claims file for the quarter, divided by the number of members in the same quarter. Specialty care physicians were identified using their primary taxonomy code, which was obtained from the NPPES file. A taxonomy code was considered primary where it was denoted in the NPPES file with a Y or an X. When searching for specialty care visits, claims were restricted to those with HCPCS/CPT codes indicating E&M visits associated with planned physician care (i.e., office visits). Both inpatient and outpatient files were included, although E&M codes used to identify physician visits should occur only in the outpatient file.

Expenditures

Weighted average expenditures were calculated on a per member per month (PMPM) basis. For each individual, PMPM expenditures were estimated as one-third of their quarterly expenditures. Expenditures were defined as payments made by the commercial insurer. Averages included all individuals enrolled during the period, so that the figures also reflect the presence of individuals with zero medical costs. The payments were not risk adjusted or price standardized

across geographic areas. Reversal (i.e., negative) claims were included in the calculations; however, negative payments at a member-year level were set to zero for total expenditures. Depending on the type of claim, claims were included in a period's total if the discharge, end date, or prescription fill date was during the year of interest. All expenditures are adjusted to 2014 dollars using the Bureau of Labor Statistics Medical Consumer Price Index.

- **Total:** This represents overall net payment amounts from all inpatient and outpatient (facility and professional) medical claims and all pharmacy claims.
- **Inpatient facility:** This represents the sum of net facility payments to a hospital for covered services provided during all inpatient admissions. Inpatient claims were identified using Bill Type = 11 or 12.
- **Non-Inpatient facility:** This represents the sum of net facility payments to a non-inpatient facility for covered services provided during a non-inpatient visit. Non-inpatient claims were identified using Bill Type not equal 11 or 12 and Claim Type = 2
- **Professional:** This represents the sum of net payments from all inpatient and outpatient professional claims. Professional claims were identified as claims for which Claim Type = 1
- **Prescription:** This represents the sum of net payments in the pharmacy claims files.

A-2.3.3 Population studied

PCMH-attributed members. As described in ***Section A-2.3.1***, Arkansas provided RTI with a list of PCMH practices from 2014 to 2016. We then applied the attribution method to identify members associated with a PCMH practice in each year. Prior to attribution, there were 1,044,205 unique members between 2013 and 2016. After applying the continuous enrollment criteria, the sample was reduced to 293,583 unique members. Of those, 124,493 members were attributed to a PCMH practice. We excluded a small number of individuals with a missing gender or insurance type as well as members 65 years of age or older as we did not observe Medicare claims. We also excluded a small number of individuals that have dental- or pharmacy-only insurance coverage. We additionally excluded a small number of individuals with a quarterly eligibility of less than 20 percent or with weighted quarterly expenditures above $1,000,000. The final analytic sample included 121,073 unique, unweighted individuals.

Table A-2-19 below shows the member characteristics by year.

Table A-2-19. Weighted annual sample characteristics, PCMH-attributed Arkansas commercial plan members, 2013–2016

Characteristic	Annualized by PCMH status		Annualized by calendar year				Annualized by PCMH group in 2013			
	Pre-PCMH	Post-PCMH	2013	2014	2015	2016	Early 2014	Mid 2014	2015	2016
Total expenditures	$4,044	$5,784	$3,427	$4,963	$6,036	$6,555	$3,243	$4,402	$4,304	$3,432
Professional expenditures	$1,561	$2,046	$1,400	$1,921	$2,125	$2,206	$1,348	$1,661	$1,600	$1,431
Prescription expenditures	$870	$1,270	$835	$1,007	$1,307	$1,466	$809	$963	$1,063	$797
Inpatient facility expenditures	$693	$1,077	$459	$873	$1,156	$1,279	$405	$721	$813	$445
Other facility expenditures	$920	$1,390	$734	$1,162	$1,448	$1,604	$681	$1,057	$828	$759
Any inpatient visit (%)	4.4	7.1	3.1	5.8	7.8	8.1	2.8	4.4	4.4	3.2
Any specialist visit (%)	42.5	48.6	43.0	46.3	49.3	50.7	42.6	50.3	41.3	39.9
Any ED visit (%)	24.1	31.5	20.7	29.6	33.0	33.5	20.5	20.7	22.8	21.7
Age	31.7	34.5	28.0	33.0	35.1	35.9	26.5	33.5	36.1	29.9
Female (%)	58.3	61.0	55.2	61.0	61.1	61.4	54.7	57.4	59.9	54.6
BH diagnosis (%)	20.0	23.9	18.5	22.6	24.6	25.4	18.2	17.6	23.2	19.6
Lives in MSA (%)	53.0	52.2	57.2	51.7	51.8	50.0	59.4	58.1	36.0	49.1
Has prescription drug coverage (%)	87.2	90.1	83.9	89.6	90.0	91.1	83.5	85.0	90.7	82.8
Marketplace plan (%)	20.8	43.0	0.0	36.3	46.2	48.2	0.0	0.0	0.0	0.0
Insurance product type										
PPO (%)	54.9	59.3	49.8	61.4	60.1	57.5	52.6	41.9	42.2	40.6
PoS (%)	21.8	14.3	27.2	16.0	13.9	12.9	25.9	32.4	30.1	30.2
Other commercial insurance (%)	15.5	19.5	14.2	16.8	19.3	22.0	13.2	14.9	16.1	19.4
Insurance type—PPO (%)	54.9	59.3	49.8	61.4	60.1	57.5	52.6	41.9	42.2	40.6
Insurance type—PoS (%)	21.8	14.3	27.2	16.0	13.9	12.9	25.9	32.4	30.1	30.2
Insurance market type										
Individual market plan (%)	43.7	60.5	27.8	54.5	63.3	64.7	26.9	33.9	30.6	27.1
Large employer plan (%)	33.3	24.4	42.7	26.6	23.2	22.4	43.5	40.0	41.6	40.2
Small employer plan (%)	9.2	6.7	11.8	7.4	6.4	5.9	11.9	11.8	12.8	11.2
Unweighted N	52,970	134,256	30,989	46,121	51,755	52,499	23,209	3,181	1,315	3,284
Weighted N	55,881	140,627	32,107	48,846	54,078	55,027	24,040	3,287	1,371	3,408

BH = behavioral health; ED = emergency department; MSA = Metropolitan Statistical Area; PCMH = patient-centered medical home; PoS = point of service; PPO = preferred provider organization.

Data source: RTI analysis of AR APC data, 2013–2016.

A-2.3.4 Statistical analysis

Regression model

Difference-in-differences regression model—*Equation A-2.9* shows our main analytic approach, which is a two-way fixed effects D-in-D model that includes both quarter fixed effects and PCMH practice fixed effects:

$$Y_{ict} = \beta_0 + \beta_1 PCMH_{ict} + \Sigma\, \delta_t Q_t + \Sigma\, \tau_c P_c + \gamma \boldsymbol{X}_{ict} + \varepsilon_{ict} \qquad \text{(A-2.9)}$$

where

Y_{ict} = outcome of interest for individual, *i*, in practice *c*, during quarter *t*

$PCMH_{ict}$ = equal to 1 if the practice is receiving Medicaid PCMH payments

X = a vector of patient and demographic and county characteristics.

Q_t = a quarterly fixed effect

P_c = a practice fixed effect

ε_{ict} = error term.

Typically, a difference-in-differences specification would control for baseline differences between members attributed treatment and comparison groups. However, because this measure did not change over time within practices, it was included in the practice fixed effects along with other practice characteristics that did not change over time. δ_t captured the change in the outcome among practices who had not achieved PCMH certification, and a linear combination of the β_1 coefficient and the δ_t coefficients captured the change in the outcome among the practices who had achieved PCMH certification. Therefore, the D-in-D parameter, β_1, shows whether the difference between practices who have and have not achieved PCMH certification increased ($\beta_1 > 0$) or decreased ($\beta_1 < 0$) after PCMH certification was achieved. Results from ***Equation A-2.9*** are presented in **Table A-14** in ***Section A.4.1*** and in **Table A-16** in ***Section A.4.2*** in ***Appendix A***. Estimates from this model also appear in ***Sub-appendix A-1.3***, ***Section A-1.3.2*** (**Table A-1-11**), ***Section A-1.3.3*** (**Table A-1-14**), and ***Section A-1.3.4*** (**Table A-1-19**).

We also estimated a variation of ***Equation A-2.9*** that estimated a separate effect for PCMH practices that adopted in early 2014 relative to those who adopted in late 2014 or later. Specifically, we estimated the specification presented in ***Equation A-2.10***:

$$Y_{ict} = \beta_0 + \beta_1 PCMH_{ict} + \beta_2 Early2014_{ict} + \beta_3 Early2014_{ict} * PCMH_{ict} + \Sigma\, \delta_t Q_t + \Sigma\, \tau_c P_c + \gamma \boldsymbol{X}_{ict} + \varepsilon_{ict} \qquad \text{(A-2.10)}$$

In this equation, β_1 was the effect of PCMH certification for the non-early 2014 groups and β_3 was the average change for the early 2014 adopters relative to the other adopting groups. β_3 captured whether the early 2014 adopters, who have 3-years of post-PCMH observations, saw

differential changes in their outcomes associated with PCMH relative to the groups with less exposure to PCMH across. Results from ***Equation A-2.10*** are presented in **Table A-15** in ***Section A.4.1*** and in **Table A-17** ***in Section A.4.2*** in ***Appendix A***. Estimates from this model also appear in ***Sub-appendix A-1.3***, ***Section A-1.3.2*** (**Table A-1-12**), ***Section A-1.3.3*** (**Table A-1-15**), and ***Section A-1.3.4*** (**Table A-1-20**).

The outcome model for expenditures was estimated using weighted ordinary least squares. For the utilization outcomes, we converted utilization counts into binary outcomes (1 = any use during the quarter) and used weighted logistic regression models. Count models were not appropriate because of the low occurrence of most types of utilization for individual members in any year; however, we multiplied the marginal effect from the logistic regression models by 100 or 1,000, as appropriate, to obtain approximate rates of utilization per 100 or 1,000 members. Multiplying the marginal effect by 100 or 1,000 did not produce an exact rate of utilization per 100 or 1,000 members because it assumed that no person has more than one visit or admission per year. However, we concluded that this was a reasonable approximation because only a small percentage of members had counts exceeding 1 for any of the utilization measures.

Control variables. In all models we controlled for the following variables:

- Age and age squared
- Gender
- Has a behavioral health diagnosis during the quarter
- Has prescription drug coverage during the quarter
- Has Marketplace coverage during the quarter
- Has a claim flagged as Medicaid private option during the quarter
- Insurance type (PPO, PoS, or other commercial plan) during the quarter
- Insurance market (individuals, small employer, large employer, or other) during the quarter
- County-level percent of population under Federal Poverty Level for the calendar year
- County-level median age for the calendar year
- County-level percent uninsured for the calendar year
- County-level percent living in a Metropolitan area for the calendar year
- County-level hospital beds per 1,000 population for the calendar year

Weighting and clustering. All the regression models were estimated using weighted regressions and weighted by eligibility fraction (the fraction of the year during which the member was eligible for the analyses). In addition, standard errors were clustered at the PCMH

practice level to account for correlation in the error term between multiple members within practices.

Sensitivity analyses

CITS regression model—As a sensitivity approach and to test for dynamic effects of PCMH adoption, we also estimated a CITS version of the two above models. The coefficients in the model could estimate both the average change in the outcome for the commercial members once the practice began receiving Medicaid PMPM payments and the change in the trend of the outcome. Modeling the change in trend was important given that there may be a gradual change in costs or utilization as providers adapted to PCMH components.

$$Y_{ict} = \beta_0 + \beta_1 PCMH_{ict} + \beta_2 Quarter_t + \beta_3 PCMH_{ict} * Quarter_{ict} + \gamma \boldsymbol{X}_{ict} + \tau_c + \varepsilon_{ict} \quad \text{(A-2.11)}$$

Equation A-2.11 modifies ***Equation A-2.9*** replacing the quarterly fixed effects with a linear time trend that is interacted with the PCMH indicator. In this alternative approach, β_1 was the average change in the outcome at the time of adoption and β_3 was the change in the outcome trend after PCMH adoption. Estimates from ***Equation A-2.11*** appear in ***Sub-appendix A-1.3, Section A-1.3.1*** (***Table A-1-8***) and ***Section A-1.3.3*** (***Table A-1-16***).

Caution should be taken in directly interpreting the coefficients from ***Equation A-2.11***: The effect of PCMH adoption varied at every quarter. For example, in ***Table A-1-8***, for the total expenditure outcome, β_1=−49.5 and β_3 = 6.0. For January 2013, $Quarter_{ict}$=1, for January 2014, $Quarter_{ict}$=4, and so forth. Thus, to calculate the effect of PCMH in January 2014, we calculated $\beta_1 + 4*\beta_3$, which was equal to −25.5. In this example, since $\beta_1 < 0$ and $\beta_3 > 0$, the treatment effect was negative in the short-run and positive in the long-run.

$$Y_{ict} = \beta_0 + \beta_1 PCMH_{ict} + \beta_2 Quarter_t + \beta_3 PCMH_{ict} * Quarter_{ict} + \beta_4 Early2014_{ict} + \beta_5 Early2014_{ict} * Quarter_t + \beta_6 Early2014_{ict} * PCMH_{ict} + \beta_7 Early2014_{ict} * PCMH_{ict} * Quarter_t + \gamma \boldsymbol{X}_{ict} \tau_c + \varepsilon_{ict} \quad \text{(A-2.12)}$$

Equation A-2.12 applies the same change to ***Equation A-2.10*** to obtain the net difference in the initial change (β_6) and the trend change associated with PCMH adoption for the early 2014 adopters (β_7). Estimates from ***Equation A-2.12*** appear in ***Sub-appendix A-1.3***, ***Section A-1.3.1*** (***Table A-1-9***) and ***Section A-1.3.3*** (***Table A-1-17***).

Using a sample attributed with a one-visit floor—One potential issue with using a three-visit floor was that we potentially introduce measurement error because the sample was limited to higher utilizers and excludes low-utilizing individuals by default. As a sensitivity test, we assessed the potential bias this introduced by estimating the models used a sample attributed by a one-visit floor. Estimates from the one-visit models are available in ***Sub-appendix A-1***, ***Sections A-1.3.3*** and ***A-1.4.4.***

A-2.4 Methods for Qualitative Data Collection and Analysis

The Arkansas SIM Initiative Round 1 evaluation team collected and analyzed a wide range of qualitative data in the fifth year of the federal SIM Initiative evaluation. These data sources included interviews with key informants and focus groups conducted during in-person site visits in previous evaluation years, a review of relevant documents, and regular evaluation calls with state officials leading the state's SIM Initiative. This report draws from past evaluation reports, where further detail is provided on previously conducted site visit interviews and focus groups.

A-2.4.1 Document review

We used Arkansas's quarterly and annual reports, operational plans, and other state documents to obtain updated information on their implementation progress during the SIM Initiative test period. To supplement these documents, we collected relevant news articles on the Arkansas SIM Initiative activities and related initiatives, and we searched reform-oriented websites that the state maintains.

In addition, we obtained numbers of providers participating in and populations reached by the different innovation models from reports Arkansas submits to the Innovation Center in conjunction with its quarterly reports. We provide Arkansas's reported numbers in ***Appendix A***. Sources for these provider and population data are detailed in the Year Four Annual Report (RTI International, 2018).

A-2.4.2 State evaluation calls

We conducted monthly federal evaluation-specific calls beginning in April 2014 and continued through the end of the SIM Initiative test period. The evaluation team for Arkansas, the state's SIM Initiative team, and the state's Innovation Center project officer typically attended each state evaluation call. The purpose of the calls was to review interim evaluation findings with the state (as available), discuss any outstanding federal evaluation data or other needs, review and discuss state implementation and self-evaluation updates, and gather more in-depth information on select topics of interest for the evaluation.

For each meeting, the evaluation team prepared a list of state-specific questions, including the status of related policy levers and implementation successes, challenges, and lessons learned. We first reviewed relevant state documents for answers to our questions. When we did not find answers in the document or needed further clarification, we sent the questions to the state ahead of the call and asked the state to have knowledgeable state officials available to answer the questions during the call.

A-2.4.3 Analysis

The evaluation team conducted thematic analysis of each source of qualitative data and then synthesized across information gleaned from site visits, focus groups, document review, and state evaluation calls. Site visit interviews and focus groups were conducted in previous years of the evaluation. For more detail on site visit interview and focus group methods, see past evaluation reports.

A-2.5 References

ACHI. (2015, January). *Statewide tracking report*. Little Rock, AR: ACHI.

ADHS. (2016). *Arkansas SIM final progress report, Attachments Vol. 2*. Little Rock, AR: ADHS.

AHRQ. (2015, October). *Appropriate testing for children with pharyngitis: Percentage of children 3 to 18 years of age who were diagnosed with pharyngitis, dispensed an antibiotic, and received a group A streptococcus (strep) test for the episode*. NQMC:010509. Rockville, MD: AHRQ.

Arkansas Medicaid. (n.d.). *Section II: Episodes of Care*. Available at https://www.medicaid.state.ar.us/Download/provider/provdocs/Manuals/EPISODE/EPISODE_II.doc

Brown, D. W., Taylor, R., Rogers, A., Weiser, R., & Kelley, M. (2003). Antibiotic prescriptions associated with outpatient visits for acute upper respiratory tract infections among adult Medicaid recipients in North Carolina. *North Carolina Medical Journal*, *64*(4), 148–156. PMID: 12971239.

Byrd, V. L., & Dodd, A. H. (2015, August). *Assessing the usability of encounter data for enrollees in comprehensive managed care 2010–2011.* Medicaid Policy Brief. Mathematica Policy Research, Brief 22. Washington, DC: Mathematica Policy Research.

KFF. (2018, January). *Medicaid and CHIP income eligibility limits for pregnant women as a percent of the federal poverty level.* Available at https://www.kff.org/health-reform/state-indicator/medicaid-and-chip-income-eligibility-limits-for-pregnant-women-as-a-percent-of-the-federal-poverty-level/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D

Li, P., Metlay, J. P., Marcus, S. C., & Doshi, J. A. (2014). Factors associated with antimicrobial drug use in Medicaid programs. *Emerging Infectious Diseases*, *20*(5), 829–832. http://dx.doi.org/10.3201/eid2005.130493

Puhani, P. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear ‘‘difference-in-differences’’ models. *Economics Letters, 115*, 85–87.

RTI International. (2018). *State Innovation Models (SIM) Initiative Evaluation: Model Test Year Four Annual Report.* Available at https://downloads.cms.gov/files/cmmi/sim-rd1-mt-fourthannrpt.pdf

Zuckerman, I. H., Perencevich, E. N., & Harris, A. D. (2007). Concurrent acute illness and comorbid conditions poorly predict antibiotic use in upper respiratory tract infections: A cross-sectional analysis. *BMC Infectious Diseases*, *7*(1), 47. http://dx.doi.org/10.1186/1471-2334-7-47

Sub-appendix B-2. Methods for Maine Analyses

B-2.1 Methods for the Maine BHH Impact Analysis

To estimate the impact of the Behavioral Health Home (BHH) model in Maine, we conducted within-state, pre-post regression analyses using MaineCare (Medicaid) data. We present results of both descriptive trends and pre-post analyses for outcomes across four evaluation domains: (1) care coordination, (2) service utilization, (3) expenditures, and (4) quality of care. This appendix details the methods we used for this analysis.

BHH in the context of MaineCare (Maine's Medicaid program). Maine implemented its BHH model in April 2014 under the authority of a Medicaid state plan amendment. MaineCare adults with serious mental illness and children with serious emotional disturbances and who meet certain clinical need criteria (based on enrollment in certain case management or treatment programs) are eligible to participate in the model. The BHH model was implemented within MaineCare's existing fee-for-service, primary care case management delivery system. MaineCare does not have Medicaid managed care. Under the BHH model, the BHH receives a monthly capitated payment from MaineCare for each Medicaid beneficiary enrolled in the BHH, to provide all care management services for the beneficiary. BHHs are not able to bill for additional fee-for-service care management services. However, payment for any non–care management services (e.g., outpatient therapy services) are billed by BHH providers to MaineCare through the existing fee-for-service system. Furthermore, Maine has not elected to expand Medicaid to all individuals under 138 percent of the federal poverty level.

Profile of participating BHH providers. Behavioral health organizations apply to become a BHH, and participation by a behavioral health organization is voluntary. Participating organizations receive financial support from MaineCare and extensive practice transformation technical assistance from Maine's SIM Initiative partners. The number of BHHs participating in the model has been stable at 24 since the start of the BHH test period. However, BHHs may have multiple locations, and the number of locations offering BHH services has increased from an estimated 51 in 2015 to 102 in 2016. We do not have information on when various locations began participating in the model; therefore, we cannot consider rolling entry of BHHs in our analysis. Furthermore, although we received a list of participating BHHs from Maine, we do not have a list of behavioral health organizations providing services to Medicaid beneficiaries who are not participating in the BHH program. Therefore, we were unable to investigate how the group of participating BHHs differs from other behavioral health organizations providing services to Medicaid beneficiaries in Maine.

Identifying the intervention group. MaineCare provided RTI with a list of MaineCare beneficiaries assigned to BHHs since the start of the BHH initiative in April 2014 through March 2016. The list described which MaineCare enrollees were assigned to a specific BHH. However,

we did not differentiate beneficiaries by the different BHHs, and enrollees were combined to create one BHH intervention group. We used the MaineCare list to identify if a BHH enrollee had been enrolled at some point in *each* BHH test period year (Year 1, April 2014-March 2015 and Year 2, April 2015-March 2016). To be eligible for the program, MaineCare enrollees need to meet certain diagnostic and functional criteria and be in need of case management services.[4] BHHs were given the latitude to identify potential enrollees. Once identified by the BHH, MaineCare staff confirmed program eligibility of potential enrollees. Furthermore, potential BHH enrollees needed to opt into the BHH model if they were also eligible for other behavioral health services, such as Section 13 targeted case management services, and BHH enrollees could leave the model at any time. Therefore, enrollment is voluntary. Maine's focus was on enrolling individuals who were already receiving care within a behavioral health organization. We do not compare the outcomes of BHH enrollees before and after BHH implementation against a comparison group for two reasons. First, BHH providers had extensive latitude to decide which beneficiaries should be enrolled in the program, making the process of mimicking their selection decisions impossible. Second, functional assessment criteria were a factor in selection of the BHH group and these data were not available to the evaluation team in the claims.

Subpopulations. In addition to the analysis on the overall population, we conducted several subpopulation analyses: (1) children and adults separately and (2) beneficiaries enrolled in Medicaid because of disability. We chose these three subpopulations under the assumption that their physical and behavioral needs differ enough between groups that the BHH intervention may show differential impacts for these groups.

Regression-adjusted pre-post analysis. We used a pre-post design, comparing changes in the outcome variables before and during the test period for the BHH group. We used an unbalanced panel longitudinal design; that is, we used all available data for beneficiaries attributed to the BHH group in any given year. We did not restrict our analysis to beneficiaries who had continuous enrollment in Medicaid or who had continuous attribution to the BHH group over the entire test period. Because the study sample comprises individuals with behavioral health conditions, the proportion of the study sample that is eligible for Medicare was quite high, so we do not exclude individuals dually enrolled in Medicare and Medicaid. Analyses used ordinary least squares for expenditure outcomes and logistic regression for the binary care coordination, utilization, and quality outcomes. All regression analyses used clustered standard errors at the provider level to account for clustering of individuals within different BHHs. The outcome models controlled for age, gender, race, disability status, Medicare-Medicaid enrollment, length of enrollment in Medicaid, health status, urban/rural area of residence, county-level characteristics, and whether the beneficiary was attributed in one or both of the

[4] See Section 92.03 of the MaineCare Benefits Manual for details on eligibility criteria: http://www.maine.gov/sos/cec/rules/10/144/ch101/c2s092.docx

demonstration years. More information on the study outcomes is available in ***Section B-2.1.2,*** and more information on the regression model is available in ***Section B-2.1.4***.

B-2.1.1 Data sources

Medicaid (MaineCare) data. Because Maine does not have MAX data available in the Chronic Conditions Data Warehouse enclave for the BHH test period, the RTI evaluation team obtained Maine Medicaid (MaineCare) data directly from the state's data vendor, Molina Medicaid Solutions. For this analysis, we used MaineCare data from April 2011 through March 2016. These data contain demographic and enrollment information, including a monthly indicator of enrollment. The data also include facility and professional medical claims for inpatient and outpatient services and pharmaceutical claims. Monthly enrollment and claims files typically became available in the following month (e.g., March files were available in April) and were updated in subsequent months to become more complete. Molina recommended a 7-month run-out period to ensure >90 percent completion of the monthly files, and data files prior to calendar year 2015 included this run-out. Data files from calendar year 2015 through March 2016 did not incorporate this run-out period because Molina was no longer able to provide the revised data, However, RTI analyses found no significant differences in file completeness between data files provided prior to 2015 and those provided after 2015.

BHH attribution data provided by MaineCare. MaineCare provided RTI with a list of MaineCare enrollees attributed to BHHs since the start of the BHH initiative in April 2014. BHH provider-level data only included a description of which MaineCare enrollees were assigned to a specific BHH. No additional information about BHHs (e.g., when they enrolled in the BHH model, specific behavioral health providers practicing in those BHHs, or other characteristics of the BHH) was provided. This list was merged with the MaineCare claims data to identify which MaineCare beneficiaries would be assigned to the BHH group and whether the beneficiary was in the BHH group in the first program year (April 2014–March 2015), the second program year (April 2015–March 2016), or both program years.

Area Health and Resource File (AHRF). The AHRF comprises data collected by the Health Resources and Services Administration from more than 50 sources containing more than 6,000 variables related to health care access at the county level. We used information on health professions supply, poverty at the county level, age, rural/urban status, and uninsured rates from 2010–2015 to select the comparison group and to use as covariates in the analysis.

B-2.1.2 Outcome measures

Care coordination

To evaluate the impact of the BHH model on care coordination, we report the following care coordination measures. These measures were calculated annually for all eligible

beneficiaries in the BHH and Accountable Community groups and their respective comparison group.

- **Percentage of beneficiaries with any visit to a primary care provider.** This is an indicator of whether the beneficiary had at least one visit to a primary care provider reported in the medical claims file for the year, divided by the number of beneficiaries in the same year. Primary care physicians were identified using their primary taxonomy code, which was obtained from the National Plan and Provider Enumeration System (NPPES) file. A taxonomy code was considered primary where it was denoted in the NPPES file with a Y or an X. When searching for primary care visits, claims were restricted to those with HCPCS/CPT codes indicating evaluation and management visits associated with planned physician care (i.e., office visits). Both inpatient and outpatient files were included, although E&M codes used to identify physician visits should occur only in the outpatient file.
- **Percentage of beneficiaries with any visit to a specialty provider.** This is an indicator of whether the beneficiary had at least one visit to a specialty provider reported in the medical claims file for the year, divided by the number of beneficiaries in the same year. Specialty care physicians were identified using their primary taxonomy code, which was obtained from the National Plan and Provider Enumeration System (NPPES) file. A taxonomy code was considered primary where it was denoted in the NPPES file with a Y or an X. When searching for specialty care visits, claims were restricted to those with HCPCS/CPT codes indicating evaluation and management visits associated with planned physician care (i.e., office visits). Both inpatient and outpatient files were included, although E&M codes used to identify physician visits should occur only in the outpatient file.
- **Percentage of mental illness–related acute inpatient hospital admissions among patients 6 years or older as of the date of discharge with a mental health follow-up visit within 7 and 30 days.** This is the number of acute inpatient hospital admissions with a primary diagnosis for a mental health disorder (identified using ICD-9 and ICD-10 diagnosis codes specified in the Healthcare Effectiveness Data and Information Set (HEDIS) measure description) followed by a visit to a provider for a mental health outpatient visit, intensive outpatient encounter, or partial hospitalization within 7 or 30 days of discharge date (respectively), divided by total number of acute inpatient hospital admissions with a primary diagnosis for a mental disorder. Admissions followed by a readmission to an acute or other facility within 7 or 30 days are excluded from the respective denominators. Discharges because of death are also excluded from the respective denominators. This measure was defined according to the HEDIS 2016 specifications.

For both indicators, 7-Day Follow-Up and 30-Day Follow-Up, any of the following meet the criteria for a follow-up visit:

- A visit with a mental health practitioner.
- A visit to a behavioral health care facility.

- A visit to a nonbehavioral health care facility with a mental health practitioner.
- A visit to a nonbehavioral health care facility with a diagnosis of mental illness.
- Transitional care management services, where the date of service on the claim is 29 days after the date the patient was discharged with a principal diagnosis of mental illness.

The following meets the criteria for only the 30-Day Follow-Up indicator:

- Transitional care management services, where the date of service on the claim is 29 days after the date the patient was discharged with a principal diagnosis of mental illness

Utilization

Utilization measures are reported as rates per 1,000 covered lives (or discharges for with a mental health follow-up visit). For each measure, we first calculate the probability of any use. To calculate the probability, the numerator was an indicator of having had at least one event (inpatient admission or emergency department [ED] visit that did not lead to a hospitalization), and the denominator is the number of eligible beneficiaries (or discharges) in the state enrolled during the period. We multiplied the probability of use by 1,000 to obtain approximate rates of utilization per 1,000 beneficiaries. Multiplying the probability by 1,000 does not produce an exact rate of utilization per 1,000 beneficiaries because it assumes no person has more than one visit or admission per quarter. However, we concluded that this is a reasonable approximation because the majority of the population had zero or one ED visit or admission per quarter. Events are included in a period's total if discharge or service date on the claim was during the period. Reversal claims (Claim Type = 22) are excluded from utilization measures because they reflect cost adjustments and not true utilization.

- **Rate of all-cause inpatient hospitalizations (per 1,000 covered persons):** This is an indicator of whether the beneficiary had at least one admission to an acute-care hospital reported in the medical claims file for the year, divided by the number of beneficiaries in the same year. Inpatient admissions were defined as claims with a Bill Type equal to 11 or 12. Some records in the inpatient claims files may appear to be multiple admissions but are in fact transfers between facilities; these records were counted as a single admission. To roll up transfers into one acute admission, we first identified claims that had no more than 1 elapsed day between discharge date of the index claim and admission date of the subsequent claim. We then combined these claims into one record by taking the earliest admission date and latest discharge date and summing all payment amounts. This same roll-up procedure was applied to claims with overlapping or identical admission and discharge dates (i.e., claims associated with the same visit).
- **Rate of inpatient hospitalizations for mental or behavioral health–related reason (per 1,000 covered persons):** This is an indicator of whether the beneficiary had at least one behavioral health–related admission to an acute-care or psychiatric hospital

reported in the inpatient file for the year divided by the number of beneficiaries in the same year. Inpatient admissions were defined as claims with a Bill Type equal to 11 or 12, and roll-up processes (described above) were applied. These claims were subsequently subset to claims with a primary diagnosis in the Mental Health Diagnosis or Chemical Dependency HEDIS value sets.

- **Rate of ED visits that did not result in an inpatient hospital admission (per 1,000 covered persons):** This is an indicator of whether the beneficiary had at least one visit to the ED that did not result in an inpatient hospital admission divided by the number of beneficiaries in the same period. ED visits (including observation stays) were identified in the outpatient file (Bill Type ≠ 11 or 12) as visits with a revenue code equal to 0450–0459, 0981, or 0762 *or* facility type equal to 23 with procedure code equal to 99281, 99282, 99283, 99284, or 99285. If the procedure code on every line item of the ED claim equaled 70000–89999, or was equal to G0106, G0120, G0122, G0130, G0202, G0204, G0206, G0219, G0235, G0252, G0255, G0288, G0389, S8035, S8037, S8040, S8042, S8080, S8085, S8092, or S9024, and no line items had a revenue center code equal to 0762, that claim was excluded (thus excluding claims for which only radiology or pathology/laboratory services were provided unless they were observation stays). Multiple ED visits on a single day were counted as a single visit.
- **Rate of 30-day readmissions (per 1,000 discharges):** This is an indicator of whether the beneficiary had at least one acute hospitalization that occurred within 30 days following a live discharge for beneficiaries ages 18 or older for the year, divided by the number of inpatient discharges in the same year. Index hospital discharges were identified as inpatient stays with a discharge date within the given measurement period (12 months) minus 30 days from the end of the period. We counted number of instances when the beneficiary had an inpatient readmission within 30 days of the index stay discharge.

Expenditures

Weighted average expenditures were calculated on a per member per month (PMPM) basis. For each individual, PMPM expenditures were estimated as one-twelfth of their annual expenditures. Expenditures were defined as payments made by MaineCare. Averages include all individuals enrolled during the period, so that the figures also reflect the presence of individuals with zero medical costs. The payments were not risk adjusted or price standardized across geographic areas. Reversal (i.e., negative) claims were included in the calculations; however, negative payments at a beneficiary-year level were set to zero for total expenditures. Depending on the type of claim, claims were included in a period's total if the discharge, end date, or prescription fill date was during the year of interest.

- **Total:** This represents overall net payment amounts from all inpatient and outpatient (facility and professional) medical claims and all pharmacy claims. Total expenditures do not include the monthly per capita BHH payment made to the BHHs because those payments were not available in the MaineCare claims.

- **Inpatient facility:** This represents the sum of net facility payments to a hospital for covered services provided during all inpatient admissions. Inpatient claims were identified using Bill Type = 11 or 12.
- **Professional:** This represents the sum of net payments from all inpatient and outpatient professional claims. Professional claims were identified as claims for which Claim Type = Professional *and* Provider Type code corresponded with the following providers:

  02-Advanced Practice Registered Nurse
  03-Advanced Practice Registered Nursing Group
  07-Assisted Living Service Provider
  08-Attendant Care
  09-Audiologist
  10-Audiology Group
  11-Behavioral Health Clinician
  12-Behavioral Health Clinician Group
  14-Case Management Services Provider
  15-Chiropractic Group
  16-Chiropractor
  26-Dietician
  28-Early Intervention Practitioner
  36-Indian Health Services Provider
  40-Nurse
  42-Occupational Therapist
  43-Occupational Therapy Assistant
  44-Occupational-Physical Therapy Group
  45-Optician
  46-Optometrist
  49-Physical Therapist
  50-Physical Therapist Assistant
  51-Physician
  52-Physician Assistant
  54-Physician Group
  56-Podiatrist
  57-Podiatry Group
  58-Private Duty Nursing Provider
  63-Speech Language Pathologist
  64-Speech Language Pathology Group
  65-Speech Therapy Assistant
  67-Speech/Hearing Therapist Group
  69-Substance Abuse Provider
  72-Vision Services Provider Group
  73-Waiver Services Provider
  74-Individual Provider
  75-Group of Providers
  76-Facility-Agency-Organization Provider

77-Multi-Disciplinary Provider
78-Facility-Agency-Organization NR Provider
80-Counselor
81-EIM Provider
82-FP Nurse
83-FP Specialist
84-Non Medicaid Provider
85-Resp. Therapist
92-Early Childhood Provider

- **Behavioral health**: This represents the sum of net payments from all inpatient and outpatient (facility and professional) medical claims for which the primary diagnosis code was related to a mental disorder, as defined by the International Classification of Diseases, versions 9 and 10. Specifically, these codes were:
  - ICD-9: 290xx–319xx *(x can be any value or missing)*
  - ICD-10: F01xxx—F99xxx; G44209, H9325, R37, R451, R457, R480, Z87890 *(x can be any value or missing)*
- **Inpatient behavioral health:** This represents the sum of net payments from all inpatient claims for which the primary diagnosis code was related to a mental disorder, as defined by the ICD-9 and ICD-10 codes included in the behavioral health expenditure measure, described in the previous measure.
- **Prescription:** This represents the sum of net payments in the pharmacy claims files.

Quality of care

To evaluate the impact on quality of care, we report the following quality measures. Quality measures followed HEDIS 2016 specifications. The measures were calculated annually for all eligible beneficiaries in the BHH group.

- **Percentage of patients ages 18 years and older diagnosed with a new episode of major depression and treated with antidepressant medication who remained on medication treatment for at least 12 weeks or 6 months (reported as separate measures).** This is percentage of patients 18 years of age and older who were diagnosed with a new episode of major depression and treated with antidepressant medication, and who remained on an antidepressant medication treatment for 12 weeks or 6 months, respectively. Two percentages are reported:
  - *Effective Acute Phase Treatment.* This is the percentage of newly diagnosed and treated patients who remained on an antidepressant medication for at least 84 days (12 weeks).
  - *Effective Continuation Phase Treatment.* This is the percentage of newly diagnosed and treated patients who remained on an antidepressant medication for at least 180 days (6 months).

For this measure, the Intake Period was defined as the 12-month window from August 1 of the year prior to the measurement year through July 31 of the measurement year. The

Index Prescription Start Date (IPSD) was defined as the earliest prescription dispensing date for an antidepressant medication during the Intake Period. Antidepressant medications are listed in ***Table B-2-1***; specific national drug codes for these medications were identified via the 2015 list developed by the National Committee for Quality Assurance.

To identify patients for inclusion in the denominator, the patient had to be at least 18 years old; continuously enrolled in Medicaid for 3 months prior to the IPSD through 7 months following the IPSD, with no more than a 1 month lapse in coverage; and have a diagnosis for major depression (as defined by the ICD-9 and ICD-10 diagnosis codes per HEDIS measure specifications) that met at least one of the following criteria:

- An outpatient visit, intensive outpatient, encounter or partial hospitalization with any diagnosis of major depression.
- An ED visit with any diagnosis of major depression.
- An acute or nonacute inpatient claim/encounter with any diagnosis of major depression.

Patients were excluded from the denominator if they filled a prescription (as indicated by variable Date Prescription Filled) in the 105 days prior to the IPSD.

Table B-2-1. Antidepressant medications

Description	Prescription		
Miscellaneous antidepressants	• Bupropion • Vortioxetine	• Vilazodone	
Monoamine oxidase inhibitors	• Isocarboxazid • Phenelzine	• Selegiline • Tranylcypromine	
Phenylpiperazine antidepressants	• Nefazodone	• Trazodone	
Psychotherapeutic combinations	• Amitriptyline-chlordiazepoxide • Amitriptyline-perphenazine		• Fluoxetine-olanzapine
SNRI antidepressants	• Desvenlafaxine • Levomilnacipran	• Duloxetine	• Venlafaxine
SSRI antidepressants	• Citalopram • Escitalopram	• Fluoxetine • Fluvoxamine	• Paroxetine • Sertraline
Tetracyclic antidepressants	• Maprotiline	• Mirtazapine	
Tricyclic antidepressants	• Amitriptyline • Amoxapine • Clomipramine	• Desipramine • Doxepin • Imipramine	• Nortriptyline • Protriptyline • Trimipramine

SNRI = serotonin-norepinephrine reuptake inhibitor; SSRI = selective serotonin reuptake inhibitor.

To identify the numerators, we summed the Days Supply variable for all identified antidepressant medications (see ***Table B-2-1***). Days that extended beyond the treatment windows, defined below, were not counted, and overlapping prescriptions were summed.[5] Specifically:

  - For the *Effective Acute Phase Treatment* numerator, we summed Days Supply for all prescriptions where IPSD <= Date Prescription Filled <= IPSD+114 days. If this sum was at least 84 (12 weeks), the numerator was set to 1.
  - For the *Effective Continuation Phase Treatment* numerator, we summed Days Supply for all prescriptions where IPSD <= Date Prescription Filled <= IPSD+231 days. If this sum was at least 180 (6 months), the numerator was set to 1.

- **Comprehensive adult diabetes care**. This comprises two measures: the percentage of patients 18–75 years old with type 1 or type 2 diabetes who had:
  - Hemoglobin A1c (HbA1c) testing

The denominator included beneficiaries age 18 to 75 with continuous Medicaid enrollment in the measurement year with no more than a 1-month gap and identified as having type 1 or type 2 diabetes in the measurement year. Diabetes was identified using any of the following criteria:

- In the outpatient claims file (Bill Type ≠ 11 or 12): At least two outpatient visits, observation visits, ED visits, or nonacute inpatient encounters on different dates of service, with a diagnosis of diabetes.
- In the inpatient claims file (Bill Type = 11 or 12): At least one acute inpatient encounter with a diagnosis of diabetes.

The numerator for Hba1c testing is set to 1 if the beneficiary is in the denominator (as defined above) and has a procedure code in the Hba1c tests value set during the measurement year.

B-2.1.3 Population studied

Intervention (BHH) group. As described in ***Appendix B*, *Section B.1*,** MaineCare provided RTI with a list of MaineCare enrollees attributed to 24 BHHs since the start of the initiative in April 2014 through March 2016. The list included identifying information (e.g., MaineCare ID) and with which BHH the individual was associated. Individuals dually eligible for Medicare and Medicaid were not excluded from the comparison group or the BHH group. A relatively high proportion of the study sample (around 36 percent in any given analysis year) are dually enrolled in Medicare and Medicaid because they have a disability that qualifies them for

[5] The decision to sum overlapping Days Supply variables was made by the analyst team. Determining the actual time frame covered by overlapping prescriptions would require significant inference as to how beneficiaries were taking their medications and when they began taking their medication after filling the prescription, both of which are beyond the scope of information provided in the claims.

Medicare enrollment.[6] Even though we are unable to capture Medicare expenditures for those dually enrolled for this analysis, we do not exclude dually enrolled individuals to match the state's attribution process and preserve sample size. We do include a covariate for having restricted benefits in regression analyses to control for potential differences across those with and without restricted benefits. The final sample is 7,560 MaineCare enrollees.

Subpopulation analysis. As described in ***Section B-2.1.4***, we also examined a select number of outcomes for adults, children, and enrollees with disabilities. Within each analysis year, adults were defined as greater than 18 years or older, and children were defined as 18 years or younger. Enrollees were defined as disabled if the BHH enrollee's reason for Medicaid enrollment in the analysis year was associated with a MaineCare program for individuals with disabilities (e.g., supplemental security income eligible and disabled).

Table B-2-2 below shows the beneficiary characteristics by year for the BHH group.

Table B-2-2. Weighted characteristics of BHH group, 2012–2016

Characteristic	2012	2013	2014	2015	2016
N	6,559	6,782	7,096	7,306	7,386
Female (%)	57.6	57.9	57.7	57.6	57.3
Age 0 (%)	0.7	0.4	0.1	0.0	0.0
Age 1 to 18 (%)	24.2	23.6	22.5	21.1	20.2
Age 19 to 64 (%)	70.5	70.7	71.9	72.7	73.1
Age 65+ (%)	4.6	5.3	5.5	6.2	6.7
Disabled (%)	50.4	52.6	53.8	56.6	57.0
Dual Medicare eligible (%)	34.5	35.5	36.1	37.1	38.6
Non-white (%)	14.8	15.4	16.2	16.7	16.3
Race missing (%)	8.6	9.2	9.9	10.4	10.0
Continuous enrollment (%)	98.9	98.8	98.8	98.6	98.4
Total months enrolled annually	11.5	11.5	11.5	11.5	11.5
Unrestricted benefits (%)	84.5	86.7	86.9	88.0	86.6
Attributed both demonstration years (%)	56.8	57.1	57.1	56.4	53.4
Lagged CDPS	1.8	1.8	1.8	1.8	1.8
MSA (%)	65.6	66.4	66.9	67.1	67.5
Uninsured rate at county level (2013)	13.5	13.5	13.5	13.5	13.4

(continued)

[6] We are unable to determine if the disability qualifying them for Medicare enrollment is a result of the individual's behavioral health condition or a co-occurring physical condition.

Table B-2-2. Weighted characteristics of BHH group, 2012–2016 (continued)

Characteristic	2012	2013	2014	2015	2016
Median age at county level (2010)	42.0	41.9	41.9	41.9	41.9
Poverty rate at county level (2013)	14.5	14.5	14.5	14.4	14.4
Hospital beds per 1,000 persons	3.2	3.2	3.2	3.2	3.2
Physicians per 1,000 persons	1.1	1.1	1.1	1.2	1.2
Number of community mental health centers at county level	0.01	0.01	0.01	0.01	0.01

BHH = Behavioral Health Home; CDPS = Chronic Illness and Disability Payment System (larger CDPS scores correspond with a larger number of comorbidities or a more severe set of comorbidities).

B-2.1.4 Statistical analysis

Regression model

Pre-post regression model—The pre-post annual fixed-effects model used for the evaluation is shown in ***Equation B-2.1***:

$$\gamma = \alpha_0 + \sum \phi_t Q_{t,p} + \delta X + \mu \qquad \text{(B-2.1)}$$

where

y = a performance measure (e.g., total expenditures) for the i-th, in period t (i,t subscripts suppressed).

X = a vector of patient and demographic characteristics.

$Q_{t,d}$ = 0,1 indicator of the n-th or t-th calendar year in the base (b) or post (p) period (n starts counting at first baseline period, whereas t starts with first model year).

μ = error term.

Table B-2-3 illustrates the interpretation of the pre-post estimate from this model. The coefficient β1 in ***Equation B-2.1*** is the difference in the measure before and after BHH implementation, holding constant other variables in the equation. Using the annual fixed-effects model, we calculated overall estimates by taking linear combinations of the yearly estimates.

Table B-2-3. Pre-post estimate

Group	Pre period	Post period	Pre-post difference
Test	$\alpha_0 + \beta_2$	$(\alpha_0 + \beta_2) + \alpha_2$	α_2

Models for unplanned readmissions and mental health follow-ups were estimated at the annual-admission level. All other outcomes were estimated with the beneficiary year as the unit of analysis.

The outcome model for total Medicaid per beneficiary per month expenditures was estimated using ordinary least squares. To show the adjusted means in the pre- and post-periods for the BHH group, we used a linear model that allows for the calculation of means that will sum to the pre-post estimate. Although this model has strong assumptions of normality of the outcome, the OLS model still produces unbiased estimates even when the normality assumptions is violated as long as errors are uncorrelated and have a constant variance (Gauss-Markov Theorem). However, we can and do control for the correlation and variance in errors with clustered standard errors at the provider level. Additionally, the model yields estimates that are readily interpretable in dollars and do not require additional transformation.

For all other outcomes, we converted utilization counts into binary outcomes (1 = any use during the year) and used weighted logistic regression models. Count models are not appropriate because of the low occurrence of most types of utilization for individual beneficiaries in any year; however, we multiplied the marginal effect from the logistic regression models by 100 or 1,000, as appropriate, to obtain approximate rates of utilization per 100 or 1,000 beneficiaries. Multiplying the marginal effect by 100 or 1,000 does not produce an exact rate of utilization per 100 or 1,000 beneficiaries because it assumes that no person has more than one visit or admission per year. However, we concluded that this is a reasonable approximation because only a small percentage of beneficiaries had counts exceeding 1 for any of the utilization measures. For expenditure outcomes, we used weighted generalized linear models with a normal distribution and identity link.

The models for total expenditures, inpatient admissions, and ED visits were run separately for children, adults, and people with disabilities. We did not examine readmissions among these subpopulations given the small number of readmissions within each of the three groups when we subdivided the total sample into these three groups.

Control variables. In all models we controlled for the following variables:

- Age (categorical): 0, 1 to 18, 19 to 64, and 65+
- Gender
- Race (non-white and missing race)
- Enrollment because of disability
- Medicare/Medicaid enrollee

- Beneficiary's classification on the Chronic Illness and Disability Payment System (CDPS)[7]
- Number of months the beneficiary was enrolled in Medicaid during the year
- If the beneficiary stayed continuously enrolled in Medicaid during the year[8]
- If the beneficiary had full Medicaid benefits during the year
- If the beneficiary was enrolled in the BHH at some point in both test years
- County-level federal poverty level, median age, and uninsured rate
- Metropolitan status of the beneficiary's county
- County-level hospital beds, physicians, and community mental health centers per capita

Weighting and clustering. All of the regression models were estimated using weighted regressions and weighted by eligibility fraction (the fraction of the year during which the beneficiary was eligible for the analyses). In addition, standard errors were clustered at the BHH level to account for clustering of beneficiaries within BHHs.

B-2.2 Methods for the Maine AC Impact Analysis

To estimate the impact of the Accountable Community (AC) model in Maine, we conducted within-state, difference-in-differences (D-in-D) regression analyses using MaineCare (Medicaid) data. This sub-appendix details the methods of both descriptive trends and D-in-D analyses for outcomes across four evaluation domains: (1) care coordination, (2) service utilization, (3) expenditures, and (4) quality of care.

ACs in the context of MaineCare (Maine's Medicaid program). Maine implemented its AC model in August 2014 under the authority of a Medicaid state plan amendment. All MaineCare beneficiaries with 6 months of continuous or 9 months of non-continuous enrollment in the program are eligible for attribution to an AC if they are either enrolled in a Heath Home in an AC or have received a plurality of their primary care services from a primary care physician at an AC or have had three or more emergency department (ED) visits at an ED in an AC. The AC model was implemented within MaineCare's existing fee-for-service system; Maine does not

[7] The CDPS is a diagnostic classification system originally developed for states to use in adjusting capitated payments for Temporary Assistance for Needy Families and Medicaid beneficiaries with disabilities and used to predict Medicaid costs. We use the CDPS to measure beneficiary morbidity. The CDPS maps selected diagnoses and prescriptions to numeric weights. Beneficiaries with a CDPS score of 0 have no diagnoses or prescriptions that factor into creating the CDPS score. The more diagnoses a beneficiary has or the greater the severity of a particular diagnosis, the larger the CDPS score.

[8] The RTI team controlled for whether the beneficiary was continuously enrolled, with no more than a 1-month break in enrollment, from the time the beneficiary first entered the Medicaid data in the year until the end of the measurement year. However, a person can have multiple occurrences of a 1-month break in enrollment and still be considered continuously enrolled. This covariate was used to control for churning in and out of Medicaid.

have Medicaid managed care. The AC model incorporates a shared savings/losses structure, under which MaineCare will compare an AC's actual total Medicaid fee-for-service expenditures to its financial benchmark expenditures. The benchmark is based upon the AC's historical total fee-for-service expenditures for its attributed MaineCare population. If the AC spends less than the benchmark, it will receive shared savings payments from MaineCare, but if it spends more, it will pay money back to MaineCare in the form of shared losses. An AC will receive more in savings from MaineCare or owe less to MaineCare if it meets certain targets on select quality measures. Also, this arrangement is predicated upon the AC's choice of Model; under Model I, ACs will only share in savings, and under Model II, ACs will share in both savings and losses. All four ACs are currently operating under Model I.

Profile of participating AC providers. There are four ACs operating in Maine: Beacon Health LLC (1,402 enrollees[9]), Maine Health Accountable Care Organization (ACO) (5,620 enrollees), Kennebec Region Health Alliance (11,899 enrollees), and Community Care Partnership of Maine (29,650 enrollees). ACs are integrated provider organizations that offer care coordination and administrative support to providers to ensure comprehensive primary, acute, and chronic health care services are made available to an attributed population. Each AC includes a lead entity, such as a regional health system, that forms contractual partnerships with providers. ACs must contract with providers that serve patients with chronic conditions (such as a Health Home [HHs]), developmental disabilities, and behavioral health needs. As of July 2017, the time of the most recently available data, these ACs had a total of 80 primary care practices participating as well as seven EDs.[10] There were 55,314 MaineCare beneficiaries enrolled in these ACs, or approximately 19 percent of the total MaineCare population. Enrollment has increased over time as some ACs have expanded their provider networks. From Year 1 to Year 2, the number of primary are practices increased from 28 to 66, and from Year 2 to Year 3, the number increased from 66 to 80 practices.

Identifying the intervention and comparison group. MaineCare provided RTI with a list of MaineCare beneficiaries assigned to each AC from August 2014 through July 2016. Enrollees in each AC were combined to create one AC intervention group. To create a comparison group, we identified all MaineCare beneficiaries who were not attributed to an AC and mimicked the AC three-step attribution process to ensure that members from the comparison group were selected in the same way AC enrollees were selected. More information on the attribution process can be found in ***Section B-2.2.1***.

Subpopulations. In addition to the analysis on the overall population, we conducted subpopulation analyses for children, adults, individuals with behavioral health conditions, and

[9] Enrollment numbers for each AC reflect the total number of nonduplicated enrollees during the time period examined for this impact analysis.
[10] See http://www.maine.gov/dhhs/oms/pdfs_doc/vbp/AC/Accountable-Communities-Providers-and-Number-of-Members.pdf

individuals also enrolled in a MaineCare HH. For adults and children, we assumed that their health care needs differ enough that the AC intervention may show differential impacts. Because of Maine's focus on improving the quality of behavioral health care within Behavioral HHs and HH in conjunction with AC requirements to include behavioral health providers in their networks, we expected that ACs may focus some care coordination and quality improvement activities on high-cost, high-users with behavioral health conditions and, therefore, examined this group. The AC program was designed to complement existing value-based delivery strategies in MaineCare, including the HH program—a program similar to the medical home but for individuals with multiple morbidities. Because many primary care practices participating in ACs are HHs, some HH enrollees are exposed to two overlapping interventions. Therefore, we investigated if there were differential impacts for AC members also enrolled in the HH program compared to comparison group of members not enrolled in an AC but enrolled in an HH.

Balancing AC and comparison groups. Following comparison group selection, we constructed a person-level propensity score weight to balance the AC group and comparison group on select observed individual characteristics. We used weighting to apply propensity scores to the analysis, as opposed to other methods (e.g., matching), to retain sample size and produce less-biased estimates for binary outcomes. After propensity score weighting, the standardized differences between the weighted comparison group means and AC group means were under the standard 10 percent threshold. More information on propensity score weighting is available in ***Section B-2.4***.

Study design. We used a D-in-D design, comparing changes in the outcome variables before and during the AC period for the AC group with changes in the outcomes before and during the AC period for the comparison group. An unbalanced panel longitudinal design was used; that is, all available data for beneficiaries attributed to the intervention and comparison groups in any given year were used.

Statistical approach. Analyses used ordinary least squares (OLS) for expenditure outcomes and logistic regression for the binary care coordination, utilization, and quality outcomes. All regression analyses used clustered standard errors at the provider organization level to account for clustering of beneficiaries within the provider. The outcome models controlled for age, gender, race, disability status, Medicare-Medicaid enrollment, health status, urban/rural area of residence, beneficiary attribution process characteristics, county-level characteristics, and length of enrollment in Medicaid. More information on the study outcomes is available in ***Section B-2.2***, and more information on the regression model is available in ***Section B-2.5.1***.

B-2.2.1 Data sources

Medicaid (MaineCare) data. Because Maine does not have Medicaid Analytic eXtract data available in the Chronic Conditions Data Warehouse enclave for the AC period, the RTI

evaluation team obtained Maine Medicaid (MaineCare) data directly from the state's data vendor, Molina Medicaid Solutions. For this analysis, we used MaineCare data from April 2011 through March 2016. These data contain demographic and enrollment information, including a monthly indicator of enrollment. The data also include facility and professional medical claims for inpatient and outpatient services and pharmaceutical claims. Monthly enrollment and claims files typically became available in the following month (e.g., March files were available in April) and were updated in subsequent months to become more complete. Molina recommended a 7-month run-out period to ensure >90 percent completion of the monthly files, and data files prior to calendar year 2015 included this run-out. Data files from calendar year 2015 through March 2016 did not incorporate this run-out period because Molina was no longer able to provide the revised data. However, RTI analyses found no significant differences in file completeness between data files provided prior to 2015 and those provided after 2015.

AC attribution data provided by MaineCare. MaineCare provided RTI with a list of MaineCare enrollees attributed to ACs since the start of the initiative in August 2014. Attribution files included which MaineCare enrollees were attributed to which AC in each program year (Year One: August 2014–July 2015 or Year Two: August 2015–July 2016). The data also included the method by which the enrollee was attributed (i.e., because the enrollee was enrolled in an HH affiliated with an AC, had a plurality of primary care visits with an AC-affiliated primary care provider, or had a plurality of ED visits to an AC-affiliated ED). Sixty-one percent of AC enrollees included in this analysis were attributed through plurality of primary care visits, 38 percent through HH enrollment, and 0.7% through ED visits.

Area Health and Resource File (AHRF). The AHRF comprises data collected by the Health Resources and Services Administration from more than 50 sources containing more than 6,000 variables related to health care access at the county level. We used information on health professions supply, poverty at the county level, age, rural/urban status, and uninsured rates from 2010–2015 to select the comparison group and to use as covariates in the analysis.

B-2.2.2 Outcome measures

Care coordination

To evaluate the impact of the AC model on care coordination, we report the following care coordination measures. These measures were calculated annually for all eligible beneficiaries in the AC group and the comparison group.

- **Probability of having any visit to a primary care provider:** This is an indicator of whether the beneficiary had at least one visit to a primary care provider reported in the medical claims file for the year, divided by the number of beneficiaries in the same year. Primary care physicians were identified using their primary taxonomy code, which was obtained from the National Plan and Provider Enumeration System (NPPES) file. A taxonomy code was considered primary where it was denoted in the

NPPES file with a Y or an X. When searching for primary care visits, claims were restricted to those with Healthcare Common Procedure Coding System (HCPCS)/Current Procedural Terminology (CPT) codes indicating evaluation and management (E&M) visits associated with planned physician care (i.e., office visits). Both inpatient and outpatient files were included, although E&M codes used to identify physician visits should occur only in the outpatient file.

- **Probability of having any visit to a specialty provider:** This is an indicator of whether the beneficiary had at least one visit to a specialty provider reported in the medical claims file for the year, divided by the number of beneficiaries in the same year. Specialty care physicians were identified using their primary taxonomy code, which was obtained from the NPPES file. A taxonomy code was considered primary where it was denoted in the NPPES file with a Y or an X. When searching for specialty care visits, claims were restricted to those with HCPCS/CPT codes indicating E&M visits associated with planned physician care (i.e., office visits). Both inpatient and outpatient files were included, although E&M codes used to identify physician visits should occur only in the outpatient file.
- **Percentage of mental illness-related acute inpatient hospital admissions among patients 6 years or older as of the date of discharge with a mental health follow-up visit within 7 and 30 days:** This is the number of acute inpatient hospital admissions with a primary diagnosis for a mental health disorder (identified using International Classification of Diseases-9 [ICD-9] and ICD-10 diagnosis codes specified in the Healthcare Effectiveness Data and Information Set [HEDIS] measure description) followed by a visit to a provider for a mental health outpatient visit, intensive outpatient encounter or partial hospitalization within 7 or 30 days of discharge date (respectively), divided by total number of acute inpatient hospital admissions with a primary diagnosis for a mental disorder. Admissions followed by a readmission to an acute or other facility within 7 or 30 days are excluded from the respective denominators. Discharges because of death are also excluded from the respective denominators. This measure was defined according to the HEDIS 2016 specifications.

For both indicators, 7-Day Follow-Up and 30-Day Follow-Up, any of the following meet the criteria for a follow-up visit:

- A visit with a mental health practitioner.
- A visit to a behavioral health care facility.
- A visit to a nonbehavioral health care facility with a mental health practitioner.
- A visit to a nonbehavioral health care facility with a diagnosis of mental illness.

The following meets the criteria for only the 30-Day Follow-Up indicator:

- Transitional care management services, where the date of service on the claim is 29 days after the date the patient was discharged with a principal diagnosis of mental illness

Utilization

Utilization measures are reported as rates per 1,000 covered lives (or discharges for with a mental health follow-up visit). For each measure, we first calculated the probability of any use. To calculate the probability, the numerator is an indicator of having had at least one event (inpatient admission or ED visit that did not lead to a hospitalization), and the denominator is the number of eligible beneficiaries (or discharges) in the state enrolled during the period. The probability of use was multiplied by 1,000 to obtain approximate rates of utilization per 1,000 beneficiaries. Multiplying the probability by 1,000 does not produce an exact rate of utilization per 1,000 beneficiaries because it assumes no person has more than one visit or admission per year. However, we concluded that this is a reasonable approximation because the majority of the population had zero or one ED visit or admission per year. Events are included in a period's total if discharge or service date on the claim was during the period. Reversal claims (Claim Type = 22) are excluded from utilization measures because they reflect cost adjustments and not true utilization.

- **Probability of having any inpatient use:** This is an indicator of whether the beneficiary had at least one admission to an acute-care hospital reported in the medical claims file for the year, divided by the number of beneficiaries in the same year. Inpatient admissions were defined as claims with a Bill Type equal to 11 or 12. Some records in the inpatient claims files may appear to be multiple admissions but are in fact transfers between facilities; these records were counted as a single admission. To roll up transfers into one acute admission, we first identified claims that had no more than 1 elapsed day between the discharge date of the index claim and the admission date of the subsequent claim. Then, these claims were combined into one record by taking the earliest admission date and latest discharge date and summing all payment amounts. This same roll-up procedure was applied to claims with overlapping or identical admission and discharge dates (i.e., claims associated with the same visit).
- **Probability of having any inpatient visits for behavioral health related reasons:** This is an indicator of whether the beneficiary had at least one behavioral health-related admission to an acute care or psychiatric hospital reported in the inpatient file for the year, divided by the number of beneficiaries in the same year. Inpatient admissions were defined as claims with a Bill Type equal to 11 or 12, and roll-up processes (described above) were applied. These claims were subsequently subset to claims with a primary diagnosis in the Mental Health Diagnosis or Chemical Dependency HEDIS value sets.
- **Probability of having any ED visits that did not lead to a hospitalization (outpatient ED) use:** This is an indicator of whether the beneficiary had at least one visit to the ED that did not result in an inpatient hospital admission, divided by the number of beneficiaries in the same period. ED visits (including observation stays) were identified in the outpatient file (Bill Type ≠ 11 or 12) as visits with a revenue code equal to 0450–0459, 0981, or 0762 *or* facility type equal to 23 with procedure code equal to 99281, 99282, 99283, 99284, or 99285. If the procedure code on every

line item of the ED claim equaled 70000–89999, or was equal to G0106, G0120, G0122, G0130, G0202, G0204, G0206, G0219, G0235, G0252, G0255, G0288, G0389, S8035, S8037, S8040, S8042, S8080, S8085, S8092, or S9024, and no line items had a revenue center code equal to 0762, that claim was excluded (thus excluding claims for which only radiology or pathology/laboratory services were provided, unless they were observation stays). Multiple ED visits on a single day were counted as a single visit.

- **Probability of having a 30-day readmission:** This is an indicator of whether the beneficiary had at least one acute hospitalization that occurred within 30 days following a live discharge for beneficiaries ages 18 or older for the year, divided by the number of inpatient discharges in the same year. Index hospital discharges were identified as inpatient stays with a discharge date within the given measurement period (12 months) minus 30 days from the end of the period. We counted the number of instances when the beneficiary had an inpatient readmission within 30 days of the index stay discharge.

Expenditures

Weighted average expenditures were calculated on a per member per month (PMPM) basis. For each individual, PMPM expenditures were estimated as one twelfth of their annual expenditures. Expenditures were defined as payments made by MaineCare. Averages include all individuals enrolled during the period, so that the figures also reflect the presence of individuals with zero medical costs. The payments were not risk adjusted or price standardized across geographic areas. Reversal (i.e., negative) claims were included in the calculations; however, negative payments at the beneficiary-year level were set to zero for total expenditures. Depending on the type of claim, claims were included in a period's total if the discharge, end date, or prescription fill date was during the year of interest.

- **Total:** This represents amounts paid from all inpatient and outpatient (facility and professional) medical claims and all pharmacy claims. This does not include any shared savings payments to providers for meeting AC-related cost and quality targets.
- **Inpatient facility:** This represents the sum of net facility payments to a hospital for covered services provided during all inpatient admissions. Inpatient claims were identified using Bill Type = 11 or 12.
- **Professional:** This represents the sum of net payments from all inpatient and outpatient professional claims. Professional claims were identified as claims for which Claim Type = Professional *and* the Provider Type code corresponded with the following providers:

  02-Advanced Practice Registered Nurse

  03-Advanced Practice Registered Nursing Group

  07-Assisted Living Service Provider

  08-Attendant Care

09-Audiologist
10-Audiology Group
11-Behavioral Health Clinician
12-Behavioral Health Clinician Group
14-Case Management Services Provider
15-Chiropractic Group
16-Chiropractor
26-Dietician
28-Early Intervention Practitioner
36-Indian Health Services Provider
40-Nurse
42-Occupational Therapist
43-Occupational Therapy Assistant
44-Occupational-Physical Therapy Group
45-Optician
46-Optometrist
49-Physical Therapist
50-Physical Therapist Assistant
51-Physician
52-Physician Assistant
54-Physician Group
56-Podiatrist
57-Podiatry Group
58-Private Duty Nursing Provider
63-Speech Language Pathologist
64-Speech Language Pathology Group
65-Speech Therapy Assistant
67-Speech/Hearing Therapist Group
69-Substance Abuse Provider
72-Vision Services Provider Group
73-Waiver Services Provider
74-Individual Provider
75-Group of Providers

76-Facility-Agency-Organization Provider

77-Multi-Disciplinary Provider

78-Facility-Agency-Organization NR Provider

80-Counselor

81-EIM Provider[11]

82-FP Nurse[12]

83-FP Specialist

84-Non-Medicaid Provider

85-Resp. Therapist

92-Early Childhood Provider

- **Behavioral health:** This represents the sum of net payments from all inpatient and outpatient (facility and professional) medical claims for which the primary diagnosis code was related to a mental disorder, as defined by ICD-9 and ICD-10. Specifically, these codes were as follows:
  - ICD-9: 290xx–319xx (*x can be any value or missing*)
  - ICD-10: F01xxx—F99xxx; G44209, H9325, R37, R451, R457, R480, Z87890 *(x can be any value or missing)*
- **Inpatient behavioral health:** This represents the sum of net payments from all inpatient claims for which the primary diagnosis code was related to a mental disorder, as defined by the ICD-9 and ICD-10 codes included in the behavioral health expenditure measure, as described in the previous measure.
- **Prescription:** This represents the sum of net payments in the pharmacy claims files.

Quality of care

To evaluate the impact on quality of care, we report the following quality measures. Quality measures followed HEDIS 2016 specifications. The measures were calculated annually for all eligible beneficiaries in the AC and comparison groups.

- **Percentage of patients ages 18 years and older diagnosed with a new episode of major depression and treated with antidepressant medication who remained on medication treatment for at least 12 weeks or 6 months (reported as separate measures):** This is the percentage of patients 18 years of age and older who were diagnosed with a new episode of major depression and treated with antidepressant medication and who remained on an antidepressant medication treatment for 12 weeks or 6 months, respectively. Two percentages are reported:

[11] EIM = Elderly Independence of Maine.
[12] FP = family practice

- *Effective Acute Phase Treatment:* This is the percentage of newly diagnosed and treated patients who remained on an antidepressant medication for at least 84 days (12 weeks).
- *Effective Continuation Phase Treatment:* This is the percentage of newly diagnosed and treated patients who remained on an antidepressant medication for at least 180 days (6 months).

For this measure, the Intake Period was defined as the 12-month window from August 1 of the year prior to the measurement year through July 31 of the measurement year. The Index Prescription Start Date (IPSD) was defined as the earliest prescription dispensing date for an antidepressant medication during the Intake Period. Antidepressant medications are listed in ***Table B-2-4***; specific national drug codes for these medications were identified via the 2015 list developed by the National Committee for Quality Assurance.

To identify patients for inclusion in the denominator, the patient had to be at least 18 years old; continuously enrolled in Medicaid for 3 months prior to the IPSD through 7 months following the IPSD, with no more than a 1 month lapse in coverage; and have a diagnosis for major depression (as defined by the ICD-9 and ICD-10 diagnosis codes per HEDIS measure specifications) that met at least one of the following criteria:

- An outpatient visit, intensive outpatient visit, encounter or partial hospitalization with any diagnosis of major depression.
- An ED visit with any diagnosis of major depression.
- An acute or nonacute inpatient claim/encounter with any diagnosis of major depression.

Patients were excluded from the denominator if they filled a prescription (as indicated by variable Date Prescription Filled) in the 105 days prior to the IPSD.

To identify the numerators, we summed the Days Supply variable for all identified antidepressant medications (see ***Table B-2-4***). Days that extended beyond the treatment windows, defined below, were not counted, and overlapping prescriptions were summed.[13] Specifically:

- For the *Effective Acute Phase Treatment* numerator, we summed Days Supply for all prescriptions where IPSD <= Date Prescription Filled <= IPSD+114 days. If this sum was at least 84 (12 weeks), the numerator was set to 1.
- For the *Effective Continuation Phase Treatment* numerator, we summed Days Supply for all prescriptions where IPSD <= Date Prescription Filled <= IPSD+231 days. If this sum was at least 180 (6 months), the numerator was set to 1.

[13] The decision to sum overlapping Days Supply variables was made by the analyst team. Determining the actual time frame covered by overlapping prescriptions would require significant inference as to how beneficiaries were taking their medications and when they began taking their medication after filling the prescription, both of which are beyond the scope of information provided in the claims.

Table B-2-4. Antidepressant medications

Description	Prescription		
Miscellaneous antidepressants	• Bupropion • Vortioxetine	• Vilazodone	
Monoamine oxidase inhibitors	• Isocarboxazid • Phenelzine	• Selegiline • Tranylcypromine	
Phenylpiperazine antidepressants	• Nefazodone	• Trazodone	
Psychotherapeutic combinations	• Amitriptyline-chlordiazepoxide • Amitriptyline-perphenazine		• Fluoxetine-olanzapine
SNRI antidepressants	• Desvenlafaxine • Levomilnacipran	• Duloxetine	• Venlafaxine
SSRI antidepressants	• Citalopram • Escitalopram	• Fluoxetine • Fluvoxamine	• Paroxetine • Sertraline
Tetracyclic antidepressants	• Maprotiline	• Mirtazapine	
Tricyclic antidepressants	• Amitriptyline • Amoxapine • Clomipramine	• Desipramine • Doxepin • Imipramine	• Nortriptyline • Protriptyline • Trimipramine

SNRI = serotonin-norepinephrine reuptake inhibitor; SSRI = selective serotonin reuptake inhibitor.

- **HbA1c testing:** The percentage of patients 18–75 years old with type 1 or type 2 diabetes who had:
  - Hemoglobin A1c (HbA1c) testing.

The denominator included beneficiaries aged 18 to 75 with continuous Medicaid enrollment in the measurement year with no more than a 1-month gap and identified as having type 1 or type 2 diabetes in the measurement year. Diabetes was identified using any of the following criteria:

- In the outpatient claims file (Bill Type ≠ 11 or 12): At least two outpatient visits, observation visits, ED visits, or nonacute inpatient encounters on different dates of service, with a diagnosis of diabetes.
- In the inpatient claims file (Bill Type = 11 or 12): At least one acute inpatient encounter with a diagnosis of diabetes.

The numerator for HbA1c testing is set to 1 if the beneficiary is in the denominator (as defined above) and has a procedure code in the HbA1c Tests value set during the measurement year.

B-2.2.3 Population studied

Intervention (AC) group. As described in the introduction of this sub-appendix, MaineCare provided RTI with a list of MaineCare beneficiaries assigned to ACs since the start of the AC initiative in August 2014 through July 2016. The list described which MaineCare

enrollees were assigned to a specific AC. We used this list to identify if an AC enrollee had been enrolled at some point in each of the first two AC test period years: Year 1, August 2014–July 2015 and Year 2, August 2015–July 2016. Individuals dually eligible for Medicare and Medicaid were not excluded from the comparison group or the AC group. A relatively high proportion of the study sample (approximately 18 percent of the AC group and 28 percent of the comparison group in any given analysis year) are dually enrolled in Medicare and Medicaid. Even though we was unable to capture Medicare expenditures for those dually enrolled for this analysis, dually enrolled individuals were not excluded because the state allowed them to participate in the AC initiative and because we want to preserve sample size. A covariate for dual enrollment was included in the regression analyses to control for potential differences across those with and without Medicare coverage. Further, we did not exclude individuals who had less than full MaineCare benefits (i.e., restricted benefits). Approximately 10 percent of the AC group has restricted benefits, and we did not exclude them because the state allowed them to participate in the AC initiative. However, a covariate for having restricted benefits in regression analyses was included to control for potential differences across those with and without restricted benefits. The final sample included 53,019 MaineCare enrollees attributed to an AC.

Subpopulation analysis. We also examined a select number of outcomes for adults, children, individuals with behavioral health conditions, and individuals enrolled in an HH. Within each analysis year, adults were defined as 18 years or older, and children were defined as 18 years or younger. Enrollees aged 18 years or older were defined as having a behavioral health condition if they had a behavioral health primary diagnosis in the year prior to entering the AC intervention group or comparison group. Enrollees were defined as being enrolled in an HH based on enrollment lists provided by MaineCare that denoted when HH enrollment began and ended.

B-2.2.4 Comparison group and propensity score weighting

For the impact analysis, we used a pre-post comparison group design, in which the comparison group provides an estimate of what would have happened among MaineCare AC beneficiaries absent the MaineCare AC. The difference in the changes over time from the pre-period to the intervention period between AC beneficiaries and their comparison group provides an estimate of the impact of the MaineCare AC. The comparison group should be similar to the intervention group on all relevant dimensions (e.g., demographic, socioeconomic, political, regulatory, and health and health systems) except for the policy change being tested.

The following section presents the procedures used to select the comparison group for the MaineCare AC in Maine.

Selection of the comparison group

Attribution to an AC is a three-step process involving HHs, primary care visits, and ED visits, so we applied the same criteria to select a comparison group. Among MaineCare enrollees

not enrolled in an AC, the team first selected beneficiaries with 6 months of continuous Medicaid eligibility or 9 months of non-continuous eligibility. Then, beneficiaries enrolled in an HH practice that is not part of an AC were selected. Among members not selected through the HH criteria, we then selected members who had a plurality of primary care visits with a primary care provider that was not part of an AC. Among members not selected through HH or plurality of primary care visits, we selected members who had three or more ED visits at a hospital that was not part of an AC. A comparison group was selected for each of the two post-implementation years under study. Similar to the AC intervention group, individuals enrolled in both Medicare and Medicaid and individuals who had restricted Medicaid benefits were retained. Although the ACs do not have strict geographic boundaries, their attributed Medicaid enrollees are more likely to be from areas clustered around the AC's participating PCPs, so comparison group enrollees may be more likely to reside in areas farther from AC participating practices. To the extent that there is geographic variation in health care use, we could introduce bias. To mitigate this risk, the comparison group was restricted to individuals residing in the same zip codes as AC enrollees. The final comparison group sample included 199,014 MaineCare beneficiaries.

Calculation of person-level weights

To balance the population characteristics for the claims-based analyses, we estimated propensity scores for all individuals from the comparison group. A propensity score is the probability that an individual is in the intervention group rather than the comparison group.

The objective of propensity score modeling is to create a weighted comparison group with characteristics equivalent to those for the AC population. To the extent that these characteristics are correlated with expenditure, utilization, and quality outcomes, propensity weighting will also help balance pre-intervention levels of the outcomes.

There are other methods to apply propensity scores to an analysis. Aside from weighting, one frequently used method is matching, whereby an intervention beneficiary is matched to a comparison group beneficiary who has a similar propensity score. After considering this method, we decided not to pursue matching for several reasons. First, propensity score weighting has been shown to produce less-biased estimates, less modeling error (e.g., mean squared error, type 1 error), and more-accurate variance estimation and confidence intervals when modeling dichotomous outcomes; this analysis includes many dichotomous utilization and quality of care outcomes. Second, matching may exclude many comparison group beneficiaries from the analysis if a good match cannot be found. Weighting has the advantage of preserving sample size.

Person-level characteristics

The initial step in the process was to select person-level characteristics to be used in each propensity score model. ***Table B-2-5*** shows the characteristics we used grouped by whether they control for demographics, enrollment, attribution, or beneficiary health status. A comprehensive set of characteristics of providers was not provided in the MaineCare claims data, so we did not include provider-level characteristics in creating propensity scores. Because of this limited information available in claims data, the team considered also including county-level characteristics to control for geographic characteristics, such as physician supply, to account for potential differences in access to care or other geographic differences. However, we found that there was little variation in county-level characteristics, which made balancing on these variables difficult. Therefore, to optimize the balance and avoid extreme weights, county-level covariates were excluded from the propensity score model. However, we did control for county-level characteristics in the outcome model and included them in the assessment of covariate balance in ***Tables B-2-6*** to ***B-2-10***.

Table B-2-5. Covariates for propensity score logistic regressions

Covariates
Individual level sociodemographic characteristics
Gender
Age and Age squared
Enrolled in Medicaid due to disability
Medicare-Medicaid enrollee
CDPS risk score
Race
Months of Medicaid enrollment
Continuously enrolled in Medicaid
Receives full Medicaid benefits
Attribution characteristics
Enrolled in an AC for 2 years
Method of attribution to the AC or the comparison group (i.e., through visits to a primary care provider, HH enrollment, or ED visits)
Health care utilization characteristics
Total annual Medicaid payments, $
Inpatient admissions per 1,000 population
ED visits per 1,000 population

AC = Accountable Community; CDPS = Chronic Illness and Disability Payment System; ED = emergency department; HH = health home.

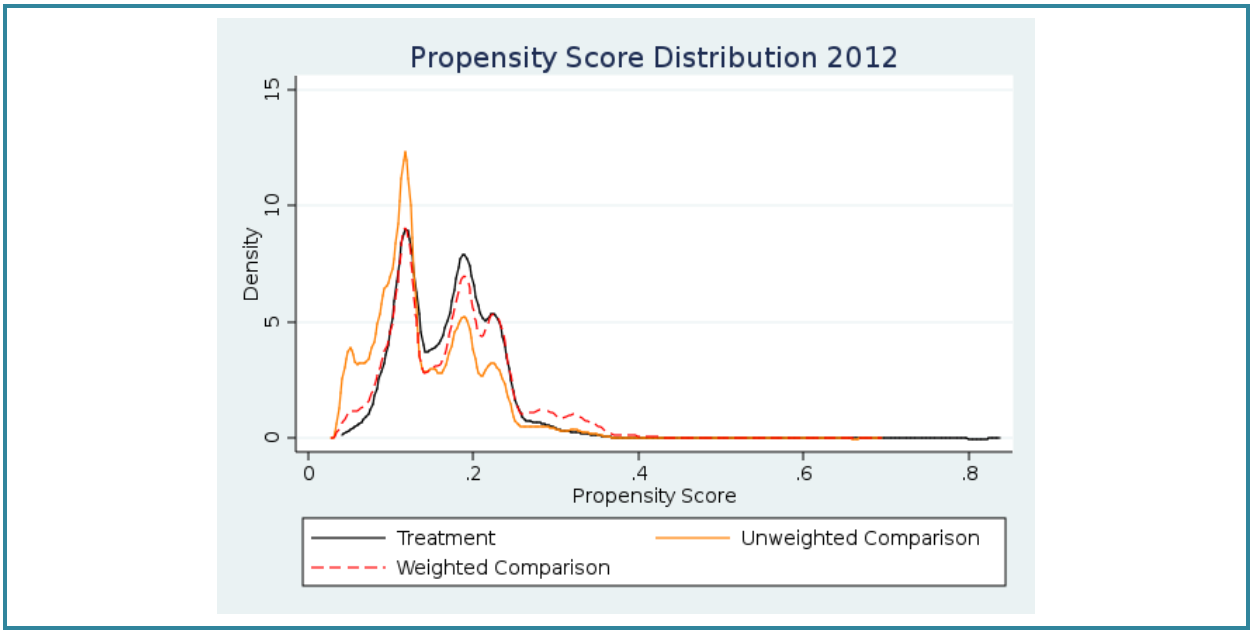
Estimation and weighting procedures

Using the characteristics listed in ***Table B-2-5***, we estimated propensity models by logistic regression, in which the outcome was one for beneficiaries attributed to the AC group and zero for the comparison group. Separate models were estimated for August 2011–July 2012, August 2012–July 2013, August 2013–July 2014, August 2014–July 2015, and August 2015–July 2016. We set propensity weights to one for all individuals in the intervention group. The propensity weight for a comparison individual was a function of his or her predicted propensity score, where weight = p/(1-p), and p is the predicted propensity. Weights that were either less than 0.05 at 0.05 or greater than 20 at 20 were trimmed. No weights had to be trimmed at 20, and depending on the year, 5,000–17,000 observations had their weights trimmed at 0.05.

B-2.2.5 Propensity model evaluation

We evaluated several aspects of the propensity score models. First, we examined plots of predicted probabilities to ensure sufficient overlap in the distributions of the intervention and comparison groups. This feature, known as common support, is critical because it provides the basis for inferring effects from group comparisons (***Figure B-2-1*** to ***B-2-5***).

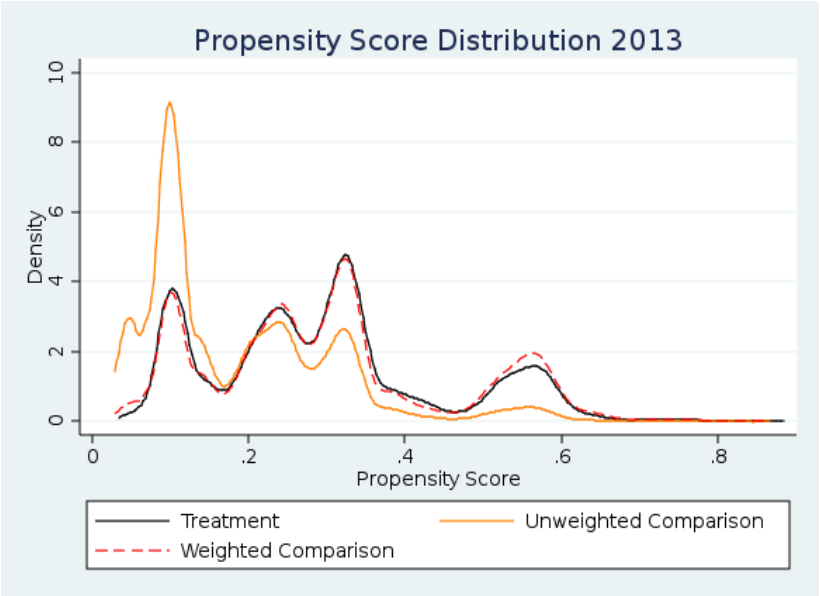
Figure B-2-1. Weighted and unweighted propensity score density plots for the AC and comparison groups[14], August 2011–July 2012



AC = Accountable Community.

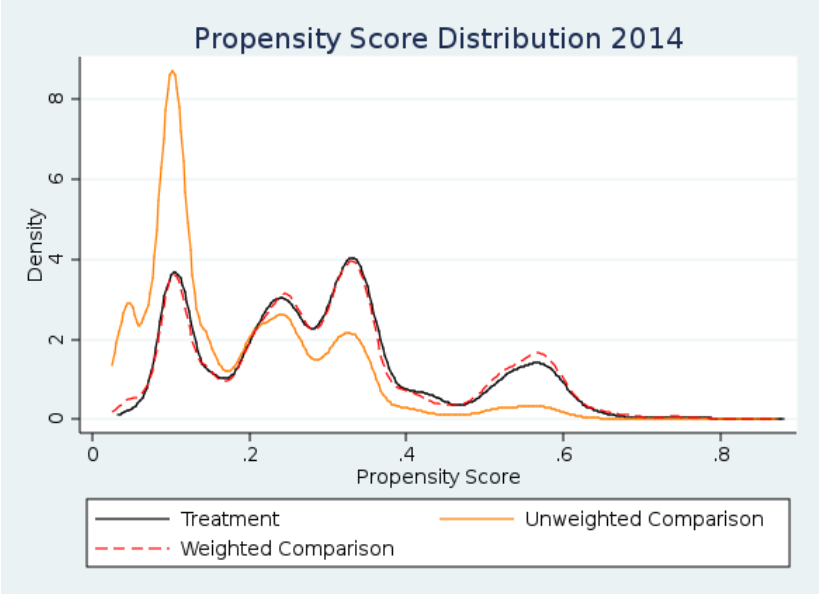
[14] In ***Figures B-2-1*** through ***B-2-5***, the Treatment lines represent those in the AC group.

Figure B-2-2. Weighted and unweighted propensity score density plots for the AC and comparison groups, August 2012–July 2013



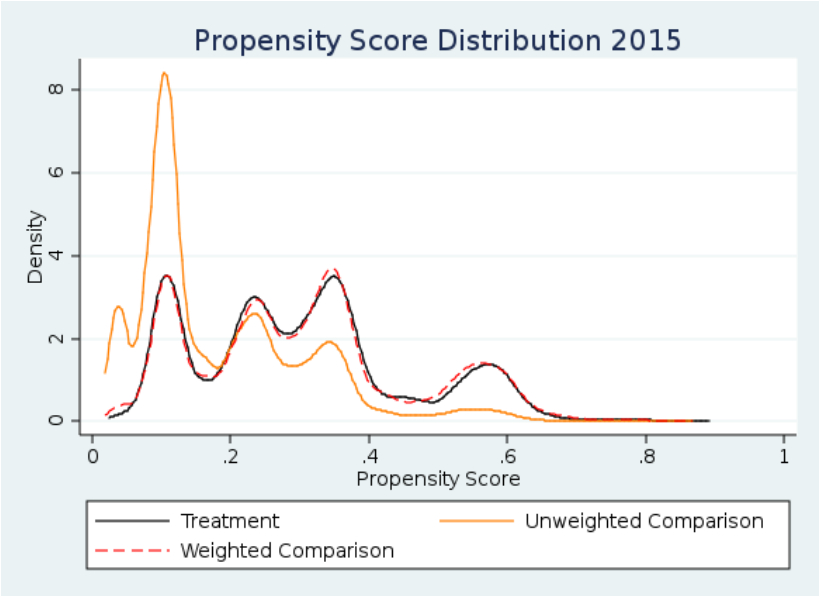
AC = Accountable Community.

Figure B-2-3. Weighted and unweighted propensity score density plots for the AC and comparison groups, August 2013–July 2014



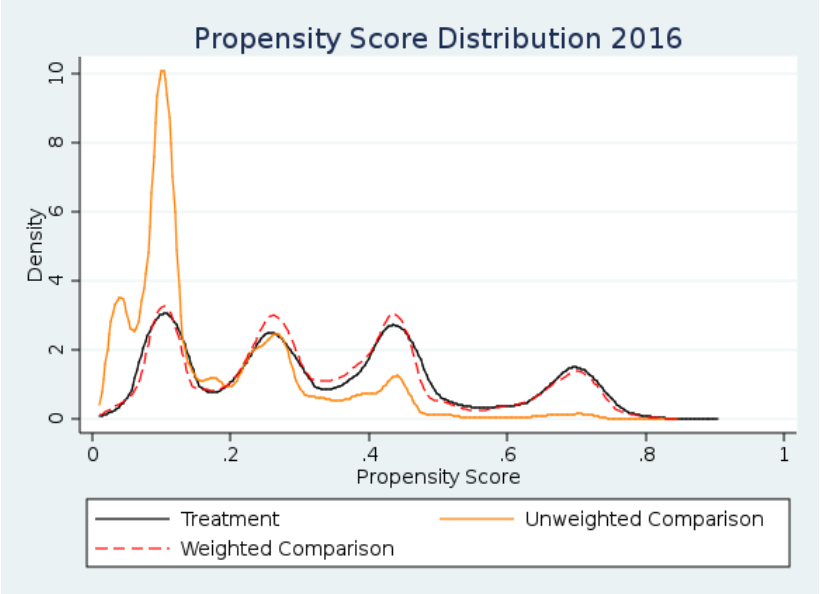
AC = Accountable Community.

Figure B-2-4. Weighted and unweighted propensity score density plots for the AC and comparison groups, August 2014–July 2015



AC = Accountable Community.

Figure B-2-5. Weighted and unweighted propensity score density plots for the AC and comparison groups, August 2015–July 2016



AC = Accountable Community.

In all years, we found that the comparison group passed the common support assumption ($P(D = 1|X)>0$) for almost the entire range of the intervention group's propensity scores. The only exceptions were in the uppermost percentiles of the intervention group's distribution (above the 99th percentile). These plots provide ample evidence that the common support assumption is upheld.

Second, we compared the logistic results of the models to see which variables had the greatest impact on the propensity score weights. AC enrollees were younger and had higher annual Medicaid expenditures relative to the comparison group. They were also more likely to be enrolled in the MaineCare HH program and have full Medicaid benefits, and they were less likely to be enrolled in the AC for 2 years.

Finally, unweighted and propensity-weighted means for the characteristics in the model were compared. As expected, we found that, after weighting, the comparison group means were within a few percentage points of the values for the AC group.

Tables B-2-6 to ***B-2-10*** show unweighted and propensity score-weighted means/proportions for August 2011–July 2016. There are notable group differences in the unweighted samples in several covariates, including age, reason for attribution, Medicare-Medicaid enrollment, total annual Medicaid expenditures, receipt of full benefits, and enrollment in the AC or comparison group for 2 years. Post-weighting, the difference between groups based on these covariates is substantially mitigated, as evidenced by the minimized standardized differences. After weighting, the county-level variables still had large standardized differences. This results from little variation in county-level characteristics among the study sample. Even though the standardized differences are large, the weighted means for these county-level variables are similar between groups.

Propensity model evaluation for subpopulation

In addition to the overall model, we also evaluated common support graphs and standardized differences of the propensity score models for the subpopulation analyses for adults, children, individuals with behavioral health conditions, and individuals also enrolled in an HH. Overall, as for the full study sample, most covariates could be balanced for relatively well, and in cases where standardized differences between groups were large even after weighting, the comparison group means were within a few percentage points of the values for the AC group, indicating small absolute differences.

Table B-2-6. Unweighted and weighted means and standardized differences, Medicaid AC and comparison groups, August 2011–July 2012

Characteristic	Unweighted			Weighted			
	Medicaid AC group	Comparison group	Standardized difference[a]	Medicaid AC group	Comparison group	Standardized difference[a]	p-value
N	23,893	140,399		23,893	24,062		
Female (%)	56.9	59.0	4.2	56.9	56.8	0.2	0.83
Age, mean	26.5	31.8	22.9	26.5	26.4	0.4	0.70
Age squared, mean	1,169.0	1,609.8	24.9	1,169.0	1162.9	0.4	0.67
Age <1 year (%)	2.9	2.4	2.7	2.9	2.6	1.6	0.07
Age 1–18 years (%)	44.4	38.6	11.7	44.4	45.8	2.9	<0.001
Age 19–64 years (%)	46.4	46.0	0.9	46.4	45.2	2.4	0.01
Age ≥65 years (%)	6.4	13.0	22.6	6.4	6.4	0.2	0.83
Enrolled in Medicaid due to disability (%)	23.7	22.3	3.3	23.7	23.5	0.3	0.73
Medicare-Medicaid enrollee (%)	19.1	28.2	21.4	19.1	18.9	0.6	0.50
Non-white (%)	18.4	16.8	4.3	18.4	18.4	0.1	0.88
Missing race (%)	12.4	12.0	1.2	12.4	12.3	0.2	0.85
Continuously enrolled in Medicaid (%)	98.5	98.7	1.4	98.5	98.5	0.0	1.00
Months enrolled in a year, mean	11.3	11.3	0.1	11.3	11.3	0.1	0.94
Attributed to AC or comparison group because enrolled in a MaineCare HH (%)	44.4	29.1	32.1	44.4	45.3	1.8	0.05
Attributed to AC or comparison group because of the number of visits to an ED (%)	0.6	0.1	8.4	0.6	0.5	1.1	0.21
Receives full Medicaid benefits (%)	90.1	82.0	23.4	90.1	90.2	0.5	0.62
Enrolled in AC or comparison group for 2 years (%)	73.4	80.1	15.9	73.4	72.5	2.0	0.03
CDPS risk score in the prior year, mean	1.0	1.0	2.1	1.0	1.0	0.1	0.90

(continued)

Table B-2-6. Unweighted and weighted means and standardized differences, Medicaid AC and comparison groups, August 2011–July 2012 (continued)

Characteristic	Unweighted			Weighted			
	Medicaid AC group	Comparison group	Standardized difference[a]	Medicaid AC group	Comparison group	Standardized difference[a]	p-value
Total annual Medicaid payments in the previous year, $	3,235	2,905	3.4	3,235	3,249	0.1	0.88
Number of inpatient admissions in the previous year, mean	0.1	0.1	0.6	0.1	0.1	0.1	0.95
Number of ED visits in the previous year, mean	0.5	0.5	0.7	0.5	0.5	0.2	0.86
Metropolitan status of the beneficiary's county	53.9	54.4	0.9	53.9	53.7	0.4	0.64
Under 65 years and uninsured (%)	13.7	13.7	2.2	13.7	13.7	2.4	0.01
Median age, mean	41.7	42.9	54.9	41.7	42.8	51.7	<0.001
Poverty rate (%)	15.3	14.4	46.1	15.3	14.5	39.8	<0.001
Hospital beds per 1,000 population	3.7	2.9	62.7	3.7	2.9	63.5	<0.001
Physicians per 1,000 population	1.2	1.0	51.6	1.2	1.0	51.7	<0.001
Community mental health centers per 1,000 population	1.8	1.1	0.8	1.8	1.1	0.8	0.36

AC = Accountable Community; CDPS = Chronic Illness and Disability Payment System; ED = emergency department; HH = health home.

[a] Absolute standardized differences are expressed as percentages.

Table B-2-7. Unweighted and weighted means and standardized differences, Medicaid AC and comparison groups, August 2012–July 2013

Characteristic	Unweighted			Weighted			
	Medicaid AC group	Comparison group	Standardized difference[a]	Medicaid AC group	Comparison group	Standardized difference[a]	p-value
N	43,994	166,587		43,994	44,476		
Female (%)	57.3	58.2	1.9	57.3	56.8	1.3	0.06
Age, mean	27.3	31.7	19.1	27.3	26.8	2.3	<0.001
Age squared, mean	1,212.6	1,615.1	22.4	1,212.6	1,177.2	2.2	<0.001
Age <1 year (%)	2.7	2.3	2.7	2.7	2.0	4.7	<0.001
Age 1–18 years (%)	43.1	39.6	7.1	43.1	45.2	4.2	<0.001
Age 19–64 years (%)	47.7	44.8	5.8	47.7	46.2	3.0	<0.001
Age ≥65 years (%)	6.5	13.3	23.0	6.5	6.6	0.5	0.45
Enrolled in Medicaid due to disability (%)	23.2	22.0	3.0	23.2	22.5	1.7	0.01
Medicare-Medicaid enrollee (%)	18.7	27.9	21.8	18.7	18.0	1.8	0.01
Non-white (%)	18.8	18.0	2.2	18.8	19.1	0.6	0.35
Missing race (%)	12.9	13.1	0.4	12.9	13.1	0.5	0.51
Continuously enrolled in Medicaid (%)	98.5	98.4	0.5	98.5	98.4	0.4	0.56
Months enrolled in a year, mean	11.2	11.2	0.4	11.2	11.2	0.9	0.17
Attributed to AC or comparison group because enrolled in a MaineCare HH (%)	42.9	26.9	34.1	42.9	43.9	2.0	<0.001
Attributed to AC or comparison group because of the number of visits to an ED (%)	0.6	0.1	8.8	0.6	0.6	0.9	0.18
Receives full Medicaid benefits (%)	91.3	82.7	25.9	91.3	91.5	0.6	0.36
Enrolled in AC or comparison group for 2 years (%)	42.4	71.5	61.5	42.4	42.2	0.3	0.67
CDPS risk score in the prior year, mean	1.3	1.3	1.5	1.3	1.3	0.5	0.44

(continued)

Table B-2-7. Unweighted and weighted means and standardized differences, Medicaid AC and comparison groups, August 2012–July 2013 (continued)

Characteristic	Unweighted			Weighted			
	Medicaid AC group	Comparison group	Standardized difference[a]	Medicaid AC group	Comparison group	Standardized difference[a]	p-value
Total annual Medicaid payments in the previous year, $	5,975	55,188	3.1	5,975	6,020	0.3	0.67
Number of inpatient admissions in the previous year, mean	0.1	0.1	0.2	0.1	0.1	0.2	0.82
Number of ED visits in the previous year, mean	1.0	0.9	3.8	1.0	1.0	0.3	0.64
Metropolitan status of the beneficiary's county	50.6	54.2	7.1	50.6	52.2	3.1	<0.001
Under 65 years and uninsured (%)	14.0	13.7	15.2	14.0	13.8	12.4	<0.001
Median age, mean	42.2	42.9	29.2	42.2	42.9	28.1	<0.001
Poverty rate (%)	15.4	14.4	47.7	15.4	14.6	39.6	<0.001
Hospital beds per 1,000 population	3.7	2.9	58.7	3.7	2.9	58.5	<0.001
Physicians per 1,000 population	1.2	1.0	37.4	1.2	1.0	37.1	<0.001
Community mental health centers per 1,000 population	3.1	2.2	0.9	3.1	4.5	1.1	0.09

AC = Accountable Community; CDPS = Chronic Illness and Disability Payment System; ED = emergency department; HH = health home.

[a] Absolute standardized differences are expressed as percentages.

Table B-2-8. Unweighted and weighted means and standardized differences, Medicaid AC and comparison groups, August 2013–July 2014

Characteristic	Unweighted			Weighted			
	Medicaid AC group	Comparison group	Standardized difference[a]	Medicaid AC group	Comparison group	Standardized difference[a]	p-value
N	47,093	177,707		47,093	47,584		
Female (%)	57.0	57.9	1.8	57.0	56.4	1.2	0.07
Age, mean	27.7	32.1	19.1	27.7	27.2	2.2	<0.001
Age squared, mean	1,240.1	1,646.5	22.2	1,240.1	1,205.9	2.1	<0.001
Age <1 year (%)	2.3	1.8	3.4	2.3	1.4	6.5	<0.001
Age 1–18 years (%)	43.2	39.7	7.0	43.2	45.3	4.3	<0.001
Age 19–64 years (%)	47.7	44.8	5.8	47.7	46.4	2.6	<0.001
Age ≥65 years (%)	6.8	13.6	22.7	6.8	6.9	0.4	0.59
Enrolled in Medicaid due to disability (%)	24.0	22.4	3.6	24.0	23.3	1.7	0.01
Medicare-Medicaid enrollee (%)	18.8	27.9	21.6	18.8	18.1	1.8	0.01
Non-white (%)	20.1	19.0	2.9	20.1	20.4	0.7	0.26
Missing race (%)	14.1	14.0	0.3	14.1	14.3	0.6	0.39
Continuously enrolled in Medicaid (%)	97.9	97.8	0.5	97.9	97.8	0.8	0.20
Months enrolled in a year, mean	11.2	11.2	2.3	11.2	11.2	1.2	0.07
Attributed to AC or comparison group because enrolled in a MaineCare HH (%)	41.8	26.2	33.4	41.8	42.8	1.8	<0.001
Attributed to AC or comparison group because of the number of visits to an ED (%)	0.6	0.1	8.7	0.6	0.5	1.2	0.07
Receives full Medicaid benefits (%)	92.0	93.0	27.4	92.0	92.1	0.4	0.56
Enrolled in AC or comparison group for 2 years (%)	42.3	71.4	61.5	42.3	42.3	0.1	0.91
CDPS risk score in the prior year, mean	1.3	1.3	1.8	1.3	1.3	0.5	0.43

(continued)

Table B-2-8. Unweighted and weighted means and standardized differences, Medicaid AC and comparison groups, August 2013–July 2014 (continued)

Characteristic	Unweighted			Weighted			
	Medicaid AC group	Comparison group	Standardized difference[a]	Medicaid AC group	Comparison group	Standardized difference[a]	p-Value
Total annual Medicaid payments in the previous year, $	6,549	6,025	3.5	6,549	6,634	0.5	0.41
Number of inpatient admissions in the previous year, mean	0.1	0.1	0.8	0.1	0.1	0.3	0.69
Number of ED visits in the previous year, mean	1.0	0.9	4.0	1.0	1.0	0.1	0.89
Metropolitan status of the beneficiary's county	51.0	54.4	6.8	51.0	52.4	2.9	<0.001
Under 65 years and uninsured (%)	14.0	13.7	15.2	14.0	13.7	12.6	<0.001
Median age, mean	42.2	42.9	29.6	42.2	42.9	28.5	<0.001
Poverty rate (%)	15.4	14.3	47.8	15.4	14.5	39.9	<0.001
Hospital beds per 1,000 population	3.7	2.9	59.8	3.7	2.9	59.6	<0.001
Physicians per 1,000 population	1.2	1.0	37.7	1.2	1.0	37.3	<0.001
Community mental health centers per 1,000 population	3.6	2.1	1.3	3.6	4.0	0.4	0.60

AC = Accountable Community; CDPS = Chronic Illness and Disability Payment System; ED = emergency department; HH = health home; SSP = Shared Savings Program.

[a] Absolute standardized differences are expressed as percentages.

Table B-2-9. Unweighted and weighted means and standardized differences, Medicaid AC and comparison groups, August 2014–July 2015

Characteristic	Unweighted			Weighted			
	Medicaid AC group	Comparison group	Standardized difference[a]	Medicaid AC group	Comparison group	Standardized difference[a]	p-value
N	50,307	188,562		50,307	50,811		
Female (%)	56.6	57.8	2.5	56.6	55.9	1.3	0.04
Age, mean	28.1	32.7	19.5	28.1	27.7	2.0	<0.001
Age squared, mean	1,268.0	1,686.0	22.4	1,268.0	1,237.2	1.8	<0.001
Age <1 year (%)	1.5	1.3	1.9	1.5	0.8	6.8	<0.001
Age 1–18 years (%)	43.2	39.3	7.9	43.2	44.9	3.4	<0.001
Age 19–64 years (%)	48.3	45.3	5.9	48.3	47.1	2.4	<0.001
Age ≥65 years (%)	7.1	14.1	23.1	7.1	7.3	0.9	0.16
Enrolled in Medicaid due to disability (%)	23.9	22.4	3.7	23.9	23.2	1.7	0.01
Medicare-Medicaid enrollee (%)	18.8	28.1	22.0	18.8	18.4	1.1	0.08
Non-white (%)	21.2	19.5	4.2	21.2	21.4	0.7	0.30
Missing race (%)	14.9	14.5	1.3	14.9	15.1	0.5	0.46
Continuously enrolled in Medicaid (%)	97.6	97.7	0.5	97.6	97.7	0.5	0.48
Months enrolled in a year, mean	11.2	11.2	1.5	11.2	11.2	1.5	0.02
Attributed to AC or comparison group because enrolled in a MaineCare HH (%)	40.1	25.4	31.7	40.1	40.7	1.3	0.04
Attributed to AC or comparison group because of the number of visits to an ED (%)	0.6	0.1	8.8	0.6	0.6	1.0	0.10
Receives full Medicaid benefits (%)	93.3	83.5	31.0	93.3	93.1	0.6	0.34
Enrolled in AC or comparison group for 2 years (%)	41.8	71.6	63.1	41.8	41.9	0.2	0.74
CDPS risk score in the prior year, mean	1.3	1.3	1.5	1.3	1.3	0.7	0.25

(continued)

Table B-2-9. Unweighted and weighted means and standardized differences, Medicaid AC and comparison groups, August 2014–July 2015 (continued)

Characteristic	Unweighted			Weighted			
	Medicaid AC group	Comparison group	Standardized difference[a]	Medicaid AC group	Comparison group	Standardized difference[a]	p-value
Total annual Medicaid payments in the previous year, $	6,326	5,716	4.1	6,326	6,385	0.4	0.55
Number of inpatient admissions in the previous year, mean	0.1	0.1	1.8	0.1	0.1	0.0	0.95
Number of ED visits in the previous year, mean	0.9	0.8	3.6	0.9	0.9	0.0	0.96
Metropolitan status of the beneficiary's county	51.3	54.6	6.6	51.3	52.6	2.5	<0.001
Under 65 years and uninsured (%)	13.9	13.7	15.5	13.9	13.7	13.0	<0.001
Median age, mean	42.2	42.9	29.9	42.2	42.9	29.3	<0.001
Poverty rate (%)	15.4	14.3	47.9	15.4	14.5	40.5	<0.001
Hospital beds per 1,000 population	3.7	2.9	60.5	3.7	2.9	60.4	<0.001
Physicians per 1,000 population	1.2	1.0	38.3	1.2	1.0	37.6	<0.001
Community mental health centers per 1,000 population	2.2	0.8	1.9	2.2	1.8	0.4	0.48

AC = Accountable Community; CDPS = Chronic Illness and Disability Payment System; ED = emergency department; HH = health home.

[a] Absolute standardized differences are expressed as percentages.

Table B-2-10. Unweighted and weighted means and standardized differences, Medicaid AC and comparison groups, August 2015–July 2016

Characteristic	Unweighted			Weighted			
	Medicaid AC group	Comparison group	Standardized difference[a]	Medicaid AC group	Comparison group	Standardized difference[a]	p-value
N	44,852	163,955		44,852	44,622		
Female (%)	56.8	58.1	2.6	56.8	55.9	1.9	0.01
Age, mean	28.6	33.2	19.7	28.6	27.6	4.6	<0.001
Age squared, mean	1,290.4	1,726.2	23.3	1,290.4	1,227.7	3.8	<0.001
Age <1 year (%)	0.8	0.8	0.2	0.8	2.0	10.2	<0.001
Age 1–18 years (%)	43.1	39.2	8.0	43.1	43.9	1.6	0.02
Age 19–64 years (%)	49.1	45.6	6.9	49.1	47.1	3.9	<0.001
Age ≥65 years (%)	7.0	14.4	24.1	7.0	7.0	0.2	0.81
Enrolled in Medicaid due to disability (%)	23.8	22.9	2.0	23.8	22.7	2.5	<0.001
Medicare-Medicaid enrollee (%)	18.5	29.0	24.8	18.5	17.6	2.5	<0.001
Non-white (%)	20.2	18.5	4.4	20.2	21.0	2.1	<0.001
Missing race (%)	13.7	13.3	1.1	13.7	14.3	1.9	<0.001
Continuously enrolled in Medicaid (%)	97.9	97.6	2.0	97.9	97.9	0.2	0.81
Months enrolled in a year, mean	11.2	11.2	3.1	11.2	11.2	3.9	<0.001
Attributed to AC or comparison group because enrolled in a MaineCare HH (%)	44.0	26.4	37.7	44.0	43.2	1.7	0.01
Attributed to AC or comparison group because of the number of visits to an ED (%)	0.5	0.1	6.9	0.5	0.4	0.8	0.24
Receives full Medicaid benefits (%)	93.9	83.7	32.7	93.9	93.8	0.1	0.84
Enrolled in AC or comparison group for 2 years (%)	47.5	82.3	78.4	47.5	48.9	2.8	<0.001
CDPS risk score in the prior year, mean	1.3	1.3	0.1	1.3	1.2	2.7	<0.001

(continued)

Table B-2-10. Unweighted and weighted means and standardized differences, Medicaid AC and comparison groups, August 2015–July 2016 (continued)

Characteristic	Unweighted			Weighted			
	Medicaid AC group	Comparison group	Standardized difference[a]	Medicaid AC group	Comparison group	Standardized difference[a]	p-value
Total annual Medicaid payments in the previous year, $	5,729	5,737	0.1	5,729	5,425	2.1	<0.001
Number of inpatient admissions in the previous year, mean	0.1	0.1	0.4	0.1	0.1	1.4	0.04
Number of ED visits in the previous year, mean	0.9	0.8	1.6	0.9	0.8	3.2	<0.001
Metropolitan status of the beneficiary's county	50.9	55.4	9.1	50.9	53.6	5.4	<0.001
Under 65 years and uninsured (%)	14.0	13.6	21.4	14.0	13.6	22.3	<0.001
Median age, mean	42.3	42.9	24.3	42.3	42.8	21.5	<0.001
Poverty rate (%)	15.4	14.3	50.7	15.4	14.5	43.2	<0.001
Hospital beds per 1,000 population	3.7	2.9	60.4	3.7	2.9	61.9	<0.001
Physicians per 1,000 population	1.1	1.0	35.4	1.1	1.0	36.5	<0.001
Community mental health centers per 1,000 population	3.7	0.1	3.9	3.7	0.0	4.0	<0.001

AC = Accountable Community; CDPS = Chronic Illness and Disability Payment System; ED = emergency department; HH = health home.

[a] Absolute standardized differences are expressed as percentages.

B-2.2.6 Statistical analysis

Regression model

The underlying assumption in D-in-D models estimating the impact of the ACs is that trends in the AC group would be similar to those in the comparison group in the absence of the initiative (i.e., that the two were on "parallel paths" prior to the start of the ACs).

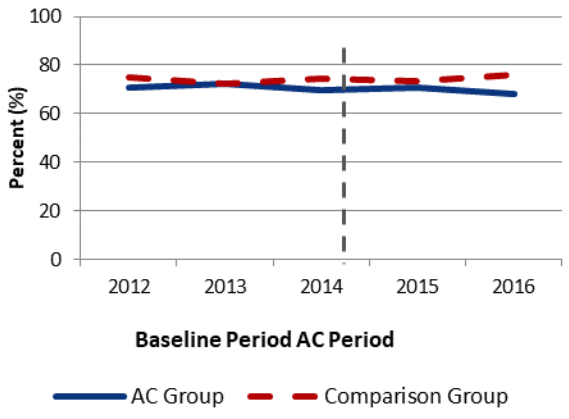
To assess the parallel assumption's validity more empirically, we modeled core expenditure and utilization outcomes during the baseline period with a linear time trend interacted with a dichotomous variable indicating the beneficiary was attributed to an AC provider (i.e., the "test" group). The following section describes the baseline analysis conducted to inform the D-in-D model.

To examine descriptively whether the trends in the AC and comparison group are parallel, we present graphs of annual unadjusted averages for the AC group and the comparison group for the baseline period (August 2011–July 2014) and the first 2 years of implementation (August 2014–July 2016).

Figures B-2-6 to ***B-2-9*** provide the unadjusted averages of the care coordination measures by year.

- The baseline trends were relatively parallel for **7-day and 30-day follow-ups following a mental health hospitalization and specialist visits,** and they were less parallel for **primary care visits**.

Figure B-2-6. Percentage of Medicaid beneficiaries with a follow-up visit within 7 days of discharge from hospitalization for mental illness, August 2011 through July 2016, Maine AC group and comparison group

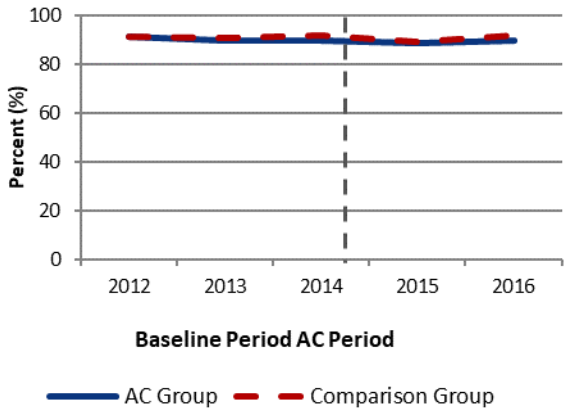


The percentage of individuals with **7-day follow-up following a mental health hospitalization** among AC and comparison group beneficiaries remained stable during the baseline period. In the AC period, the percent increased slightly over time for the comparison group beneficiaries while the percent decreased over time for AC beneficiaries (***Figure B-2-6***).

AC = Accountable Community.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figure B-2-7. Percentage of Medicaid beneficiaries with a follow-up visit within 30 days of discharge from hospitalization for mental illness, August 2011 through July 2016, Maine AC group and comparison group

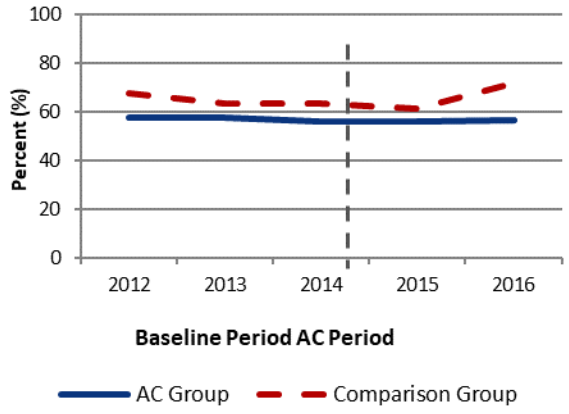


The percentage of individuals with **30-day follow-up following a mental health hospitalization** among AC and comparison group beneficiaries remained stable during the baseline and AC period (***Figure B-2-7***).

AC = Accountable Community.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figure B-2-8. Percentage of beneficiaries with a visit to a primary care provider, August 2011 through July 2016, Maine AC group and comparison group

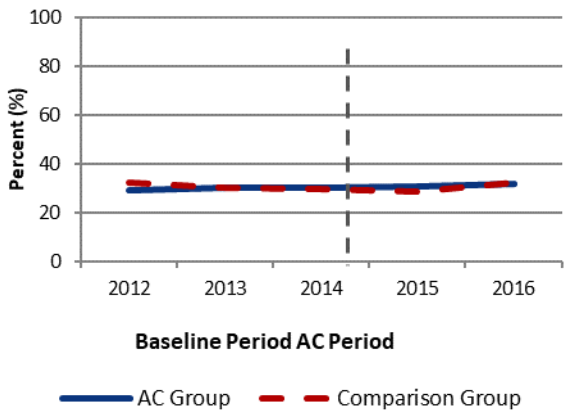


The percentage of beneficiaries with a **visit to a primary care provider** among AC beneficiaries remained stable over time, while the percentage increased only for the comparison group during the AC period (***Figure B-2-8***).

AC = Accountable Community.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figure B-2-9. Percentage of beneficiaries with a visit to a specialty provider, August 2011 through July 2016, Maine AC group and comparison group



The percentage of beneficiaries with a **visit to a specialty provider** among AC and comparison group beneficiaries remained stable over time (***Figure B-2-9***).

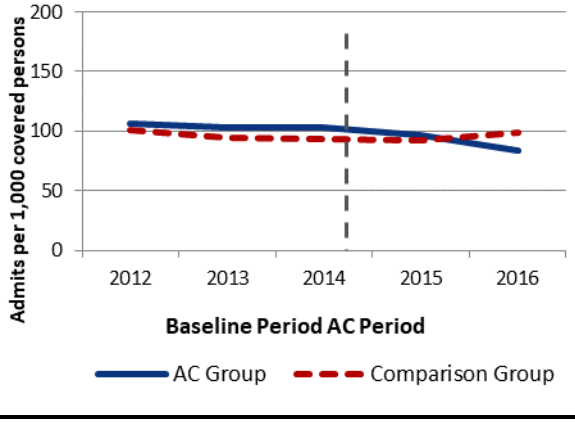
AC = Accountable Community.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figures B-2-10 to ***B-2-12*** provide unadjusted annual averages of inpatient admissions, outpatient ED visits, and 30-day readmissions per 1,000 Medicaid beneficiaries by year, respectively.

- The baseline trends were parallel for **acute inpatient admission** and **outpatient ED visit rates** but not parallel for the rate of **30-day readmissions**.

Figure B-2-10. All-cause acute inpatient admissions per 1,000 Medicaid beneficiaries, August 2011 through July 2016, Maine AC group and comparison group

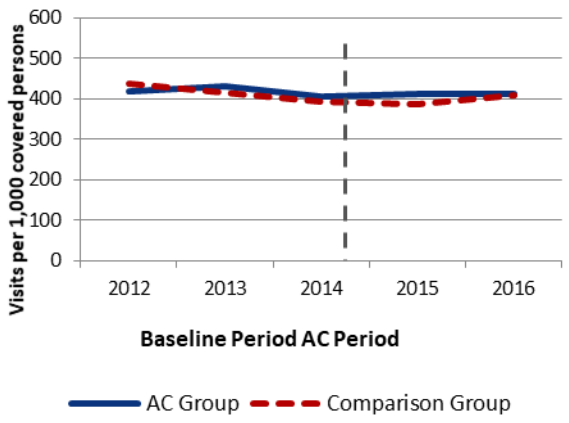


The rate of **acute inpatient admissions** among AC beneficiaries and comparison group beneficiaries was similar throughout the baseline period. During the AC period, the rate began trending downward for AC beneficiaries (***Figure B-2-10***).

AC = Accountable Community.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figure B-2-11. ED visits that did not lead to a hospitalization per 1,000 Medicaid beneficiaries, August 2011 through July 2016, Maine AC group and comparison group

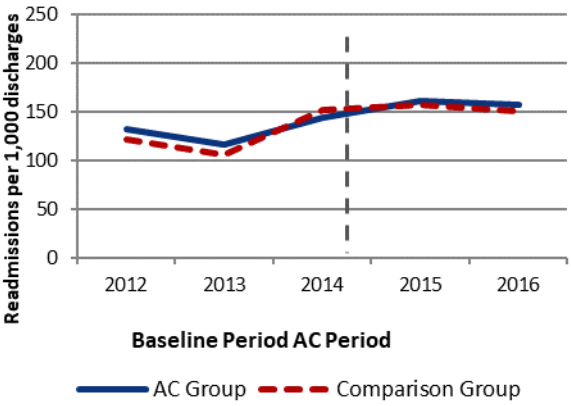


The rate of **outpatient ED visits** was similar throughout the baseline and AC periods for AC and comparison beneficiaries (***Figure B-2-11***).

AC = Accountable Community; ED = emergency department.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figure B-2-12. Discharges with a readmission within 30 days per 1,000 Medicaid beneficiaries, August 2011 through July 2016, Maine AC group and comparison group



The rate of discharges with a **readmission within 30 days** among AC beneficiaries and comparison group beneficiaries fluctuated in the baseline period and was steady during the AC period (***Figure B-2-12***). Both groups experienced an increase in the readmission rate over time.

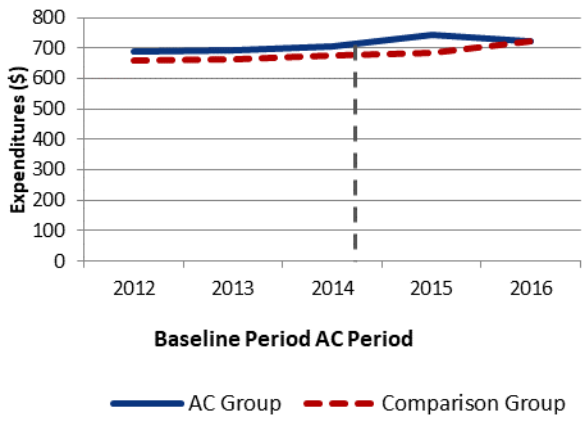
AC = Accountable Community.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figures B-2-13 to ***B-2-16*** provide unadjusted annual averages of total, inpatient, professional, and pharmaceutical Medicaid per beneficiary per month (PBPM) expenditures.

- The baseline trends were parallel for **total, inpatient, professional, and pharmaceutical Medicaid expenditures**.

Figure B-2-13. Average total PBPM expenditures, August 2011 through July 2016, Maine AC group and comparison group

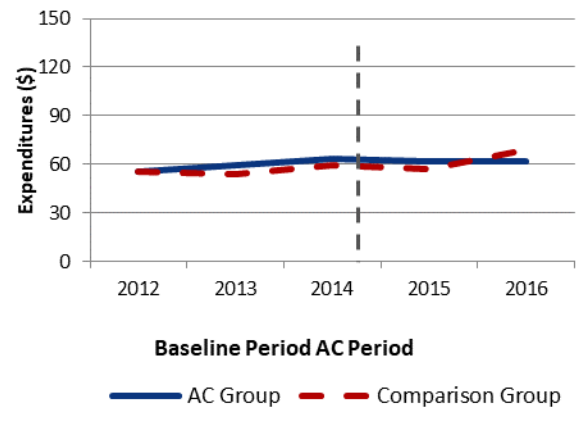


The average **total PBPM** expenditures trend was similar for AC and comparison group beneficiaries in the baseline period, with a slight trend toward higher payments in the AC period for both groups (***Figure B-2-13***).

AC = Accountable Community; PBPM = per beneficiary per month.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figure B-2-14. Average inpatient PBPM expenditures, August 2011 through July 2016, Maine AC group and comparison group

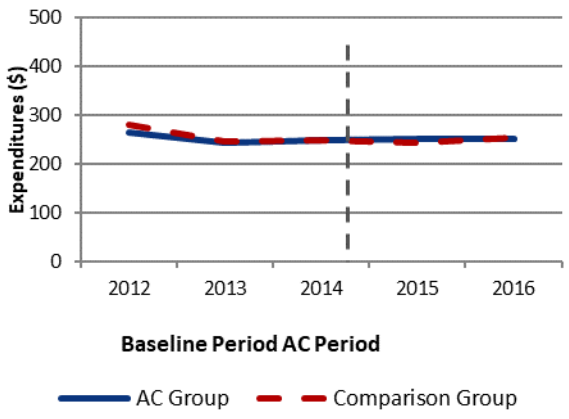


The average **inpatient PBPM** expenditures trend was similar for AC and comparison group beneficiaries in the baseline period, with a slight trend toward higher payments in the AC period for the comparison group (**Figure B-2-14**).

AC = Accountable Community; PBPM = per beneficiary per month.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figure B-2-15. Average professional PBPM expenditures, August 2011 through July 2016, Maine AC group and comparison group

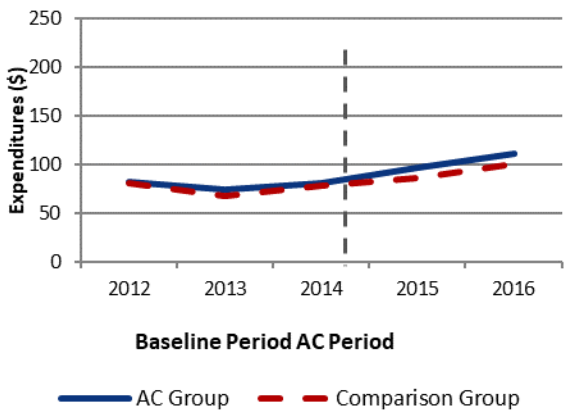


The average **professional PBPM** expenditures trend was similar for AC and comparison group beneficiaries in the baseline and AC periods. Professional payments began trending downward in the beginning of the baseline period before leveling out (***Figure B-2-15***).

AC = Accountable Community; PBPM = per beneficiary per month.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figure B-2-16. Average pharmaceutical PBPM expenditures, August 2011 through July 2016, Maine AC group and comparison group



The average **pharmaceutical PBPM** expenditures was similar for AC and comparison group beneficiaries in the baseline and AC periods, with a trend toward higher payments starting midway through the baseline period and extending throughout the AC period (***Figure B-2-16***).

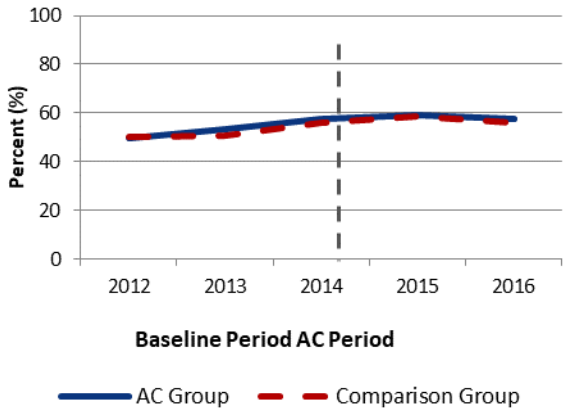
AC = Accountable Community; PBPM = per beneficiary per month.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figures B-2-17 to ***B-2-19*** provide the unadjusted averages of the quality of care measures by year.

- The baseline trends were parallel for **84- and 180-day antidepressant medication management and receipt of an HbA1c test**.

Figure B-2-17. Percentage of Medicaid beneficiaries aged 18 years or older with depression who remained on antidepressant medication for at least 84 days, August 2011 through July 2016, Maine AC group and comparison group

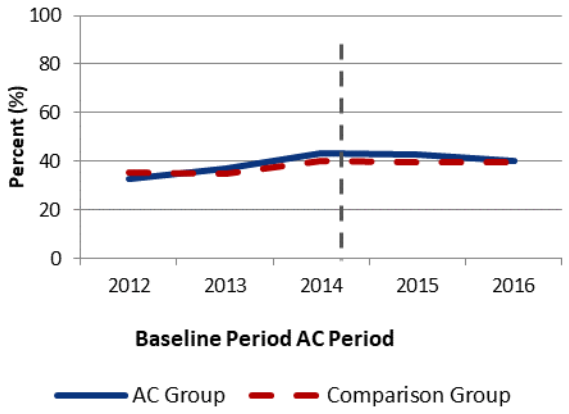


The percentage of patients who remained on antidepressant medication for at least 84 days among AC and comparison group beneficiaries was similar throughout the baseline and AC periods (***Figure B-2-17***). The percentage for both outcomes trended slightly upward during the baseline period and remained steady during the AC period.

AC = Accountable Community.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figure B-2-18. Percentage of Medicaid beneficiaries aged 18 years or older with depression who remained on antidepressant medication for at least 180 days, August 2011 through July 2016, Maine AC group and comparison group

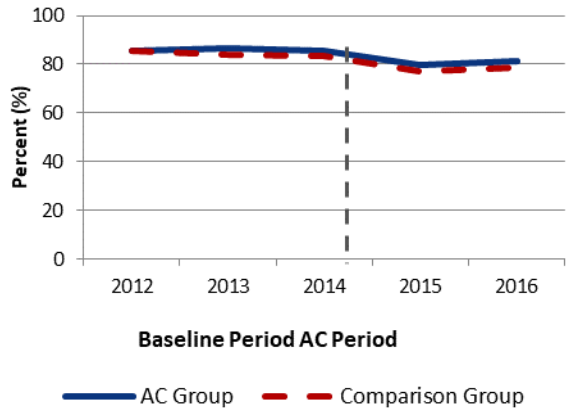


The percentage of **patients who remained on antidepressant medication for at least 180 days** among AC and comparison group beneficiaries was similar throughout the baseline and AC periods (***Figure B-2-18***). For both groups, the percentage trended slightly upward during the baseline period and began to decrease slightly during the AC period.

AC = Accountable Community.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

Figure B-2-19. Percentage of Medicaid beneficiaries aged 18–75 years with diabetes who received an HbA1c test, August 2011 through July 2016, Maine AC group and comparison group



The rate of **patients with diabetes who received an HbA1c test** among AC and comparison group beneficiaries trended slightly downward during the baseline period through the beginning of the AC period. The rate began increasing slightly midway through the AC period (***Figure B-2-19***).

AC = Accountable Community; HbA1c = hemoglobin A1c.

Note: Years are defined as follows, 2012=8/2011–7/2012; 2013=8/2012–7/2013; 2014=8/2013–7/2014; 2015=8/2014–7/2015; 2016=8/2015–7/2016.

An annual fixed-effects model considered for the evaluation is shown in ***Equation B-2.2***:

$$\gamma = \alpha_0 + \alpha_1 I + \sum \beta_n Q_{n,b} + \sum \phi_t Q_{t,p} \bullet I + \delta X + \mu \qquad \text{(B-2.2)}$$

where

y = a performance measure (e.g., total PBPM cost per year) for the i-th beneficiary in the j-th group (AC or comparison), in period t (i,j,t subscripts suppressed).

I = a 0,1 indicator (0 = comparison group, 1 = AC group).

X = a vector of patient and demographic characteristics.

$Q_{n,b}$, $Q_{t,d}$ = 0,1 indicator of the n-th or t-th calendar year in the base (b) or post (p) period (n starts counting at first baseline period, whereas t starts with first AC Model year).

μ = error term.

The model in ***Equation B-2.2*** assumes that, except for an intercept difference α_1, the outcomes for beneficiaries in the AC group and comparison group followed a similar growth trend during the baseline period. We investigated whether the baseline period before the start of AC model satisfied the baseline trend assumptions of the D-in-D model in ***Equation B-2.2***—that is, whether the outcome trends for beneficiaries in the AC and comparison groups were similar during this period.

To test the similarity of baseline trends, a model with a linear trend during the baseline period was used. We tested whether this trend differed for AC group participants relative to comparison group participants. Specifically, the model for the outcomes may be written as follows:

$$y = \alpha_0 + \alpha_1 I + \theta \bullet t + \lambda I \bullet t + \delta X + \mu. \quad \text{(B-2.3)}$$

In ***Equation B-2.3***, *y, I,* X, and μ are defined as in ***Equation B-2.2***. The variable *t* is linear time ranging from 1 to 3. The linear time trend in the comparison group is θ•*t*, whereas for AC group beneficiaries ($I = 1$), it is $(\theta + \lambda) * t$. Hence, λ measures the difference in linear trends, and the *t*-statistic for this coefficient can be used to test the null hypothesis of equal trends (λ = 0). In other words, rejecting the null hypothesis would suggest that the assumption of equal trends underlying the outcome models is not met.

The parameters of ***Equation B-2.3*** were estimated using weighted least-squares regression models for three key outcomes. The weights are a function of the eligibility fraction and propensity scores. For each outcome, estimates and standard errors of the difference between the baseline trend in the AC and comparison groups (λ) are reported.

Table B-2-11 shows estimates of the baseline trend differences for the following outcomes:

- Total Medicaid PBPM expenditures
- Probability of an acute inpatient stay
- Probability of an outpatient ED visit.

Table B-2-11. Differences in average expenditure and utilization outcomes during the baseline period, AC group and comparison group beneficiaries

Parameter estimate	Total PBPM ($)	Any inpatient	Any outpatient ED visit
AC–comparison group trend difference	−2.102	.001	.011
	(11.823)	(.001)	(.003)

AC = Accountable Community; ED = emergency department; PBPM = per beneficiary per month.

Baseline is the period August 2011–July 2014. The trend (slope) is the year-to-year change in the outcome variable. Standard errors are given in parentheses. *p < 0.10; **p < 0.05; ***p < 0.01.

Relative to the comparison group, there was no statistically significant difference in the baseline trends for total Medicaid PBPM expenditures and the likelihood of having an inpatient admissions or ED visit for the AC group. Based on these results, we concluded that in general, beneficiaries in the AC group were on a trajectory similar to that of comparison beneficiaries prior to the AC model, and thus, the parallel trend assumption of the D-in-D model was satisfied.

D-in-D regression model. The D-in-D model is shown in ***Equation B-2.4***. This model is an annual fixed effects model, as shown in ***Equation B-2.2.*** As in ***Equation B-2.2***, Y_{ijt} is the outcome for individual *i* (AC group or comparison group) in state *j* in year *t*; I_{ij} (=0,1) is an indicator equal to 1 if the individual is in the AC group and 0 if the individual is in its comparison group; *Qn* is a series of yearly dummies for the baseline period (Years 1 to 3); and Q_t is a series of yearly dummies for the post years (Years x to x). The interaction of the AC group indicator and Q_t (I_{ij}* Q_t) measures the difference in the pre-post change between the test group and its comparison states.

$$Y_{ijt} = \alpha_0 + \beta_1 I_{ij} + \sum \beta_2 Q_n + \sum \alpha_2 Q_t + \sum \gamma I_{ijt} * Q_t + \lambda X_{ijt} + \varepsilon_{ijt} \quad \text{(B-2.4)}$$

Table B-2-12 illustrates the interpretation of the D-in-D estimate from this model. The coefficient β_1 in ***Equation B-2.4*** is the difference in the measure between AC beneficiaries and comparison beneficiaries at the start of the baseline period, holding constant other variables in the equation. The β_2 and α_2 coefficients are for the annual fixed effects and capture differences over time for each baseline and post year, respectively. The coefficient of the interaction term between Q_t and *I* measures any differences for the AC group relative to the comparison group in the post years relative to baseline years. Thus, in the post period, the comparison group mean is captured by $\alpha_0 + \alpha_2$, whereas the AC group mean is captured by $(\alpha_0 + \beta_1) + (\alpha_2 + \gamma)$. In other words, the between-group difference changes from β_1 during the baseline years to $\beta_1 + \gamma$ during the post period. The D-in-D parameter, γ, shows whether the between-group difference increased ($\gamma>0$) or decreased ($\gamma<0$) after the AC was implemented. Using the annual fixed effects model, we calculated overall estimates by taking linear combinations of the yearly estimates.

Table B-2-12. D-in-D estimate

Group	Pre period	Post period	Pre-post difference
AC group	$\alpha_0 + \beta_1 + \beta_2$	$(\alpha_0 + \beta_1) + (\alpha_2 + \gamma)$	$\alpha_2 + \gamma$
Comparison group	$\alpha_0 + \beta_2$	$\alpha_0 + \alpha_2$	α_2
Between group	β_1	$\beta_1 + \gamma$	γ

AC = Accountable Community; D-in-D = difference-in-differences.

Models for unplanned readmissions and mental health follow-ups were estimated at the annual-admission level. All other outcomes were estimated with the beneficiary year as the unit of analysis.

The outcome model for total Medicaid PBPM expenditures was estimated using OLS. To show the adjusted means in the pre- and post-periods for the AC and comparison groups, a linear model that allows for the calculation of means that will sum to the D-in-D estimate was used. Although this model has strong assumptions of normality of the outcome, the OLS model still produces unbiased estimates, even when the normality assumptions is violated, as long as errors

are uncorrelated and have a constant variance (Gauss-Markov Theorem). However, we could and did control for the correlation and variance in errors with clustered standard errors at the provider organization level. Additionally, the model yields estimates that are readily interpretable in dollars and do not require additional transformation.

For all other outcomes, we converted utilization counts into binary outcomes (1 = any use during the year) and used weighted logistic regression models. Count models are not appropriate because of the low occurrence of most types of utilization for individual beneficiaries in any year; however, the marginal effects from the inpatient admission and ED visit logistic regression models were multiplied by 1,000 to obtain approximate rates of inpatient and ED utilization per 1,000 beneficiaries. Multiplying the marginal effect by 1,000 does not produce an exact rate of inpatient or ED utilization per 1,000 beneficiaries because it assumes no person has more than one visit or admission per year. However, we concluded that this is a reasonable approximation because only a small percentage of beneficiaries had counts exceeding one for the inpatient admission and ED visit measures. In addition, the marginal effects from the primary care visit and specialty visit logistic regression models were multiplied by 100 to obtain a percentage probability of any visits during the year. For expenditure outcomes, weighted generalized linear models with a normal distribution and identity link were used.

The models for total expenditures, inpatient admissions, ED visits, and readmissions were run separately for children and adults. In addition, we ran these outcomes and behavioral health related expenditures and admissions separately for people with behavioral health conditions.

Control variables. In all models, we controlled for the following variables:

- Age and age-squared
- Gender
- Race (non-white and missing race)
- Enrollment due to disability
- Medicare/Medicaid enrollee
- Beneficiary's classification on the CDPS[15]
- Number of months the beneficiary was enrolled in Medicaid during the year

[15] The chronic illness and disability payment system (CDPS) is a diagnostic classification system originally developed for states to use in adjusting capitated payments for Temporary Assistance for Needy Families and disabled Medicaid beneficiaries and used to predict Medicaid costs. The RTI team used the CDPS to measure beneficiary morbidity. The CDPS maps selected diagnoses and prescriptions to numeric weights. Beneficiaries with a CDPS score of 0 have no diagnoses or prescriptions that factor into creating the CDPS score. The more diagnoses a beneficiary has or the greater the severity of a particular diagnosis, the larger the CDPS score will be.

- If the beneficiary stayed continuously enrolled in Medicaid during the year[16]
- If the beneficiary had full Medicaid benefits during the year
- If the beneficiary was enrolled in the AC at some point in both test years
- Method of attribution to the AC or the comparison group, i.e., whether the beneficiary was enrolled in the AC or the comparison group because he/she had a majority of visits to a primary care provider, because he/she was enrolled in an HH, or because he/she had a majority of visits to an ED
- County-level federal poverty level, median age, and uninsured rate
- Metropolitan status of the beneficiary's county
- County-level hospital beds and physicians per capita for all regression models and community mental health centers per capita for behavioral health-related outcomes regression models.

Weighting and clustering. All the regression models were estimated using weighted regressions and weighted by the propensity score times the eligibility fraction. In addition, standard errors were clustered at the provider organization level (e.g., HH, primary care provider, or ED) to account for clustering of beneficiaries within the provider through which they were attributed to the AC or comparison group.

Adjusted means. The regression-adjusted D-in-D estimate and the D-in-D calculated from regression-adjusted means will differ for one of two reasons. First, in nonlinear specifications the D-in-D calculated from the regression-adjusted means is known to be a biased estimator for the treatment effect. To address this bias, we use the nonlinear D-in-D approach described in Puhani (2012). In some cases the bias may be extreme, leading to substantial differences between the regression-adjusted D-in-D estimates versus the D-in-D calculated from regression-adjusted means.

Second, in linear specifications the D-in-D calculated from the overall regression-adjusted means may be substantially different than the overall regression-adjusted D-in-D estimate because we use different weights to obtain the overall figures. Specifically, the regression-adjusted D-in-D estimates are weighted using the number of treatment beneficiaries observed in each year relative to the total number of treatment beneficiaries ever observed during the test period. This is mathematically equivalent to weighting the test-period adjusted means for both groups with the same weights that are applied to the treatment group. However, the test-period adjusted means that are presented for the comparison group are weighted using the

[16] The RTI team controlled for whether or not the beneficiary was continuously enrolled, with no more than a one month break in enrollment, from the time the beneficiary first entered the Medicaid data in the year until the end of the measurement year. However, a person can have multiple occurrences of a one month break in enrollment and still be considered continuously enrolled This covariate was used to control for churning in and out of Medicaid.

number of comparison beneficiaries observed in each year relative to the total number of comparison beneficiaries ever observed during the test period. The implication of this is that in cases where there are large differences in the rates of rolling entry or exit across the two groups, we may observe large differences in the D-in-D calculated from the overall regression-adjusted means versus the overall regression-adjusted D-in-D estimate.

B-2.3 Methods for Qualitative Data Collection and Analysis

The Maine SIM Initiative Round 1 Evaluation team collected and analyzed a wide range of qualitative data in the fifth year of the federal SIM Initiative evaluation. These data sources included interviews with key informants and focus groups conducted during in-person site visits in previous evaluation years, a review of relevant documents, and regular evaluation calls with state officials leading the state's SIM Initiative. This report draws from past evaluation reports, where further detail is provided on previously conducted site visit interviews and focus groups.

B-2.3.1 Document review

We used Maine's quarterly and annual reports, operational plans, SIM partner presentations, and other state documents to obtain updated information on the state's implementation progress during the SIM Initiative test period. To supplement these documents, we collected relevant articles, briefs, and reviews of Maine's SIM Initiative activities and related initiatives. We also used sites that the state maintains on the initiative, such as the Maine SIM Steering Committee site, which includes end-of-project summaries for each of the Maine SIM partners.

In addition, we obtained numbers of providers and practices participating in and populations served by the different innovation models from reports Maine submits to the Innovation Center, quarterly reports submitted to CMS, state department websites, and personal communication with state officials. We provide Maine's reported numbers in ***Appendix B***. Counts of providers and practices reached are state reported numbers as of March 2016 for BHHs and HHs (CMS, 2016) and July 2017 for ACs (Maine Accountable Communities Webpage, 2018). Counts of populations reached are state reported numbers as of July 2017 for ACs (Maine Accountable Communities Webpage, 2018) and September 2017 for BHHs (Maine Department of Health and Human Services, 2017) and HHs (personal communication, June 6, 2018). Denominators used to compute percentage of population reached are Kaiser Family Foundation population estimates based on the Census Bureau's March 2017 Current Population Survey (Kaiser Family Foundation, 2018).

B-2.3.2 State evaluation calls

We conducted monthly federal evaluation-specific calls beginning in April 2014 and continued through the end of the SIM Initiative test period. The RTI/NASHP evaluation team for Maine, the state's SIM Initiative team, and the state's Innovation Center project officer typically

attended each state evaluation call. The purpose of the calls was to review interim evaluation findings with the state (as available), discuss any outstanding federal evaluation data or other needs, review and discuss state implementation and self-evaluation updates, and gather more in-depth information on select topics of interest from state officials leading the SIM Initiative in Maine.

For each meeting, the evaluation team prepared a list of state-specific questions, including the status of related policy levers and implementation successes, challenges, and lessons learned. We first reviewed relevant state documents for answers to our questions. When we did not find answers in the document or needed further clarification, we sent the questions to Maine ahead of the call and asked the state to have knowledgeable state officials available to answer the questions during the call.

B-2.3.3 Analysis

The RTI/NASHP evaluation team conducted thematic analysis of each source of qualitative data and then synthesized across information gleaned from site visits, focus groups, document review, and state evaluation calls. Site visit interviews and focus groups were conducted in previous evaluation years. For more detail on site visit and focus group methods, see past evaluation reports.

B-2.4 References

Centers for Medicare & Medicaid Services (CMS). (2016). *Maine State Innovation Model Progress Report, Quarter 1 2016*. Supplied by CMS.

Kaiser Family Foundation. (2018). *Health insurance coverage of the total population*. Available at http://kff.org/other/state-indicator/total-population/

Maine Accountable Communities Webpage. (2018). *Accountable Communities Initiative*. Available at http://www.maine.gov/dhhs/oms/vbp/accountable.html

Maine Department of Health and Human Services. (2017). *Maine State Innovation Model final report*. Available from the State of Maine.

Puhani, P. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters, 115*, 85–87.

[this page intentionally left blank]

Sub-appendix C-2. Methods for Massachusetts Analysis

The Massachusetts SIM Initiative Round 1 Evaluation team collected and analyzed a wide range of qualitative data in the fifth year of the federal SIM Initiative evaluation. These data sources included interviews with key informants and focus groups conducted during in-person site visits, a review of relevant documents, and regular evaluation calls with the state officials leading the state's SIM Initiative. ***Appendix C, Section C.1*** draws from previous site visits and the final site visit conducted in January 2018. ***Appendix C, Section C.2*** includes information only from the final site visit. See past evaluation reports for detail on previously conducted site visit interviews and focus groups. Methods for the final site visit interviews and focus groups are detailed below.

C-2.1 Site Visit Key Informant Interviews

The RTI/NASHP evaluation team conducted interviews with a variety of SIM Initiative stakeholders in Massachusetts, usually in person but sometimes by telephone. In the interviews, we focused on implementation successes, challenges, lessons learned, and model sustainability. Discussion topics included (1) policy impacts, (2) stakeholder participation, (3) health care delivery transformation, (4) payment system reform, (5) quality measurement and reporting, (6) population health, (7) health information technology and other infrastructure investments, (8) workforce and practice transformation, and (9) overall outcomes and lessons learned.

Stakeholders interviewed included the states' SIM Initiative teams, other state officials, managed care organizations, Pilot accountable care organization (ACO) providers, community partners, and e-Referral stakeholders. We solicited suggestions from the state SIM teams for interview candidates and identified additional candidates from review of relevant documents. We contacted interview candidates by e-mail or phone to offer them the opportunity to participate. Final lists of site visit interviewees were not shared with state SIM Initiative teams or CMS staff; the list remained confidential.

We held interviews in the offices or locations of the interview participants. All interviews were conducted by at least two evaluation team members. The interview lead used discussion guides to structure each interview, and a designated note taker recorded feedback from each session. We also audio-recorded each of the interviews to confirm the notes' accuracy and to clarify areas in the notes that were unclear; however, we did not transcribe the recordings. Prior to audio recording, we obtained permission from all interview participants and instructed them that recordings could be stopped at any time.

Different discussion guides were used for each major type of stakeholder and tailored for each state. The interviews were interactive; participants were encouraged to share feedback most relevant to their particular roles in the Massachusetts SIM Initiative. To encourage candid

discussion, we were clear that we would not identify the specific interview participants or attribute specific comments to individuals in subsequent reporting. Specific interview sessions typically lasted no more than 1 hour.

The Massachusetts team conducted 27 total interviews in Boston and Worcester during the week of January 22–25, 2018. Interviewees gave their perspective on the various components of the SIM Initiative and the Full ACO, focusing especially on the Pilot ACO program, the Full ACO program, and e-Referral. ***Table C-2-1*** provides a distribution of the completed interviews by interviewee type.

Table C-2-1. Key informant interviews conducted in Massachusetts, January 2018

Key informant	Number of interviews
State officials	7
Managed care organizations	2
Accountable care organization providers	11
Community Partners	4
E-Referral stakeholders	3
Total	27

C-2.2 Focus Groups

The Massachusetts team also conducted focus groups with providers and consumers involved in the pilot ACO. The providers selected for focus groups were primary care providers (PCPs) in Pilot ACOs in Boston and Worcester; a total of 32 PCPs participated in four focus groups. The consumers were Medicaid beneficiaries attributed to Pilot ACOs in Boston and Worcester; a total of 20 beneficiaries participated in three focus groups.

Recruitment. At the request of the evaluation team, Massachusetts created lists of potentially eligible consumer and provider focus group participants. The state sent a letter to the targeted population asking them to call The Henne Group if they were interested in participating. The Henne Group screened participants by phone based on the eligibility criteria developed by the evaluation team. If participants were eligible and interested, The Henne Group scheduled them for a focus group. We sought to recruit nine participants and two alternates for each group.

Methods. Prior to the start of the group, all participants were required to sign a consent form that outlined the purpose of the discussion and how the information from the discussion would be used and stated that the discussions would be audio-recorded. We used a semistructured moderator guide, audio-recorded the discussions, took notes during the groups for analysis purposes, and had the audio recordings transcribed verbatim. The consumer focus

groups lasted 90 minutes, and the provider groups lasted 1 hour. At the conclusion of the group, we provided $75 to each consumer and $300 to each provider as a gesture of appreciation for their time.

The purpose of the focus groups was to understand consumers' and providers' current experience and reflections of care delivery during the Pilot ACO and the changes they have observed over time. To capture this information, the moderator's guide addressed consumer and provider perspectives on quality of care, care coordination, delivery reform, and provider reaction to opportunities for participation in new delivery systems, payment models, or other infrastructure supports.

C-2.3 Document Review

We used Massachusetts' quarterly and annual reports, operational plans, self-evaluation reports, and other state documents to obtain updated information on their implementation progress during the SIM Initiative test period. To supplement these documents, we collected relevant news articles on Massachusetts' SIM Initiative activities and related initiatives, and we searched reform-oriented websites that the state maintains.

In addition, we obtained the number of population served by the Full ACO model from personal communication with the state. The denominator used to compute percentage of population reached is a Kaiser Family Foundation population estimate of the Medicaid population based on the Census Bureau's March 2017 Current Population Survey (Kaiser Family Foundation, 2018).

C-2.4 State Evaluation Calls

We conducted monthly federal evaluation-specific calls beginning in April 2014 and continued through the end of the SIM Initiative test period. The RTI/NASHP evaluation team for Massachusetts, the state's SIM Initiative team, and the state's Innovation Center project officer typically attended each state evaluation call. The purpose of the calls was to review interim evaluation findings with the state (as available), discuss any outstanding federal evaluation data or other needs, review and discuss state implementation and self-evaluation updates, and gather more in-depth information on select topics of interest for the evaluation.

For each meeting used to collect additional information and perspective from state officials leading the SIM Initiative in Massachusetts, the evaluation team prepared a list of state-specific questions—including the status of related policy levers and implementation successes, challenges, and lessons learned. We first reviewed relevant state documents for answers to our questions. When we did not find answers in the document or needed further clarification, we sent the questions to the state ahead of the call and asked the state to have knowledgeable state officials available to answer the questions during the call.

C-2.5 Analysis

The RTI/NASHP evaluation team conducted thematic analysis of each source of qualitative data and then synthesized across information gleaned from site visits, focus groups, document review, and state evaluation calls. For example, for the focus group data, the team examined the transcripts of each focus group to identify emerging themes for consumer and provider groups and produced an internal topline report to guide further state analyses. Members of the state team who were present at the groups reviewed the topline reports and provided feedback. Using the final topline reports from the focus groups and other qualitative data collection activity, the team produced the synthesized analysis contained in this report.

C-2.6 References

Kaiser Family Foundation. (2018). *Health insurance coverage of the total population*. Available at http://kff.org/other/state-indicator/total-population/

Sub-appendix D-2. Methods for Minnesota Analyses

D-2.1 Methods for the Minnesota IHP Impact Analysis Using DHS Medicaid Claims

To estimate the impact of the Integrated Health Partnerships (IHPs) among Medicaid beneficiaries in Minnesota, we conducted difference-in-differences (D-in-D) regression analyses comparing beneficiaries attributed to an IHP to those not attributed to an IHP that were otherwise eligible. The Minnesota Department of Human Services (DHS) provided Medicaid data that indicated which beneficiaries were attributed to an IHP during 2013, 2014, 2015, and 2016. We replicated attribution with a pool of beneficiaries who were eligible for, but not attributed to, IHPs during these same time periods to generate a comparison group. We conducted descriptive trend and D-in-D analyses for outcomes for three evaluation domains: (1) care coordination, (2) service utilization, and (3) quality of care. The Medicaid claims provided by the Minnesota DHS did not have reliable expenditure data and we therefore conducted a complementary analyses of costs using expenditure data from the Minnesota All Payer Claims Database (MN APCD). The methods and additional details around these analyses are presented in ***Section D-2.2***.

IHPs in the context of Minnesota Medicaid. In Minnesota, approximately 14 percent of the population is covered by Medicaid. Minnesota has a longstanding Medicaid managed care program and in 2014, just over 70 percent of Medicaid beneficiaries were enrolled in a managed care organization.[17]

The Minnesota IHPs were created through 2010 legislation and function as the Medicaid accountable care organization (ACO). IHP implementation introduced the opportunity for groups of providers to share one- or two-sided risk with the Medicaid program, regardless of their contracts with Minnesota's Medicaid MCOs; thus, Medicaid managed care enrollment does not determine nor exclude eligibility for IHP attribution.[18] By 2017, there were 21 participating IHPs in the program covering 58 percent of the total Minnesota Medicaid population. Among these covered beneficiaries, the majority were receiving care from one of the more than 10,000 IHP providers. To achieve such high levels of participation, DHS expanded its contracts with providers each year of the demonstration.

[17] Source: 2014 Medicaid Managed Care Enrollment Report, Centers for Medicare and Medicaid Services, U.S. Department of Health and Human Services, Spring 2016. Summarized and available from: https://www.kff.org/medicaid/state-indicator/total-medicaid-mc-enrollment

[18] For additional details on how IHPs work in this managed care environment, refer to ***Appendix D, Section D.1***.

Between 22 and 33 percent of the Medicaid population is *not* eligible to be attributed to an IHP. As stated in the state's Payment Model Overview,[19] Minnesotans not eligible for IHP attribution (and therefore excluded from both IHP and comparison groups) include those with:

- No health care home or evaluation and management claims (3 to 5 percent) during the attribution year
- Medicare eligibility, or enrollment in partial benefit plans such as the Family Planning Program or Emergency Medical Assistance Program (12 to 18 percent) during the attribution year
- Limited enrollment duration, including fewer than 6 months of continuous enrollment or fewer than 9 months of noncontinuous enrollment (7 to 10 percent) during the attribution year.

Of IHP-eligible Minnesotans, 85 percent are enrolled with a Medicaid managed care plan as compared to 70 percent of the total Minnesotan Medicaid population.

Profile of IHP participating providers. IHPs are a diverse group, with some representing large integrated delivery systems (known as "integrated IHPs") while others are smaller provider-led organizations (known as "virtual IHPs"). For our analysis, both integrated and virtual IHPs are combined into a single treatment group. Implementation of the first six IHPs, whose contracts with Medicaid started on January 1, 2013, occurred prior to the SIM Initiative implementation period beginning October 1, 2013. These first six IHPs included North Memorial Health Care, Centracare Health System, Essentia Health, Children's Hospitals and Clinics of Minnesota, and Federally Qualified Health Center Urban Health Network (known as FUHN). Three more IHPs received contracts starting January 1, 2014: the Mayo Clinic, Southern Prairie Community Care (providers in rural counties), and Hennepin Health, which is a county-based provider that had been operating as an ACO under a Medicaid waiver for two years prior to being recognized as an IHP. Seven more IHPs received contracts starting January 1, 2015: Lake Region Health Care, Lakewood Health System, Wilderness Health, Winona Health, Mankato Clinic, Courage Kenny (Allina), and Bluestone Physician Services. Three more IHPs received contracts starting January 1, 2016: Integrity Health Network, Allina Health System, and Gillette Children's Specialty Healthcare.

Minnesota leveraged its SIM activities to get broad participation in IHPs over time, which increased the diversity of providers involved in IHPs as more joined the program annually since 2013. DHS designed the IHP requirements to offer flexibility in how providers participated, with the intent of reducing barriers to entry in the program.[20] Some IHPs represent providers treating specific populations, such as people with disabilities or children. IHPs are

[19] http://www.dhs.state.mn.us/main/idcplg?IdcService=GET_FILE&RevisionSelectionMethod=LatestReleased&Rendition=Primary&allowInterrupt=1&noSaveAs=1&dDocName=dhs16_177106

[20] http://www.dhs.state.mn.us/main/groups/sim/documents/pub/dhs16_197631.pdf

present throughout the state, in both urban and rural areas. The only information with which we can compare providers who participated in an IHP by 2016 versus those who had not are the typical claims-based information (e.g., provider specialty). The intent of our analysis is to determine whether, as of 2016, those providers in an IHP are yielding better outcomes for IHP-eligible Medicaid beneficiaries than those providers who are not in IHPs. As noted earlier, IHPs represent a diverse group of providers. Although we do not have information to account for practice-level variation, we use beneficiary and county-level factors to balance the IHP and comparison groups.

Study design. Our analysis compares the pre-period (2010–2012) and post-period (2013–2016) trends for the IHP and comparison groups using a longitudinal design with an unbalanced panel. That is, we used all available data for beneficiaries attributed to the IHP and comparison groups in any given year and did not restrict our analysis to beneficiaries who had continuous enrollment in Medicaid. This means we included beneficiaries who were eligible for Medicaid for the first time in the year of their attribution. The rationale for an unbalanced panel—as opposed to using a balanced panel approach—is to provide estimation of the program's impact encompassing the entire population of attributed beneficiaries—not a subset based on prior eligibility. The disadvantage of an unbalanced design, however, is the inclusion of beneficiaries without baseline observational data prior to their attribution. In the Minnesota Medicaid claims analyses, 21 percent of the sample does have a full panel of data (i.e., are observed in each year from 2010 through 2016). To ensure that our comparison group and IHP group are comparable, we conduct a test to determine if trends in key outcomes were similar (parallel) prior to the start of the IHP program. As shown in ***Figures D-2-2*** through ***D-2-3***, we pass the parallel trends test in the baseline period. And finally, as shown in the balance tables by year (***Tables D-2-2*** through ***D-2-13***), average beneficiary characteristics do not differ substantially year to year within the IHP or comparison group, suggesting that even though some beneficiaries may not have baseline data, the characteristics of the sample are stable over time. Lastly, if a beneficiary was ever attributed to an IHP, they were excluded from the comparison group, regardless of attribution year. The difference in the changes over time from the pre-period to the post-period between the IHP group and comparison group provides an estimate of the impact of the IHP in its first four years of implementation.

Profiles of IHP and comparison groups. Minnesota attributes beneficiaries yearly, retrospectively. A beneficiary is attributed to an IHP if (1) the beneficiary received a health care home service (billing codes S0280 or S0281) from any provider in the IHP, or, (2) the beneficiary did not receive any health care home services during the attribution year, they received the plurality of primary care services during the attribution year from providers in the IHP. In cases where a beneficiary was attributed to an IHP in one year, but received no primary care services in a subsequent year, the lookback period will be extended to 2 years in an effort to not penalize IHPs for effectively managing patients outside of the office setting. Dually eligible beneficiaries in Medicaid and Medicare were not eligible to participate in IHPs. As noted earlier,

we received beneficiary-level files from Minnesota's DHS, which included IHP attribution status for each year between 2013 and 2016. Additionally, the file noted which beneficiaries were eligible for participation but not attributed to an IHP. The state also provided a list of the organizational NPIs for practices participating in an IHP. Prior to attributing the comparison group we excluded beneficiaries that were ever attributed to an IHP during 2013–2016 to prevent comparison group beneficiaries from becoming IHP group beneficiaries when including additional years in future analyses. Subsequently, we used professional claims and Minnesota's IHP attribution methodology to attribute eligible but non-IHP-attributed beneficiaries to a non-IHP participating provider based on receipt of either health care home services, or the plurality of their primary care. Comparison group attribution was done separately for each demonstration year (2013–2016) to confirm that (1) each beneficiary in the comparison group had at least one health care home or evaluation and management service in a given year (a requirement to be eligible for IHP-attribution), and (2) the comparison group beneficiaries received the plurality of their primary care services from a non-IHP provider during that year. If a beneficiary could not be attributed to either an IHP organization or a non-IHP provider during any given year, their claims were omitted from the analysis for that year. More information on sample construction is available in ***Section D-2.1.3.***

Subpopulations. In addition to the analysis on the overall population, we conducted two subpopulation analyses: (1) children and adults separately and (2) beneficiaries diagnosed with behavioral health conditions. Children are an important subpopulation to look at with respect to the Medicaid population because two IHPs were pediatric focused: Children's Hospitals and Clinics of Minnesota (2013) and Gillette Children's Specialty Healthcare (2016). As part of Minnesota's SIM Initiative, there was considerable effort made in integrating behavioral health. In light of these efforts, and Minnesota's successes, it is important to investigate whether there were similar findings among beneficiaries diagnosed with behavioral health conditions, relative to the full population.

Balancing IHP and comparison groups. Because Medicaid beneficiaries were not randomly assigned to IHPs or the comparison group, there may be observed sociodemographic and geographic differences between IHP-attributed beneficiaries and comparison group beneficiaries that may influence results. To address this, we used propensity score weighting to statistically adjust the study sample to reduce these differences. To apply propensity score weighting, we first used logistic regression to predict a Medicaid beneficiary's likelihood of being attributed to an IHP based on select sociodemographic and geographic characteristics. This predicted likelihood is known as the propensity score. We then took the inverse of the propensity score using the formula (propensity score/(1-propensity score)) to create what is known as the inverse probability of treatment weight. We then applied each comparison group member's inverse probability of treatment weight to our regression models. IHP-attributed beneficiaries receive an inverse probability of treatment weight of one. By applying these weights, comparison group beneficiaries are made to look more like the IHP beneficiaries. After propensity score

weighting, the standardized differences between the weighted comparison group and IHP group means were all well under the standard 10 percent threshold. More information on propensity score weighting is available in ***Sections D-2.1.4.***

Regression-adjusted difference-in-differences outcomes. Analyses used logistic regression for binary outcomes. All analyses used clustered standard errors at the beneficiary level to account for repeated observations from the same beneficiaries over time.

Even though clustering at an organizational level is a commonly applied strategy for obtaining unbiased standard errors in D-in-D models (Bertrand, Duflo, & Mullainathan, 2004), we did not do this because accurately identifying organizational clusters over time would require making several ad hoc assumptions to track organization NPIs across observation periods and because organization NPIs do not represent all of the treatment providers with whom beneficiaries actually engage. Accordingly, these factors would greatly reduce our confidence that clustering at an organizational level is correcting the bias in our standard errors. Furthermore, it is likely that this is a minor limitation because ignoring organizational-level interdependence is associated with a downward bias in standard errors in D-in-D models, suggesting that some hypotheses with p-values just below 0.10 should in fact be rejected. However, we had virtually no marginally significant findings (i.e., with p-values just below 0.10). In fact, many of our statistically significant findings would remain significant even if standard errors were downwardly biased by as much as 100 percent. Considering these factors, we clustered all models at the beneficiary level.

The outcome models controlled for demographic, health plan, health status, and county-level characteristics. More information on outcomes is available in ***Section D-2.1.2.*** More information on the regression model is available in ***Section D-2.1.5***.

D-2.1.1 Data sources

Medicaid data. We used Medicaid claims data provided by the Minnesota Department of Human Services (DHS) to derive eligibility information and measure claims-based outcomes. In this report, we used data from 2010 to 2016 to examine the 3 years before (2010–2012) and 4 year after (2013–2016) the start of the IHP model. The Medicaid claims data included three linkable types of files: (1) an enrollment file, containing beneficiary characteristics, monthly enrollment indicators, and coverage information; (2) a provider file, containing National Provider Identifier, specialty, and name; and (3) Medicaid claims files, including inpatient, outpatient, professional, and pharmaceutical claims. These files include information for Medicaid beneficiaries attributed to an IHP in years 2013–2016 and those that were not attributed but were otherwise eligible. The analytic sample included individuals of all ages and excluded Medicare-Medicaid enrollees.

Attribution file. We received a list of IHP Medicaid beneficiaries attributed to an IHP in each year of the baseline and intervention years. We also received a list of beneficiaries that the state designated as eligible but not participating in an IHP. Beneficiaries in this eligible but not participating group form the comparison group for our analysis. Provider attribution was independent across years, and we received annual practice participation lists for each year of the intervention period as provider participation was independent across years. The Minnesota DHS conducts attribution on a rolling monthly basis, looking back 12 months to determine if a beneficiary meets the attribution criteria to be attributed to any of the IHPs with active contracts. This means that in the data provided by the state, newly on-boarded IHPs have some beneficiaries with attribution flags indicating that they were attributed prior to that IHP's start date. To address this, we merged on start dates for each IHP and only included beneficiaries who are attributed to IHPs after each IHP's start date in the intervention group for each year. Beneficiaries included in the attribution lists were linked to the enrollment and claims data to form the analysis sample. Only beneficiaries attributed to an IHP or the comparison group in at least 1 year were included in the sample.

Area Health Resource File (AHRF). The AHRF comprises data collected by the Health Resources and Services Administration from more than 50 sources containing more than 6,000 variables related to health care access at the county level. We used 2010 and 2012 information on health professions supply, population characteristics, and economic data as covariates in the outcome model analyses.

D-2.1.2 Outcome measures

Care coordination

To evaluate the impact of the IHP demonstration in Minnesota on care coordination, we report the following care coordination measures. Each measure was calculated annually for all eligible beneficiaries in the IHP group and comparison group overall and for children and adults separately.

- **Percentage of beneficiaries with any visit to a primary care provider:** This is an indicator for whether the beneficiary had at least one visit to a primary care provider reported in the medical claims file for the year, divided by the number of beneficiaries in the same year. Primary care physicians were identified using their primary taxonomy code, which was obtained from the National Plan and Provider Enumeration System (NPPES) file. A taxonomy code was considered primary where it was denoted in the NPPES file with a Y or an X. When searching for primary care visits, claims were restricted to those with Healthcare Common Procedure Coding System (HCPCS)/Current Procedural Terminology (CPT) codes indicating evaluation and management (E&M) visits associated with planned physician care (i.e., office visits). Both inpatient and outpatient files were included, although E&M codes used

to identify physician visits should occur only in the outpatient file. ***Table D-2-1*** below provides a cross-walk between provider taxonomy codes and the categorization of either primary care physician or specialty provider designation.

Table D-2-1. Cross-walk between provider taxonomy codes and primary care physician (PCP) or specialty provider (SPE) designation

Taxonomy	PCP/SPE	Type	Classification	Specialization
207K00000X	SPE	Allopathic & Osteopathic Physicians	Allergy & Immunology	
207L00000X	SPE	Allopathic & Osteopathic Physicians	Anesthesiology	
207LP2900X	SPE	Allopathic & Osteopathic Physicians	Anesthesiology	Pain Medicine
207LC0200X	SPE	Allopathic & Osteopathic Physicians	Anesthesiology	Critical Care Medicine
208U00000X	SPE	Allopathic & Osteopathic Physicians	Clinical Pharmacology	
208C00000X	SPE	Allopathic & Osteopathic Physicians	Colon & Rectal Surgery	
207N00000X	SPE	Allopathic & Osteopathic Physicians	Dermatology	
207P00000X	SPE	Allopathic & Osteopathic Physicians	Emergency Medicine	
207PE0004X	SPE	Allopathic & Osteopathic Physicians	Emergency Medicine	Emergency Medical Services
207PT0002X	SPE	Allopathic & Osteopathic Physicians	Emergency Medicine	Medical Toxicology
207Q00000X	**PCP**	Allopathic & Osteopathic Physicians	Family Medicine	
207QS0010X	SPE	Allopathic & Osteopathic Physicians	Family Medicine	Sports Medicine
207QA0000X	**PCP**	Allopathic & Osteopathic Physicians	Family Medicine	Adolescent Medicine
207QA0505X	**PCP**	Allopathic & Osteopathic Physicians	Family Medicine	Adult Medicine
208D00000X	**PCP**	Allopathic & Osteopathic Physicians	Family Medicine	General Practice
208M00000X	SPE	Allopathic & Osteopathic Physicians	Hospitalist	
207R00000X	**PCP**	Allopathic & Osteopathic Physicians	Internal Medicine	
207RR0500X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Rheumatology
207RC0000X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Cardiovascular Disease
207RX0202X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Medical Oncology
207RG0100X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Gastroenterology
207RE0101X	**PCP**	Allopathic & Osteopathic Physicians	Internal Medicine	Endocrinology
207RH0003X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Hematology & Oncology
207RI0200X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Infectious Disease
207RH0000X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Hematology
207RP1001X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Pulmonary Disease

(continued)

Table D-2-1. Cross-walk between provider taxonomy codes and primary care physician (PCP) or specialty provider (SPE) designation (continued)

Taxonomy	PCP/SPE	Type	Classification	Specialization
207RN0300X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Nephrology
207RI0011X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Interventional Cardiology
207RC0200X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Critical Care Medicine
207RC0001X	SPE	Allopathic & Osteopathic Physicians	Internal Medicine	Clinical Cardiac Electrophysiology
207RG0300X	**PCP**	Allopathic & Osteopathic Physicians	Internal Medicine	Geriatric Medicine
207RH0002X	**PCP**	Allopathic & Osteopathic Physicians	Internal Medicine	Hospice and Palliative Medicine
207SG0201X	SPE	Allopathic & Osteopathic Physicians	Medical Genetics	Clinical Genetics (M.D.)
207T00000X	SPE	Allopathic & Osteopathic Physicians	Neurological Surgery	
207V00000X	**PCP**	Allopathic & Osteopathic Physicians	Obstetrics & Gynecology	
207VM0101X	SPE	Allopathic & Osteopathic Physicians	Obstetrics & Gynecology	Maternal & Fetal Medicine
207VX0201X	SPE	Allopathic & Osteopathic Physicians	Obstetrics & Gynecology	Gynecologic Oncology
207W00000X	SPE	Allopathic & Osteopathic Physicians	Ophthalmology	
207X00000X	SPE	Allopathic & Osteopathic Physicians	Orthopedic Surgery	
207XS0106X	SPE	Allopathic & Osteopathic Physicians	Orthopedic Surgery	Hand Surgery
207XX0005X	SPE	Allopathic & Osteopathic Physicians	Orthopedic Surgery	Sports Medicine
207Y00000X	SPE	Allopathic & Osteopathic Physicians	Otolaryngology	
207ZP0105X	SPE	Allopathic & Osteopathic Physicians	Pathology	Clinical Pathology/Laboratory Medicine
207ZP0102X	SPE	Allopathic & Osteopathic Physicians	Pathology	Anatomic Pathology & Clinical Pathology
207ZN0500X	SPE	Allopathic & Osteopathic Physicians	Pathology	Neuropathology
207ZH0000X	SPE	Allopathic & Osteopathic Physicians	Pathology	Hematology
207ZB0001X	SPE	Allopathic & Osteopathic Physicians	Pathology	Blood Banking & Transfusion Medicine
208000000X	**PCP**	Allopathic & Osteopathic Physicians	Pediatrics	
2080P0205X	**PCP**	Allopathic & Osteopathic Physicians	Pediatrics	Pediatric Endocrinology

(continued)

Table D-2-1. Cross-walk between provider taxonomy codes and primary care physician (PCP) or specialty provider (SPE) designation (continued)

Taxonomy	PCP/SPE	Type	Classification	Specialization
2080P0207X	SPE	Allopathic & Osteopathic Physicians	Pediatrics	Pediatric Hematology-Oncology
2080P0006X	SPE	Allopathic & Osteopathic Physicians	Pediatrics	Developmental—Behavioral Pediatrics
2080P0202X	SPE	Allopathic & Osteopathic Physicians	Pediatrics	Pediatric Cardiology
2080N0001X	SPE	Allopathic & Osteopathic Physicians	Pediatrics	Neonatal-Perinatal Medicine
2080P0203X	SPE	Allopathic & Osteopathic Physicians	Pediatrics	Pediatric Critical Care Medicine
208100000X	SPE	Allopathic & Osteopathic Physicians	Physical Medicine & Rehabilitation	
2081P0004X	SPE	Allopathic & Osteopathic Physicians	Physical Medicine & Rehabilitation	Spinal Cord Injury Medicine
2081P0010X	SPE	Allopathic & Osteopathic Physicians	Physical Medicine & Rehabilitation	Pediatric Rehabilitation Medicine
2081P2900X	SPE	Allopathic & Osteopathic Physicians	Physical Medicine & Rehabilitation	Pain Medicine
2083X0100X	SPE	Allopathic & Osteopathic Physicians	Preventive Medicine	Occupational Medicine
2083P0901X	SPE	Allopathic & Osteopathic Physicians	Preventive Medicine	Public Health & General Preventive Medicine
2083P0500X	SPE	Allopathic & Osteopathic Physicians	Preventive Medicine	Preventive Medicine/ Occupational Environmental Medicine
2084N0400X	SPE	Allopathic & Osteopathic Physicians	Psychiatry & Neurology	Neurology
2084P0800X	SPE	Allopathic & Osteopathic Physicians	Psychiatry & Neurology	Psychiatry
2084A0401X	SPE	Allopathic & Osteopathic Physicians	Psychiatry & Neurology	Addiction Medicine
2085R0001X	SPE	Allopathic & Osteopathic Physicians	Radiology	Radiation Oncology
2085R0202X	SPE	Allopathic & Osteopathic Physicians	Radiology	Diagnostic Radiology
2085R0203X	SPE	Allopathic & Osteopathic Physicians	Radiology	Therapeutic Radiology
2085R0204X	SPE	Allopathic & Osteopathic Physicians	Radiology	Vascular & Interventional Radiology
208600000X	SPE	Allopathic & Osteopathic Physicians	Surgery	

(continued)

Table D-2-1. Cross-walk between provider taxonomy codes and primary care physician (PCP) or specialty provider (SPE) designation (continued)

Taxonomy	PCP/SPE	Type	Classification	Specialization
2086S0122X	SPE	Allopathic & Osteopathic Physicians	Surgery	Plastic and Reconstructive Surgery
2086S0129X	SPE	Allopathic & Osteopathic Physicians	Surgery	Vascular Surgery
2086S0127X	SPE	Allopathic & Osteopathic Physicians	Surgery	Trauma Surgery
208G00000X	SPE	Allopathic & Osteopathic Physicians	Thoracic Surgery (Cardiothoracic Vascular Surgery)	
208800000X	SPE	Allopathic & Osteopathic Physicians	Urology	
261Q00000X	**PCP**	Ambulatory Health Care Facilities	Clinic/Center	
101YM0800X	SPE	Behavioral Health & Social Service Providers	Counselor	Mental Health
103T00000X	SPE	Behavioral Health & Social Service Providers	Psychologist	
1041C0700X	SPE	Behavioral Health & Social Service Providers	Social Worker	Clinical
104100000X	SPE	Behavioral Health & Social Service Providers	Social Worker	
111N00000X	SPE	Chiropractic Providers	Chiropractor	
111NI0013X	SPE	Chiropractic Providers	Chiropractor	Independent Medical Examiner
133V00000X	SPE	Dietary and Nutritional Service Providers	Dietitian, Registered	
152W00000X	SPE	Eye and Vision Services Providers	Optometrist	
291U00000X	SPE	Laboratories	Clinical Medical Laboratory	
176B00000X	**PCP**	Other Service Providers	Midwife	
174400000X	SPE	Other Service Providers	Specialist	
367A00000X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Advanced Practice Midwife	
367H00000X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Anesthesiologist Assistant	
364SM0705X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Clinical Nurse Specialist	Medical-Surgical
364SP0809X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Clinical Nurse Specialist	Psych/Mental Health
364S00000X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Clinical Nurse Specialist	

(continued)

Table D-2-1. Cross-walk between provider taxonomy codes and primary care physician (PCP) or specialty provider (SPE) designation (continued)

Taxonomy	PCP/SPE	Type	Classification	Specialization
364SA2200X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Clinical Nurse Specialist	Adult Health
364SP0807X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Clinical Nurse Specialist	Psych/Mental Health, Child & Adolescent
364SP0808X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Clinical Nurse Specialist	Psych/Mental Health
364SN0000X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Clinical Nurse Specialist	Neonatal
367500000X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Nurse Anesthetist	Certified Registered
363LF0000X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Family
363LP0200X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Pediatrics
363L00000X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	
363LA2200X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Adult Health
363LW0102X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Women's Health
363LG0600X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Gerontology
363LP0808X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Psych/Mental Health
363LX0001X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Obstetrics & Gynecology
363LN0005X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Neonatal Critical Care
363LN0000X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Neonatal
363A00000X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Physician Assistant	
363AM0700X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Physician Assistant	Medical
363AS0400X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Physician Assistant	Surgical
363LP0222X	SPE	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Pediatrics, Critical Care

(continued)

Table D-2-1. Cross-walk between provider taxonomy codes and primary care physician (PCP) or specialty provider (SPE) designation (continued)

Taxonomy	PCP/SPE	Type	Classification	Specialization
363LP2300X	**PCP**	Physician Assistants & Advanced Practice Nursing Providers	Nurse Practitioner	Primary Care
213E00000X	SPE	Podiatric Medicine & Surgery Service Providers	Podiatrist	
213ES0103X	SPE	Podiatric Medicine & Surgery Service Providers	Podiatrist	Foot & Ankle Surgery
213ES0131X	SPE	Podiatric Medicine & Surgery Service Providers	Podiatrist	Foot Surgery
225100000X	SPE	Rehabilitative & Restorative Service Providers	Developmental	Physical Therapist
390200000X	**PCP**	Student in an Organized Health Care Training Program		
333600000X	SPE	Suppliers	Pharmacy	

- **Percentage of beneficiaries with any visit to a specialty provider:** This is an indicator of whether the beneficiary had at least one visit to a specialty provider reported in the medical claims file for the year, divided by the number of beneficiaries in the same year. Specialty care physicians were identified using their primary taxonomy code, which was obtained from the NPPES file. A taxonomy code was considered primary where it was denoted in the NPPES file with a Y or an X. When searching for specialty care visits, claims were restricted to those with HCPCS/CPT codes indicating E&M visits associated with planned physician care (i.e., office visits). Both inpatient and outpatient files were included, although E&M codes used to identify physician visits should occur only in the outpatient file.

Percentage of acute inpatient hospital admissions with a follow-up visit within 14 days: This measure calculates the percentage of patients who had an acute care hospitalization[21] and who had a qualifying evaluation and management outpatient visit within 14 days of discharge. For an acute care hospitalization to be included in the denominator, the beneficiary was required to (1) be eligible at both the time of admission and 14 days post-discharge, (2) be alive both at discharge and 14 days post-discharge, and (3) not have a readmission within 14 days post-discharge. We used the following CPT codes to identify a follow-up visit:

- 99201 through 99205; New Patient, Office/Other Outpatient Services
- 99211 through 99215; Established Patient, Office/Other Outpatient Services

[21] Psychiatric hospital admissions are included in the short-term acute care hospitalizations. They cannot be identified separately in the Minnesota Medicaid data.

- 99241 through 99245; Evaluation and Management Consultation Services
- 99341 through 99350; Home-based ambulatory care visits
- 99381 through 99387; New Patient, Preventive Medicine Services
- 99391 through 99397; Established patient, Preventive Medicine Services
- 99401 through 99412; New or established patient; Counseling Risk Factor Reduction and Behavior Change Intervention
- 99420 through 99429; Other preventive medicine services, Counseling Risk Factor Reduction and Behavior Change Intervention
- G0402, G0438, G0439, G0463; Preventive Physical Examination and Wellness Visits

Percentage of patients age 5–64 years with persistent asthma who were appropriately prescribed medication during the year: This is the percentage of patients age 5 to 64 years with persistent asthma who were appropriately prescribed asthma medications. The denominator excludes persons with 11 or 12 months of enrollment during the measurement year. Persistent asthma patients were identified as follows:

- At least one ED visit with asthma as the principal diagnosis;
- At least one acute inpatient discharge with asthma as the principal diagnosis;
- At least four outpatient visits on different dates with asthma as a diagnosis (does not have to be primary) *and* at least two asthma medication dispensing events; or
- At least four asthma medication dispensing events.

If a beneficiary is only identified as having persistent asthma because of having four asthma medication dispensing events, where leukotriene modifiers were the sole asthma medication dispensed in the measurement year, they must also have at least one diagnosis of asthma in any setting, during the measurement year. Beneficiaries with emphysema, COPD, cystic fibrosis, or acute respiratory failure diagnoses are also excluded.

The numerator indicates whether each beneficiary in the denominator were dispensed at least one prescription for asthma controller medication during the measurement year. The list of asthma controller medications is a subset of the list of all asthma medications searched under the fourth criteria for inclusion in the denominator. ***Table D-2-2*** provides the list of asthma medications included in the fourth criteria for identifying the denominator.

Table D-2-3 lists the asthma controller medications included in the numerator.

Table D-2-2. Asthma medications

Description	Prescriptions		
Antiasthmatic combinations	• Dyphylline-guaifenesin	• Guaifenesin-theophylline	
Antibody inhibitors	• Omalizumab		
Inhaled steroid combinations	• Budesonide-formoterol	• Fluticasone-salmeterol	• Mometasone-formoterol
Inhaled corticosteroids	• Beclomethasone • Budesonide • Ciclesonide	• Flunisolide • Fluticasone CFC free • Mometasone	
Leukotriene modifiers	• Montelukast	• Zafirlukast	• Zileuton
Mast cell stabilizers	• Cromolyn		
Methylxanthines	• Aminophylline	• Dyphylline	• Theophylline
Short-acting, inhaled beta-2 agonists	• Albuterol	• Levalbuterol	• Pirbuterol
Long-acting, inhaled beta-2 agonists	• Aformoterol	• Formoterol	• Salmeterol

CFC = chlorofluorocarbon.

Table D-2-3. Asthma controller medications

Description	Prescriptions		
Antiasthmatic combinations	• Dyphylline-guaifenesin	• Guaifenesin-theophylline	
Antibody inhibitors	• Omalizumab		
Inhaled steroid combinations	• Budesonide-formoterol	• Fluticasone-salmeterol	• Mometasone-formoterol
Inhaled corticosteroids	• Beclomethasone • Budesonide • Ciclesonide	• Flunisolide • Fluticasone CFC free • Mometasone	
Leukotriene modifiers	• Montelukast	• Zafirlukast	• Zileuton
Mast cell stabilizers	• Cromolyn		
Methylxanthines	• Aminophylline	• Dyphylline	• Theophylline

CFC = chlorofluorocarbon.

Percentage of patients age 18 years and older diagnosed with a new episode of major depression and treated with antidepressant medication who remained on medication treatment at least 84 or 180 days (reported as separate measures): This is the percentage of patients 18 years of age and older who were diagnosed with a new episode of major depression and treated with antidepressant medication, and who remained on an antidepressant medication

treatment for 84 days (12 weeks) or 180 days (6 months), respectively. Two percentages are reported:

- *Effective Acute Phase Treatment.* This is the percentage of newly diagnosed and treated patients who remained on an antidepressant medication for at least 84 days (12 weeks).
- *Effective Continuation Phase Treatment.* This is the percentage of newly diagnosed and treated patients who remained on an antidepressant medication for at least 180 days (6 months).

For this measure, the Intake Period was defined as the 8-month window from May 1 through December 31 of the year prior to the measurement year. The Index Prescription Start Date (IPSD) was defined as the earliest prescription dispensing date for an antidepressant medication during the Intake Period. Antidepressant medications are listed in ***Table D-2-4*** below; specific national drug codes for these medications were identified via the 2015 list developed by the National Committee for Quality Assurance.

Table D-2-4. Antidepressant medications

Description	Prescription		
Miscellaneous antidepressants	• Bupropion • Vortioxetine	• Vilazodone	
Monoamine oxidase inhibitors	• Isocarboxazid • Phenelzine	• Selegiline • Tranylcypromine	
Phenylpiperazine antidepressants	• Nefazodone	• Trazodone	
Psychotherapeutic combinations	• Amitriptyline-chlordiazepoxide • Amitriptyline-perphenazine		• Fluoxetine-olanzapine
SNRI antidepressants	• Desvenlafaxine • Levomilnacipran	• Duloxetine	• Venlafaxine
SSRI antidepressants	• Citalopram • Escitalopram	• Fluoxetine • Fluvoxamine	• Paroxetine • Sertraline
Tetracyclic antidepressants	• Maptrotiline	• Mirtazapine	
Tricyclic antidepressants	• Amitriptyline • Amoxapine • Clomipramine	• Desipramine • Doxepin • Imipramine	• Nortriptyline • Protriptyline • Trimipramine

SNRI = serotonin-norepinephrine reuptake inhibitor; SSRI = selective serotonin reuptake inhibitor.

To identify patients for inclusion in the denominator, the patient had to be at least 18 years old at the beginning of the Intake Period; continuously enrolled in Medicaid for 3 months prior to the IPSD through 7 months following the IPSD with no more than a 1 month lapse in

coverage; and have a diagnosis for major depression (as defined by the ICD-9 and ICD-10 diagnosis codes per HEDIS measure specifications) that met at least one of the following criteria:

- An outpatient visit, intensive outpatient encounter, or partial hospitalization with any diagnosis of major depression.
- An ER visit with any diagnosis of major depression.
- An acute or nonacute inpatient claim/encounter with any diagnosis of major depression.

Patients were excluded from the denominator if they filled a prescription (as indicated by the Date Prescription Filled) in the 105 days prior to the IPSD.

To calculate numerators, we summed the number of days' supply for all identified antidepressant medications (see ***Table D-2-4***). Days that extended beyond the treatment window, defined below, were not counted, and overlapping prescriptions were summed.[22] Specifically:

- For the *Effective Acute Phase Treatment* numerator, we summed days' supply for all prescriptions where IPSD <= date prescription was filled <= IPSD + 114. If this sum was at least 84 days (12 weeks), the numerator was set to 1; otherwise, the numerator was set to 0.
- For the *Effective Continuation Phase Treatment* numerator, we summed days' supply for all prescriptions where IPSD <= date prescription was filled <= IPSD + 231. If this sum was at least 180 days (6 months), the numerator was set to 1; otherwise, the numerator was set to 0.

Utilization

Utilization measures are reported as rates per 1,000 covered lives (or discharges for readmissions). For each measure, we first calculate the probability of any use. To calculate the probability, the numerator was an indicator of having had at least one event (inpatient admission or ER visit that did not lead to a hospitalization) and the denominator is the number of eligible beneficiaries (or discharges) in the state enrolled during the period. We multiplied the probability of use by 1,000 to obtain approximate rates of utilization per 1,000 beneficiaries. Multiplying the probability by 1,000 does not produce an exact rate of utilization per 1,000 beneficiaries because it assumes no person has more than one visit or admission per year. However, we concluded that this is a reasonable approximation because only a small percentage of beneficiaries had counts exceeding one for any of the utilization measures. Events are included in a period's total if discharge or service date on the claim was during the period.

[22] The decision to sum overlapping days' supply was made by the analyst team. Determining the actual time frame covered by overlapping prescriptions would require significant inference as to how beneficiaries were taking their medications and when they began taking their medication after filling the prescription, both of which are beyond the scope of information provided in claims.

- **Rate (per 1,000 covered persons) of all-cause inpatient hospitalizations:** This is an indicator of whether the beneficiary had at least one admission to an acute-care hospital reported in the inpatient file for the year, divided by the number of beneficiaries in the same year. For Minnesota, we identified all-cause acute care hospital admissions using a combination of type of bill for inpatient services (11 or 12) and category of service for general inpatient, inpatient neo-natal ICU, and an undefined category (001, 073, or 999). We included the undefined category at the state's recommendation, because all inpatient encounters with a consolidated pay-to provider (a large percentage of the claims) are assigned to this category. Some records in the inpatient claims files may appear to be multiple admissions but are in fact transfers between facilities; these records were counted as a single admission. To roll up transfers into one acute admission, the RTI team first identified claims that had no more than 1 elapsed day between the discharge date of the index claim and the admission date of the subsequent claim. Then, these claims were combined into one record by taking the earliest admission date and latest discharge date and summing all payment amounts. This same roll-up procedure was applied to claims with overlapping or identical admission and discharge dates (i.e., claims associated with the same visit).
- **Rate of ER visits that did not result in an inpatient hospital admission (per 1,000 covered persons):** This is an indicator of whether the beneficiary had at least one visit to the ED that did not result in an inpatient hospital admission, divided by the number of beneficiaries in the same period. For all data sources, ED visits (including observation stays) are identified in the outpatient services file as visits with a revenue center line item equal to 045X, 0981 (ED care) or 0762 (treatment or observation room, thus including observation stays in the overall count). If the procedure code on every line item of the ED claim equaled 70000–89999 and no line items had a revenue center code equal to 0762, that claim was excluded (thus excluding claims for which only radiology or pathology/laboratory services were provided unless they were observation stays). Multiple ED visits on a single day were counted as a single visit. If there was an inpatient ED visit on the same day, the outpatient encounter was excluded.
- **Rate of 30-day readmissions (per 1,000 discharges):** This is an indicator of whether the beneficiary had at least one acute hospitalization that occurred within 30 days following a live discharge for beneficiaries ages 18 or older for the year, divided by the number of inpatient discharges in the same year. Index hospital discharges were identified as inpatient stays with a discharge date within the given measurement period (12 months) minus 30 days from the end of the period. The RTI team counted the number of instances when the beneficiary had an inpatient readmission within 30 days of the index stay discharge.

Quality of care

To evaluate the impact on quality of care, we report the following quality measures. The measures were calculated annually for all eligible beneficiaries in the IHP and comparison group.

Percentage of patients age 18–75 years with diabetes (type 1 and type 2) who had Hemoglobin A1c (HbA1c) testing: The denominator included beneficiaries age 18 to 75 with at least 11 months of enrollment and who were enrolled in the last month of each measurement year. Diabetes was identified under the following criteria:

- In the outpatient claims file (Bill Type 11 or 12): At least two outpatient visits, observation visits, ER visits, or nonacute inpatient encounters on different dates of services, with a diagnosis of diabetes.
- In the inpatient claims file (Bill Type 11 or 12): At least one acute inpatient visit with a diagnosis of diabetes.

The numerator is set to 1 if the beneficiary in the denominator had a procedure code in the HbA1c Tests value set during the measurement year.

D-2.1.3 Populations studied

The study population includes Medicaid beneficiaries attributed to IHPs at any point between 2013–2016. The Minnesota DHS provided a list of IHP-attributed beneficiaries. The Minnesota DHS conducts retrospective attribution on a monthly basis and finalizes attributed beneficiaries for financial reconciliation purposes at the end of the second quarter of the subsequent year. We replicated the state's attribution method in identifying the comparison group, and additional details of the methodology are presented in Section F-2.4. In summary, 12-months of claims were analyzed for all eligible beneficiaries, and were attributed to IHPs if (1) the plurality of their primary care practice or specialty physician visits were with an IHP-affiliated provider, or (2) they had health home visits with an IHP-affiliated provider. Eligibility was determined based on having continuous Medicaid enrollment for at least six months, or a total of 9+ months of discontinuous enrollment and not being dually eligible for Medicare.

In addition to the overall population of IHP-attributed beneficiaries, we also conducted subpopulation-specific analyses of children, adults, and beneficiaries with behavioral health conditions. Children were defined as persons aged 18 or younger, and adults were defined as persons aged 19–64 years old. Beneficiaries were identified as having mental or behavioral health conditions prior to their attribution to the IHP or comparison group. To be included in the subpopulation of beneficiaries with mental or behavioral health conditions, a beneficiary had to have two or more outpatient claims with a primary mental health or chemical dependency diagnosis or at least one inpatient claim with a primary diagnosis of mental health or chemical dependence.

Prior to 2014, Minnesota provided Medicaid coverage to childless adults up to 75 percent of the federal poverty line (FPL). In 2014, Minnesota expanded its Medicaid coverage to childless non-disabled adults whose household income does not exceed 138 percent of the FPL. We cannot identify which beneficiaries in our data are newly eligible as a result of this

expansion, but have no reason to believe the expansion affects the IHP and comparison groups differently. Because expansion of coverage to a new set of adults in the 75–138 percent FPL range occurred early in the post-period, it is possible that changes in outcomes in the post-period when compared to the baseline may have been due new beneficiaries in the sample, and possibly bias outcomes to show greater use of primary care due to preexisting demand for services. However, among all childless adults covered by Medicaid in 2013, fewer than 40 percent fall into the 75–138 percent FPL range,[23] and potentially even fewer will meet the basic eligibility requirements to be enrolled in an IHP (e.g., 6 months of continuous enrollment or 9 months of non-continuous enrollment) and therefore will not be eligible for our sample in either the IHP or comparison groups. Additionally, adults newly eligible in 2014 would still need 6 months of continuous or 9 months of noncontinuous enrollment to be attributed to the IHP or comparison group—restricting the proportion of the sample affected by the change in eligibility policy in 2014 to an even smaller group, i.e., those who enrolled in the first six months of 2014. Our model uses an unbalanced panel longitudinal design to maximize utilization of available data; that is, all available data for beneficiaries attributed to the IHP and comparison group in any given year were used for analysis. More information on study design and same construction is presented later in this section.

D-2.1.4 Comparison group and propensity score weighting

For the impact analysis, we used a pre-post comparison group design, in which the comparison group provides an estimate of what would have happened in the IHP group absent the effect on the intervention. The difference in the changes over time between the IHP group and its comparison group provides an estimate of the impact of the IHP. Ideally, the comparison group should be similar to the IHP group on all relevant dimensions (e.g., demographic, socioeconomic, political, regulatory, health, and health systems) except for the policy change being tested.

In the following section, we detail the procedures we used to select the comparison group for the IHP demonstration in Minnesota.

Selection of comparison group

As noted earlier, we received a list of all IHP eligible beneficiaries as well as an indication of IHP participation for years 2013, 2014, and 2015 from DHS. Prior to attributing the comparison group we excluded beneficiaries that were ever attributed to an IHP in any demonstration year (2013 through 2016). Attribution was determined by all professional medical claims (claim types 'A' or 'V'). Dental services were excluded. We then replicated Minnesota's

[23] Estimates based on issue brief available here: http://www.mnbudgetproject.org/research-analysis/economic-security/health-care/covering-more-minnesotans-through-medicaid-would-improve-health-outcomes-and-reduce-state-costs

IHP attribution methodology among potential comparison group beneficiaries for each demonstration year (2013 through 2016) separately.

There were three steps in attribution:

1. **Health Home Claim Code Attribution.** If procedure code S0280 or S0281 occurred on a line item, we attributed the beneficiary to the billing provider that had the most occurrences of those S codes for each year.

   For beneficiaries remaining unattributed, go to step 2.

2. **Primary Care E&M Attribution.** Line items from the professional claims were selected if the following E&M codes occurred on the line item; 99201–99215, 99304–99350, 99381–99387, 99391–99397, G0402, G0438, or G0439. These claims were then subset to those provided by primary care providers (PCPs). A PCP was identified using the provider taxonomy crosswalk provided by DHS and included family medicine, internal medicine, obstetrics & gynecology, pediatrics, ambulatory health care facilities/clinics, midwives, clinical nurse specialist, nurse practitioner, physician assistant, and students in an organized health care training program. A beneficiary was attributed to the PCP billing provider that had the most E&M service claims within the attribution year. If an equal number of such claims were present for different providers, then the beneficiary was attributed to the provider with the most recent E&M date of service.

   For those beneficiaries remaining unattributed, go to step 3.

3. **Specialist E&M Attribution.** Using the E&M line items selected in step 2, use the taxonomy crosswalk to select those provided by specialists in including surgery, mental and behavioral health,, emergency medicine, oncology, neonatal critical care, allergy & immunology, dermatology, and ophthalmology. A beneficiary was attributed to the specialty billing provider that had the most E&M service claims within the attribution year. If an equal number of such claims were present for different providers, then the beneficiary was attributed to the provider with the most recent E&M date of service.

   Beneficiaries who did not receive any of these services from any of these providers during the applicable year remain unattributed to either the IHP or comparison group during the applicable year. A full description of Minnesota's attribution methodology is available from the Minnesota Department of Human Services.[24]

[24] http://www.dhs.state.mn.us/main/idcplg?IdcService=GET_FILE&RevisionSelectionMethod=LatestReleased&Rendition=Primary&allowInterrupt=1&noSaveAs=1&dDocName=dhs16_177106

Calculation of person-level weights

To balance the population characteristics for the claims-based analyses, we estimated propensity scores for all individuals from the comparison group. A propensity score is the probability that an individual is in the IHP group rather than the comparison group.

The objective of propensity score modeling is to create a weighted comparison group with characteristics equivalent to those for the IHP population. To the extent that these characteristics are correlated with care coordination, utilization, and quality outcomes, propensity weighting will help balance variation in outcomes from the baseline period.

There are other methods to apply propensity scores to an analysis. Aside from weighting, one frequently used method is matching, whereby an IHP beneficiary is matched to a comparison group beneficiary who has a similar propensity score. While we considered this method, we decided not to pursue matching for several reasons. First, propensity score weighting has been shown to produce less biased estimates, less modeling error (e.g., mean squared error, type 1 error), and more accurate variance estimation and confidence intervals when modeling dichotomous outcomes; and this analysis includes many dichotomous utilization and quality of care outcomes. Second, matching may exclude many comparison group beneficiaries from the analysis if a good match cannot be found. Weighting has the advantage of preserving sample size.

Person-level characteristics

The initial step in the process was to select person-level characteristics to be used in each propensity score model. ***Table D-2-5*** shows the characteristics we used grouped by whether they control for demographics, enrollment, attribution, or beneficiary health status. Because there is limited information available in claims data, we considered also including county level characteristics to control for geographic characteristics such as physician supply and median income to account for potential differences in access to care or other geographic differences. However, we found that there was little variation in county level characteristics which made it difficult to balance on these variables. To optimize the balance and to avoid extreme weights, we therefore excluded county level covariates from the propensity score model. However, we do control for county level characteristics in the outcome model.

Table D-2-5. Covariates for propensity score logistic regressions

Characteristic	Variable type
Demographic characteristics	
Female	Dichotomous
Age (age and age squared)	Continuous
Disabled	Dichotomous
Non-white	Dichotomous
Disabled	Dichotomous
Enrollment	
Continuously enrolled for the entire calendar year	Dichotomous
Had 9+ months of eligibility in the prior calendar year	Dichotomous
Health status measures	
Chronic Illness and Disability Payment score (count of major comorbidities)	Continuous
Had any ED visits in the prior calendar year (except in the 2010 model)	Dichotomous
Had any inpatient admissions in the prior calendar year (except in the 2010 model)	Dichotomous

ED = emergency department.

Estimation and weighting procedures

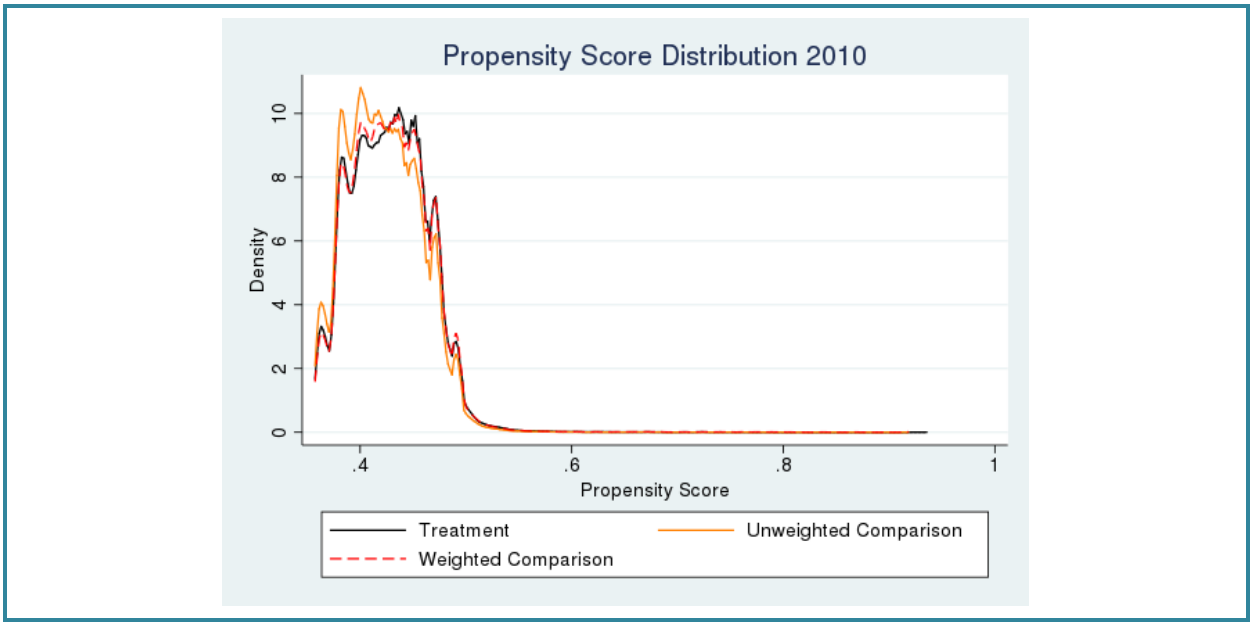
Using the characteristics listed in ***Table D-2-5***, we estimated propensity models by logistic regression, in which the outcome was one for beneficiaries attributed to an IHP and zero for the comparison group. Separate models were estimated for 2010, 2011, 2012, 2013, 2014, 2015, and 2016 data.

We set propensity weights to one for all individuals in the IHP group. The propensity weight for a comparison individual was a function of his or her predicted propensity score—where weight = p/(1-p), with p the predicted propensity. Our procedure includes trimming weights that are either less than 0.05 or greater than 20. In this analysis no weights were trimmed because they were less than 0.05, and in most years and subpopulations no weights were trimmed because they were greater than 20. However, occasionally the propensity score models produced weights that were larger than 20 for some years and subpopulation although no more than 35 observations were ever dropped in a single analysis.

Propensity model evaluation

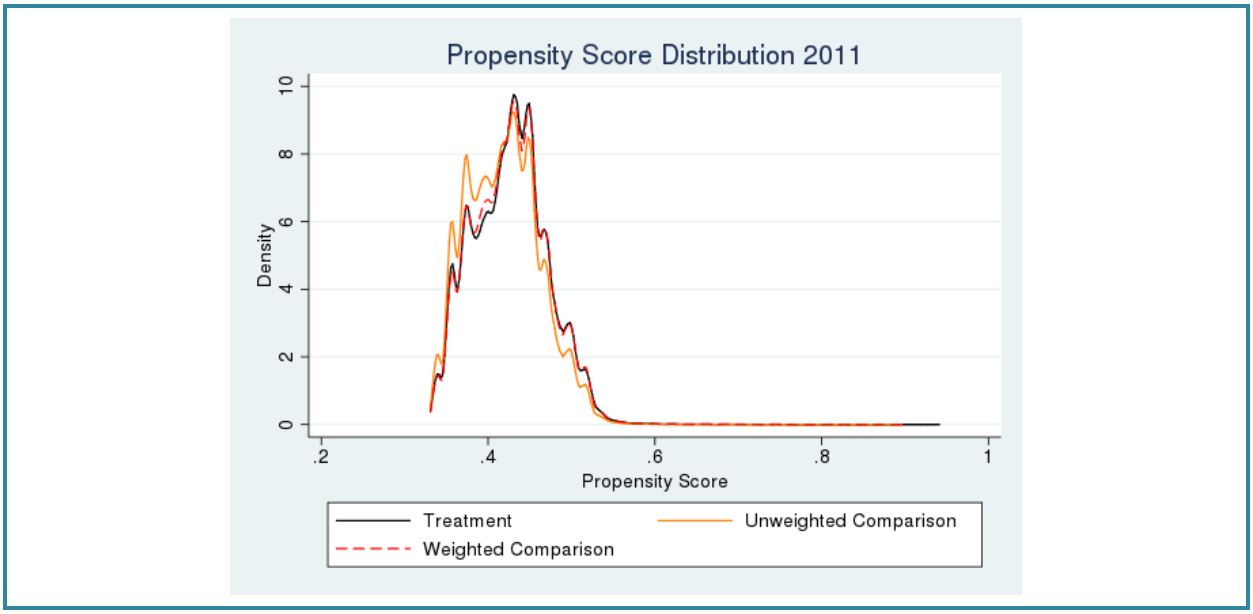
We evaluated several aspects of the propensity score models. First, we examined plots of predicted probabilities to ensure sufficient overlap in the distributions of the IHP and comparison groups. This feature, known as common support, is critical because it provides the basis for inferring effects from group comparisons (***Figures D-2-1*** to ***D-2-7***).

Figure D-2-1. Weighted and unweighted propensity score density plots for the IHP and comparison group, 2010, using DHS Medicaid claims data[25]



DHS = Department of Human Services; IHP = Integrated Health Partnership.

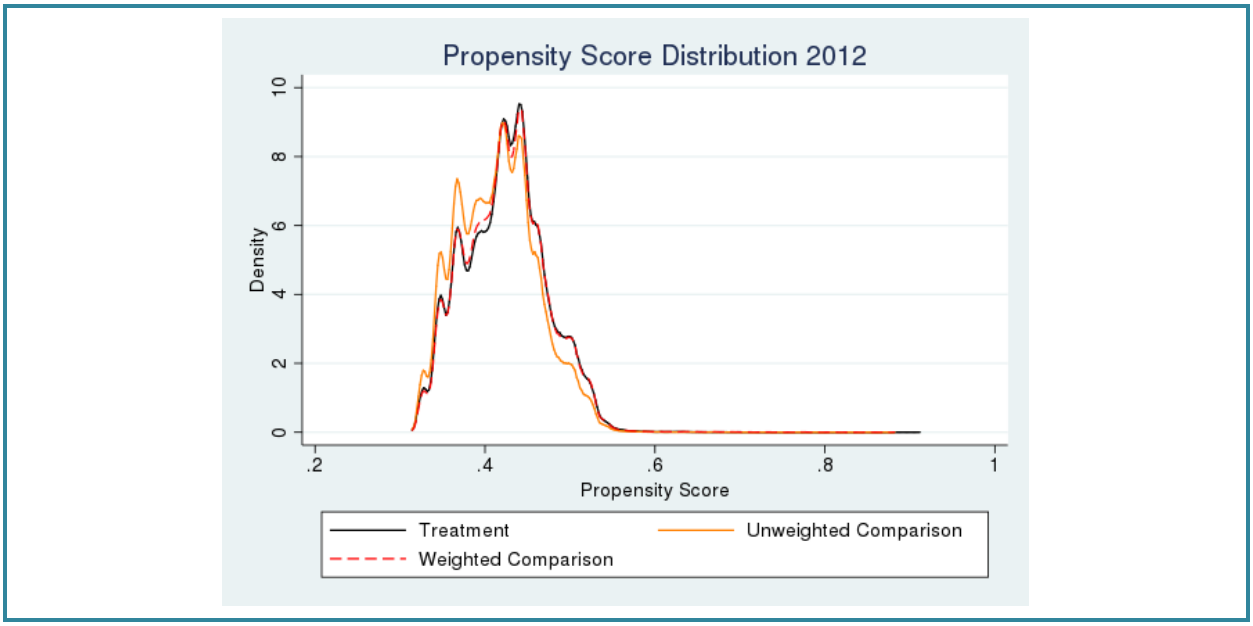
Figure D-2-2. Weighted and unweighted propensity score density plots for the IHP and comparison group, 2011, using DHS Medicaid claims data



DHS = Department of Human Services; IHP = Integrated Health Partnership.

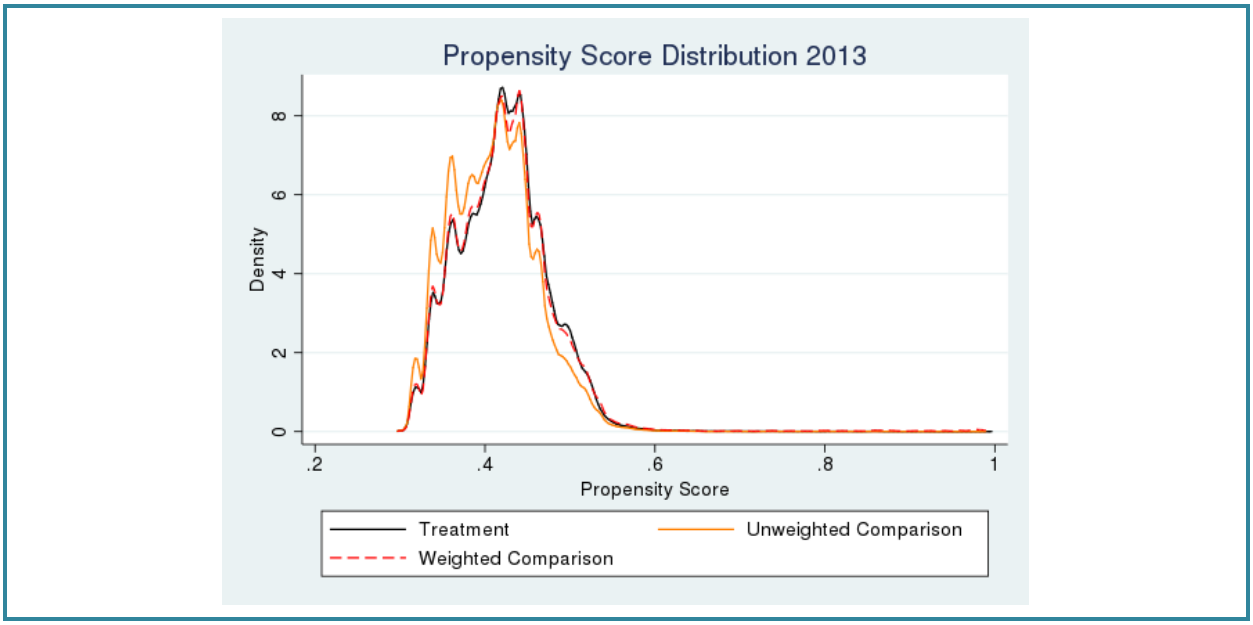
[25] In ***Figures D-2-1*** through ***D-2-5***, the Treatment lines represent those in the IHP group.

Figure D-2-3. Weighted and unweighted propensity score density plots for the IHP and comparison group, 2012, using DHS Medicaid claims data



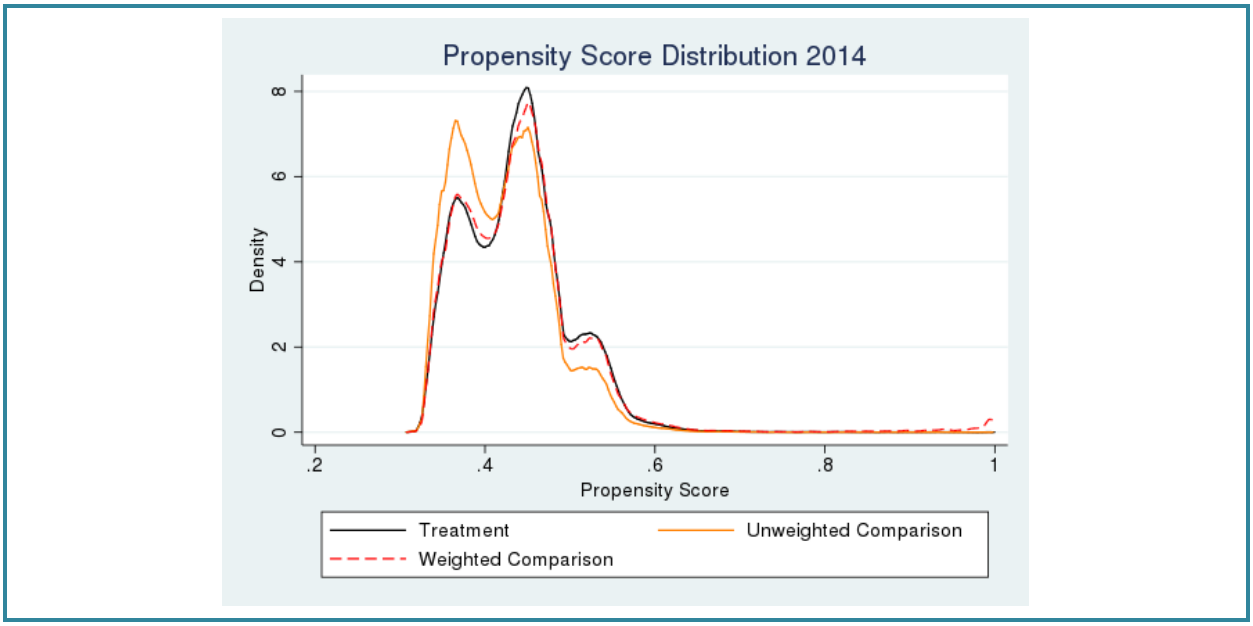
DHS = Department of Human Services; IHP = Integrated Health Partnership.

Figure D-2-4. Weighted and unweighted propensity score density plots for the IHP and comparison group, 2013, using DHS Medicaid claims data



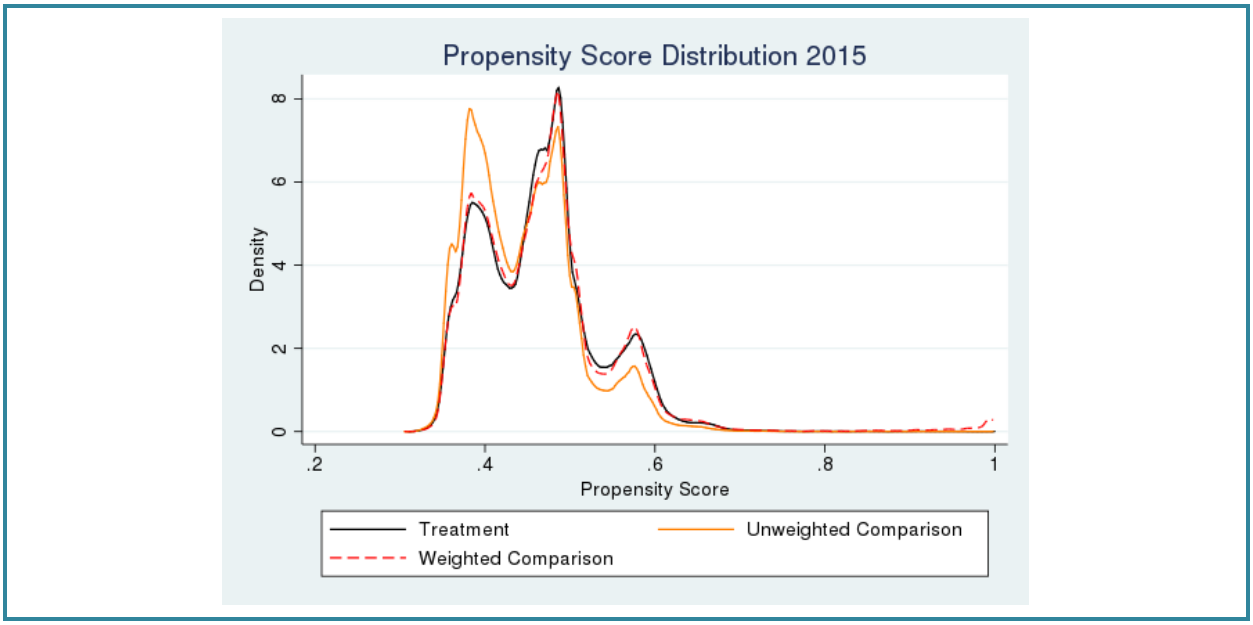
DHS = Department of Human Services; IHP = Integrated Health Partnership.

Figure D-2-5. Weighted and unweighted propensity score density plots for the IHP and comparison group, 2014, using DHS Medicaid claims data



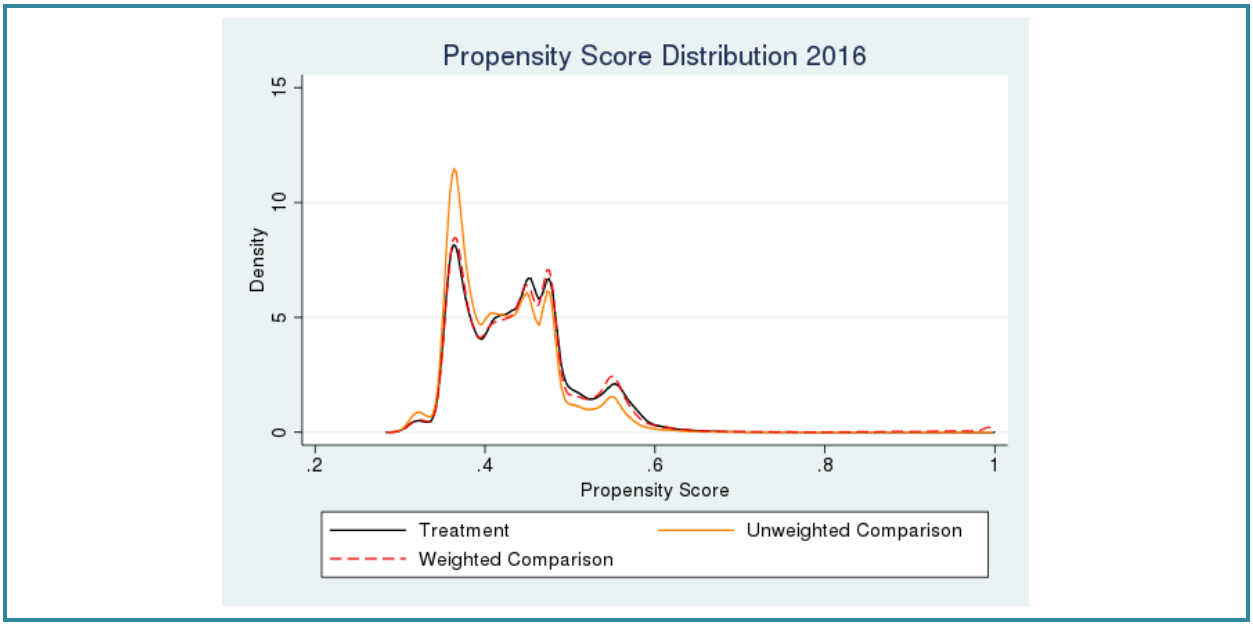
DHS = Department of Human Services; IHP = Integrated Health Partnership.

Figure D-2-6. Weighted and unweighted propensity score density plots for the IHP and comparison group, 2015, using DHS Medicaid claims data



DHS = Department of Human Services; IHP = Integrated Health Partnership.

Figure D-2-7. Weighted and unweighted propensity score density plots for the IHP and comparison group, 2016, using DHS Medicaid claims data



DHS = Department of Human Services; IHP = Integrated Health Partnership.

In all years, we found the comparison group passed the common support assumption (P(D = 1|X)>0) for almost the entire range of the IHP group's propensity scores. The only exceptions were in the uppermost percentiles of the IHP group's distribution [above the 99th percentile]. These plots provide ample evidence that the common support assumption is upheld.

Second, we compared the logistic results of the models to see which variables had the greatest impact on the propensity score weights. Overall, there were few substantial differences between the IHP and comparison groups in terms of beneficiary- or county-level characteristics. Furthermore, most differences were not consistent across measurement years.

Finally, we compared unweighted and propensity-weighted means for the characteristics in the model. As expected, we found that, after weighting, the comparison group means were within a few percentage points of the values for the IHP group.

Tables D-2-6 to ***D-2-12*** show unweighted and propensity score weighted means/proportions for 2010–2016. The notable group differences in the unweighted samples—age, attribution and socioeconomic factors—are substantially mitigated post-weighting as evidenced by the minimized standardized differences. With the exception of several county level characteristics in select years such as the number of hospital beds per 1,000 residents and the percent uninsured, all covariates were well under the commonly accepted threshold of less than 10 percent standardized difference after weighting.

Table D-2-6. Unweighted and weighted means and standardized differences, IHP and comparison groups, 2010

Characteristic	Unweighted			Weighted			
	Comparison group	IHP group	Standardized difference[a]	Comparison group	IHP group	Standardized difference[a]	p-value
N	327,954	242,125		242,191	242,125		
Individual-level sociodemographic characteristics							
Infants (%)	3.5	4.3	4.1	4.4	4.3	0.4	0.17
Age 1–18 years old (%)	48.5	52.7	8.5	52.8	52.7	0.1	0.67
Age 19–64 (%)	48.0	43.0	10.1	42.8	43.0	0.3	0.30
CDPS Risk Score	1.1	1.2	5.0	1.2	1.2	1.2	0.01
Female (%)	56.0	57.1	2.2	57.2	57.1	0.1	0.80
Disabled (%)	7.7	8.3	2.0	8.2	8.3	0.1	0.67
Non-white (%)	46.9	49.5	5.2	49.5	49.5	0.0	0.95
Continuously enrolled in 2010 (%)	92.3	92.4	0.5	92.4	92.4	0.0	0.96
County-level characteristics							
Median age	37.7	37.5	5.9	37.6	100.0	4.7	<0.001
Percent below the poverty line	12.2	12.4	6.1	12.2	12.4	5.3	<0.001
Hospital beds per 1,000 residents	3.1	3.5	14.1	3.1	3.5	14.4	<0.001
Percent without health insurance	9.8	9.7	6.4	9.8	9.7	6.9	<0.001
Rural/urban continuum	71.2	77.0	13.1	71.6	77.0	12.4	<0.001

CDPS = Chronic Illness and Disability Payment System; IHP = Integrated Health Partnerships.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-7. Unweighted and weighted means and standardized differences, IHP and comparison groups, 2011

Characteristic	Unweighted			Weighted			
	Comparison group	IHP group	Standardized difference[a]	Comparison group	IHP group	Standardized difference[a]	p-value
N	368,064	268,599		268,624	268,599		
Individual-level sociodemographic characteristics							
Infants (%)	3.3	4.1	4.0	3.8	4.1	1.5	<0.001
Age 1–18 years old (%)	47.0	51.7	9.6	52.0	51.7	0.5	0.08
Age 19–64 (%)	49.7	44.2	11.1	44.2	44.2	0.1	0.62
CDPS Risk Score	0.0	0.0	0.2	0.0	0.0	0.2	0.48
Female (%)	1.1	1.2	4.9	1.2	1.2	0.9	0.03
Disabled (%)	55.5	56.6	2.4	56.7	56.6	0.0	0.94
Non-white (%)	7.7	8.3	2.2	8.2	8.3	0.2	0.54
Continuously enrolled in 2011 (%)	46.8	49.5	5.4	49.4	49.5	0.1	0.67
Enrolled for at least 9 months in 2010 (%)	92.3	92.5	0.7	92.5	92.5	0.1	0.72
Any ED visits in 2010 (%)	65.5	68.4	6.2	68.4	68.4	0.0	0.90
Any inpatient admissions in 2010 (%)	25.2	29.6	9.9	29.6	29.6	0.0	0.94
County-level characteristics							
Median age	9.2	10.3	3.8	10.3	10.3	0.0	0.94
Percent below the poverty line	37.7	37.4	5.7	37.6	100.0	4.9	<0.001
Hospital beds per 1,000 residents	12.2	12.4	6.1	12.2	12.4	5.1	<0.001
Percent without health insurance	3.1	3.5	14.5	3.1	3.5	14.5	<0.001
Rural/urban continuum	9.8	9.7	6.7	9.8	9.7	7.4	<0.001

CDPS = Chronic Illness and Disability Payment System; ED = emergency department; IHP = Integrated Health Partnerships.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-8. Unweighted and weighted means and standardized differences, IHP and comparison groups, 2012

Characteristic	Unweighted			Weighted			
	Comparison group	IHP group	Standardized difference[a]	Comparison group	IHP group	Standardized difference[a]	p-value
N	409,240	294,923		294,946	294,923		
Individual-level sociodemographic characteristics							
Infants (%)	3.7	4.2	2.6	4.1	4.2	0.4	0.12
Age 1–18 years old (%)	45.6	51.2	11.3	50.9	51.2	0.6	0.02
Age 19–64 (%)	50.7	44.6	12.3	45.0	44.6	0.7	0.003
CDPS Risk Score	0.0	0.0	0.2	0.0	0.0	0.2	0.42
Female (%)	1.2	1.3	4.5	1.3	1.3	0.9	0.01
Disabled (%)	55.0	56.2	2.4	56.2	56.2	0.0	0.94
Non-white (%)	7.8	8.3	1.8	8.2	8.3	0.2	0.49
Continuously enrolled in 2012 (%)	46.6	49.7	6.2	49.6	49.7	0.1	0.67
Enrolled for at least 9 months in 2011 (%)	91.5	91.9	1.4	91.8	91.9	0.1	0.68
Any ED visits in 2011 (%)	66.6	69.8	7.0	69.8	69.8	0.1	0.80
Any inpatient admissions in 2011 (%)	26.7	31.8	11.2	31.8	31.8	0.0	0.94
County-level characteristics							
Median age	9.2	10.3	3.8	10.3	10.3	0.0	0.95
Percent below the poverty line	37.6	37.4	5.7	37.6	100.0	4.7	<0.001
Hospital beds per 1,000 residents	12.2	12.4	6.2	12.2	12.4	4.8	<0.001
Percent without health insurance	3.1	3.5	14.9	3.1	3.5	14.9	<0.001
Rural/urban continuum	9.8	9.7	7.0	9.8	9.7	8.0	<0.001

CDPS = Chronic Illness and Disability Payment System; ED = emergency department; IHP = Integrated Health Partnerships.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-9. Unweighted and weighted means and standardized differences, IHP and comparison groups, 2013

Characteristic	Unweighted			Weighted			
	Comparison group	IHP group	Standardized difference[a]	Comparison group	IHP group	Standardized difference[a]	p-value
N	460,204	326,449		327,267	326,449		
Individual-level sociodemographic characteristics							
Infants (%)	3.4	4.0	3.0	3.8	4.0	0.9	<0.001
Age 1–18 years old (%)	44.5	50.8	12.7	50.4	50.8	0.7	0.001
Age 19–64 (%)	52.1	45.2	13.8	45.7	45.2	1.1	<0.001
CDPS Risk Score	0.0	0.0	0.4	0.0	0.0	0.5	0.04
Female (%)	1.1	1.2	8.1	1.4	1.2	5.4	<0.001
Disabled (%)	54.6	55.8	2.4	55.8	55.8	0.0	0.84
Non-white (%)	7.7	8.2	1.8	8.3	8.2	0.2	0.52
Continuously enrolled in 2013 (%)	46.3	49.8	7.0	49.8	49.8	0.1	0.69
Enrolled for at least 9 months in 2012 (%)	92.2	92.3	0.2	92.3	92.3	0.0	0.87
Any ED visits in 2012 (%)	65.2	68.8	7.6	68.9	68.8	0.3	0.15
Any inpatient admissions in 2012 (%)	26.4	31.6	11.4	31.8	31.6	0.4	0.07
County-level characteristics							
Median age	9.3	10.0	2.2	10.3	10.0	1.0	<0.001
Percent below the poverty line	37.6	37.4	5.6	37.6	100.0	4.5	<0.001
Hospital beds per 1,000 residents	12.2	12.4	6.1	12.2	12.4	4.7	<0.001
Percent without health insurance	3.1	3.5	15.1	3.1	3.5	15.2	<0.001
Rural/urban continuum	9.8	9.7	7.3	9.9	9.7	8.4	<0.001

CDPS = Chronic Illness and Disability Payment System; ED = emergency department; IHP = Integrated Health Partnerships.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-10. Unweighted and weighted means and standardized differences, IHP and comparison groups, 2014

Characteristic	Unweighted			Weighted			
	Comparison group	IHP group	Standardized difference[a]	Comparison group	IHP group	Standardized difference[a]	p-value
N	499,288	373,238		374,827	373,238		
Individual-level sociodemographic characteristics							
Infants (%)	3.1	3.8	4.0	3.7	3.8	0.9	<0.001
Age 1–18 years old (%)	40.6	47.8	14.5	47.3	47.8	0.9	<0.001
Age 19–64 (%)	56.0	48.4	15.4	48.8	48.4	0.8	<0.001
CDPS Risk Score	0.3	0.0	6.3	0.2	0.0	5.7	<0.001
Female (%)	1.0	1.2	9.8	1.4	1.2	7.4	<0.001
Disabled (%)	54.9	55.5	1.1	55.4	55.5	0.1	0.56
Non-white (%)	6.5	6.6	0.3	6.7	6.6	0.6	0.02
Continuously enrolled in 2014 (%)	46.0	49.2	6.5	49.2	49.2	0.1	0.80
Enrolled for at least 9 months in 2013 (%)	98.4	98.0	3.6	98.0	98.0	0.1	0.62
Any ED visits in 2012 (%)	54.7	59.0	8.7	59.3	59.0	0.6	0.01
Any inpatient admissions in 2012 (%)	19.2	25.2	14.5	25.5	25.2	0.7	0.003
County-level characteristics							
Median age	7.7	8.7	3.5	9.1	8.7	1.5	<0.001
Percent below the poverty line	37.6	37.3	6.2	37.5	100.0	5.6	<0.001
Hospital beds per 1,000 residents	12.0	12.3	7.7	12.1	12.3	6.2	<0.001
Percent without health insurance	3.0	3.5	17.3	3.1	3.5	17.3	<0.001
Rural/urban continuum	9.7	9.6	3.9	9.7	9.6	5.2	<0.001

CDPS = Chronic Illness and Disability Payment System; ED = emergency department; IHP = Integrated Health Partnerships.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-11. Unweighted and weighted means and standardized differences, IHP and comparison groups, 2015

Characteristic	Unweighted			Weighted			
	Comparison group	IHP group	Standardized difference[a]	Comparison group	IHP group	Standardized difference[a]	p-value
N	496,433	411,808		413,302	411,808		
Individual-level sociodemographic characteristics							
Infants (%)	3.0	3.5	3.1	3.4	3.5	0.4	0.06
Age 1–18 years old (%)	39.3	47.3	16.2	46.7	47.3	1.3	<0.001
Age 19–64 (%)	57.2	49.1	16.2	49.5	49.1	0.7	0.002
CDPS Risk Score	0.5	0.0	9.3	0.4	0.0	7.9	<0.001
Female (%)	0.9	1.1	9.7	1.3	1.1	7.0	<0.001
Disabled (%)	55.4	55.0	0.7	54.9	55.0	0.2	0.45
Non-white (%)	6.2	5.7	2.2	5.8	5.7	0.7	0.002
Continuously enrolled in 2015 (%)	47.9	49.8	3.8	49.8	49.8	0.1	0.69
Enrolled for at least 9 months in 2014 (%)	97.6	96.6	6.2	96.6	96.6	0.2	0.38
Any ED visits in 2012 (%)	64.4	68.3	8.1	68.4	68.3	0.2	0.41
Any inpatient admissions in 2012 (%)	17.3	24.5	17.6	24.7	24.5	0.6	0.02
County-level characteristics							
Median age	7.6	8.7	4.0	9.0	8.7	1.4	<0.001
Percent below the poverty line	37.6	37.3	7.3	37.6	100.0	7.4	<0.001
Hospital beds per 1,000 residents	12.0	12.2	7.8	12.0	12.2	6.8	<0.001
Percent without health insurance	3.0	3.5	17.1	3.0	3.5	17.0	<0.001
Rural/urban continuum	9.6	9.6	3.5	9.7	9.6	4.7	<0.001

CDPS = Chronic Illness and Disability Payment System; ED = emergency department; IHP = Integrated Health Partnerships.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-12. Unweighted and weighted means and standardized differences, IHP and comparison groups, 2016

Characteristic	Unweighted			Weighted			
	Comparison group	IHP group	Standardized difference[a]	Comparison group	IHP group	Standardized difference[a]	p-value
N	480,549	360,480		361,740	360,480		
Individual-level sociodemographic characteristics							
Infants (%)	1.7	2.1	2.7	2.1	2.1	0.0	0.83
Age 1–18 years old (%)	42.4	49.8	14.9	49.4	49.8	0.6	0.01
Age 19–64 (%)	55.4	48.1	14.6	48.0	48.1	0.2	0.36
CDPS Risk Score	0.5	0.0	9.6	0.4	0.0	8.5	<0.001
Female (%)	0.9	1.1	9.4	1.3	1.1	6.8	<0.001
Disabled (%)	55.0	55.2	0.4	55.2	55.2	0.1	0.56
Non-white (%)	6.7	5.7	4.3	5.8	5.7	0.6	0.02
Continuously enrolled in 2016 (%)	50.2	50.9	1.4	51.0	50.9	0.1	0.82
Enrolled for at least 9 months in 2015 (%)	95.3	94.7	2.4	94.8	94.7	0.1	0.68
Any ED visits in 2012 (%)	75.5	76.7	2.8	76.6	76.7	0.2	0.42
Any inpatient admissions in 2012 (%)	17.9	25.0	17.3	25.2	25.0	0.5	0.06
County-level characteristics							
Median age	8.3	9.5	4.2	9.8	9.5	1.1	<0.001
Percent below the poverty line	37.6	37.2	10.4	37.7	100.0	11.1	<0.001
Hospital beds per 1,000 residents	12.0	12.1	4.1	12.0	12.1	3.7	<0.001
Percent without health insurance	3.1	3.4	14.5	3.1	3.4	14.3	<0.001
Rural/urban continuum	9.7	9.5	7.4	9.7	9.5	8.2	<0.001

CDPS = Chronic Illness and Disability Payment System; ED = emergency department; IHP = Integrated Health Partnerships.

[a] Absolute standardized differences are expressed as percentages.

Propensity model evaluation for subpopulation

In addition to the overall model, we also evaluated common support graphs and standardized differences of the propensity score models for the subpopulation analyses. We found that we maintained balance across the IHP-attributed and comparison group beneficiaries among the subpopulations. Common support overlap looked very similar to plots presented for the overall population.

D-2.1.5 Statistical analysis

Regression model

The underlying assumption in D-in-D models estimating the impact of IHPs is that trends in the test group would be similar to that of the comparison group in the absence of the initiative (i.e., that the two were on "parallel paths" prior to the start of the IHPs).

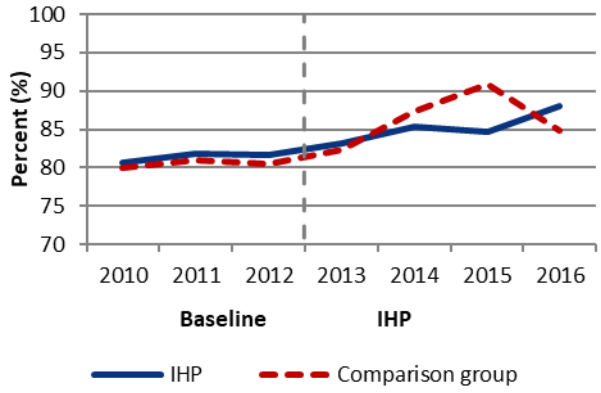
To assess the parallel assumption's validity more empirically, we modelled core expenditure and utilization outcomes during the baseline period with a linear time trend interacted with a dichotomous variable indicating the beneficiary was attributed to an IHP provider (i.e., the "test" group). The following section describes the baseline analysis we conducted to inform the D-in-D model.

To examine descriptively whether the trends in the IHP and comparison group are parallel, we present graphs of annual unadjusted averages for IHP-attributed beneficiaries and the comparison group for the baseline period (2010–2012) and the first 4 years of the IHP implementation (2013–2016).

Figures D-2-8 to ***D-2-13*** provide the unadjusted averages of the care coordination measures by year.

- The baseline trends were parallel for all of the care coordination measures.

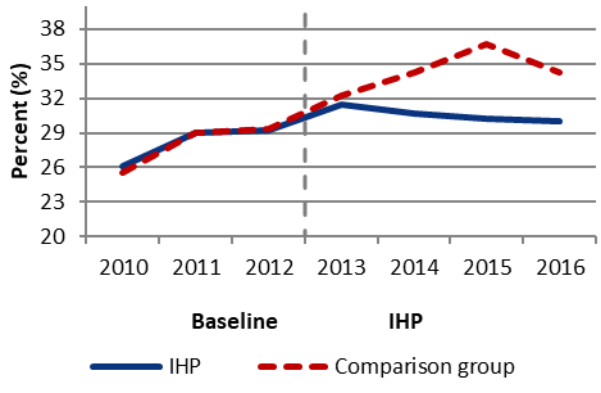
Figure D-2-8. Percentage of Medicaid beneficiaries with a visit to a primary care provider, 2010 through 2016, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



The percentage of IHP-attributed and comparison group beneficiaries with a **primary care provider visit** was similar throughout the baseline period (***Figure D-2-8***). During the IHP implementation period there were some noticeable differences, with comparison group trends increasing more rapidly than the IHP group for most of the IHP implementation period.

IHP = Integrated Health Partnership.

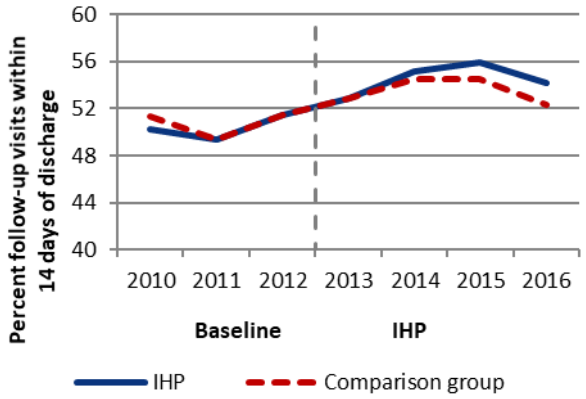
Figure D-2-9. Percentage of Medicaid beneficiaries with a visit to a specialty care provider, 2010 through 2016, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



The percentage of IHP-attributed and comparison group beneficiaries with a **specialty care provider visit** was similar throughout the baseline period (***Figure D-2-9***). During the IHP implementation period, the likelihood of a visit increased more rapidly for comparison group beneficiaries than for IHP-attributed beneficiaries. As with primary care visits, there was a modest convergence in 2016 between the two groups.

IHP = Integrated Health Partnership.

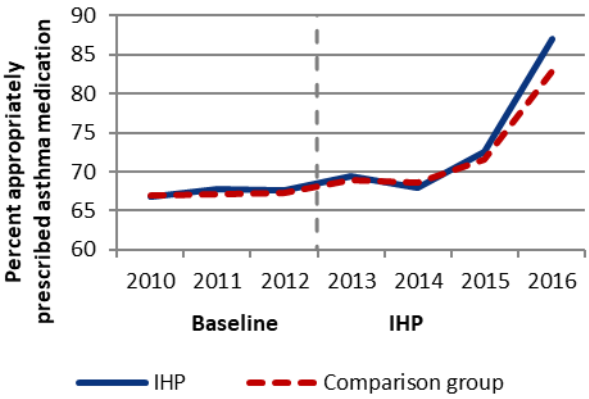
Figure D-2-10. Percentage of Medicaid beneficiaries with a follow-up visit within 14 days of discharge, 2010 through 2016, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



The percentage of hospitalizations with a **follow-up visit within 14 days of discharge** among IHP-attributed and comparison group beneficiaries was similar throughout the baseline and IHP implementation period (***Figure D-2-10***). The percentage of hospitalizations in the IHP group that had a follow-up visit within 14 days of discharge increased more than the comparison group across the latter years of the IHP implementation period.

IHP = Integrated Health Partnership.

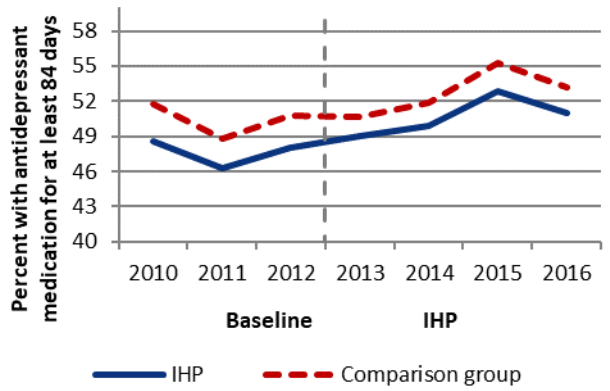
Figure D-2-11. Percentage of Medicaid beneficiaries with persistent asthma who were appropriately prescribed medication during the year, 2010 through 2016, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



The percentage of Medicaid beneficiaries **with persistent asthma who were appropriately prescribed medication during the year** among IHP-attributed and comparison group beneficiaries was similar throughout the entire baseline period and the IHP implementation period. Both groups experienced a remarkable increase in 2016. However, the IHP group experienced a larger increase in 2016 than the comparison group.

IHP = Integrated Health Partnership.

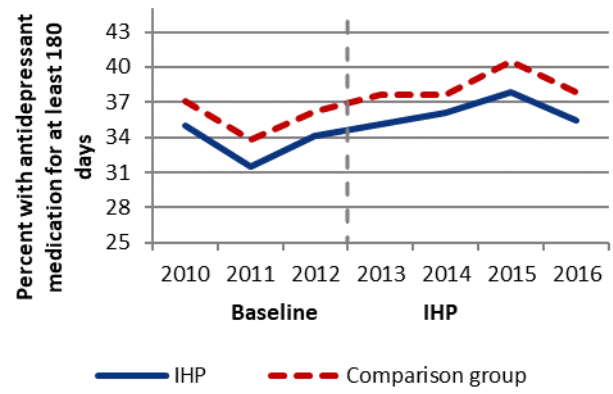
Figure D-2-12. Percentage of Medicaid beneficiaries age 18 years and older diagnosed with a new episode of major depression and treated with antidepressant medication who remained on medication treatment at least 84 days, 2010 through 2016, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



The percentage of Medicaid beneficiaries **age 18 years and older diagnosed with a new episode of major depression and treated with antidepressant medication who remained on medication treatment at least 84 days** among IHP-attributed and comparison group beneficiaries was similar throughout the baseline and IHP implementation periods. There were consistent differences across the two groups, with the IHP group percentages lower than the comparison group percentages.

IHP = Integrated Health Partnership.

Figure D-2-13. Percentage of Medicaid beneficiaries age 18 years and older diagnosed with a new episode of major depression and treated with antidepressant medication who remained on medication treatment at least 180 days, 2010 through 2016, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



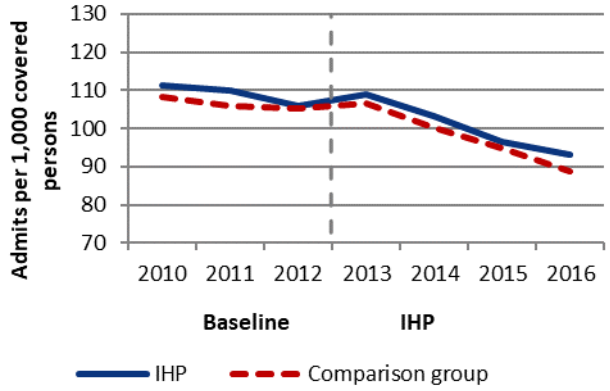
The percentage of Medicaid beneficiaries **age 18 years and older diagnosed with a new episode of major depression and treated with antidepressant medication who remained on medication treatment at least 180 days** among IHP-attributed and comparison group beneficiaries was similar throughout the baseline and IHP implementation periods. There were consistent differences across the two groups, with the IHP group percentages lower than the comparison group percentages.

IHP = Integrated Health Partnership.

Figures D-2-14 to ***D-2-16*** provide unadjusted annual averages of inpatient admissions, outpatient ED visits, and 30-day readmissions per 1,000 Medicaid beneficiaries by year, respectively.

- The baseline trends were parallel for **acute inpatient admission** and **outpatient ED visit rates**, but not parallel for the rate of **30-day readmissions**.

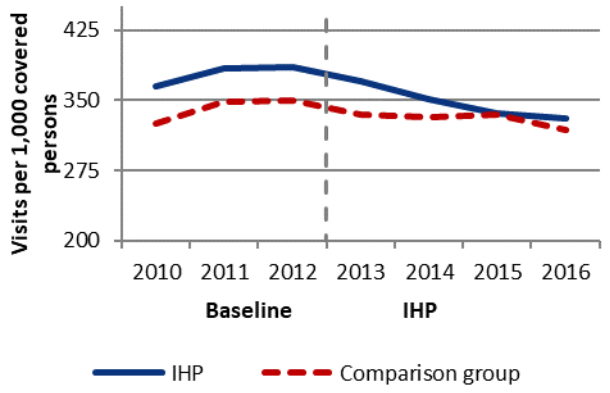
Figure D-2-14. All-cause acute inpatient admissions per 1,000 Medicaid beneficiaries, 2010 through 2016, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



The rate of **acute inpatient admissions** among IHP-attributed Medicaid beneficiaries and comparison group beneficiaries was similar throughout the baseline and IHP implementation periods. The admission rate remained steady during the baseline period and trended slightly downward during the implementation period for both groups (***Figure D-2-14***).

IHP = Integrated Health Partnership.

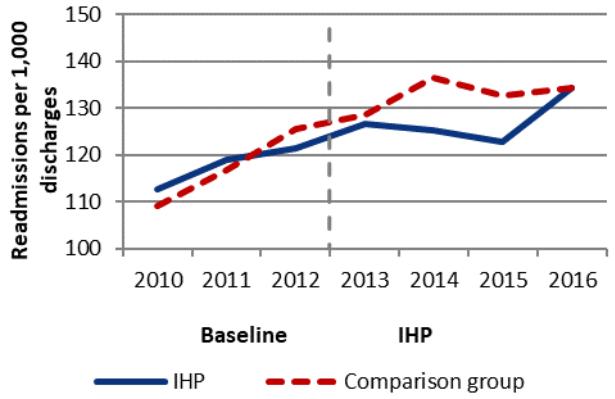
Figure D-2-15. ED visits that did not lead to a hospitalization per 1,000 Medicaid beneficiaries, 2010 through 2016, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



The rate of **outpatient ED visits** among IHP-attributed Medicaid beneficiaries and comparison beneficiaries was different throughout the baseline period and parallel (***Figure D-2-15***). The ED visit rate decreased more rapidly among IHP-attributed beneficiaries than for comparison group beneficiaries during the IHP implementation period.

ED = emergency department; IHP = Integrated Health Partnership.

Figure D-2-16. Discharges with a readmission within 30 days per 1,000 Medicaid beneficiaries, 2010 through 2016, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



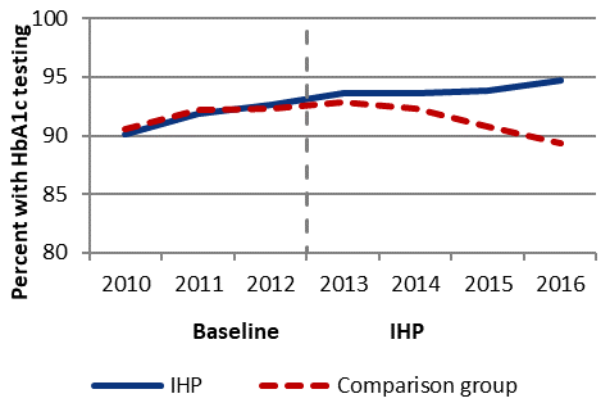
The rate of discharges with a **readmission within 30 days** among IHP-attributed Medicaid beneficiaries increased through the baseline period. The rate among IHP-attributed beneficiaries decreased slightly during the IHP implementation period, while the comparison group rate continued to steadily increase (***Figure D-2-16***).

IHP = Integrated Health Partnership.

Figure D-2-17 provide the unadjusted averages of the quality of care measure by year.

- The baseline trends were parallel for **HbA1c testing rates**.

Figure D-2-17. Percentage of Medicaid beneficiaries age 18–75 years with diabetes (type 1 and type 2) who had Hemoglobin A1c (HbA1c) testing, 2010 through 2016, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



The percentage of **patients age 18–75 years with diabetes who had HbA1c testing** among IHP-attributed Medicaid beneficiaries and comparison group beneficiaries was similar throughout the baseline period. The percentage continued to increase during the IHP implementation period among IHP-attributed beneficiaries, and declined among comparison group beneficiaries (***Figure D-2-17***).

IHP = Integrated Health Partnership.

An annual fixed-effects model considered for the evaluation is shown in ***Equation D-2.1***:

$$\gamma = \alpha_0 + \alpha_1 I + \sum \beta_n Q_{n,b} + \sum \phi_t Q_{t,p} \bullet I + \delta X + \mu \qquad \text{(D-2.1)}$$

where

y = a performance measure (e.g., total PBPM cost per year) for the *i*-th beneficiary in the *j*-th group (test or comparison), in period t (i,j,t subscripts suppressed).

I = a 0,1 indicator (1 = beneficiary is currently attributed to an IHP, 0 = beneficiary is not currently attributed to an IHP or beneficiary is part of the comparison group).

X = a vector of patient and demographic characteristics.

$Q_{n,b}$, $Q_{t,d}$ = 0,1 indicator of the *n*-th or *t*-th calendar year in the base (*b*) or post (*p*) period (*n* starts counting at first baseline period, whereas *t* starts with first IHP year).

μ = error term.

The model in ***Equation D-2.1*** assumes that, except for an intercept difference α_1, the outcomes for beneficiaries in the IHP group and beneficiaries in the comparison groups followed a similar growth trend during the baseline period. We investigated whether the baseline period before the IHP implementation satisfied the baseline trend assumptions of the D-in-D model in ***Equation D-2.1***—that is, whether the outcome trends for beneficiaries in IHP and in the comparison group were similar during this period.

To test the similarity of baseline trends, we used a model with a linear trend during the baseline period. We tested whether this trend differed for IHP-attributed beneficiaries relative to comparison group beneficiaries. Specifically, the model for the outcomes may be written as follows.

$$y = \alpha_0 + \alpha_1 I + \theta \bullet t + \lambda I \bullet t + \delta X + \mu. \qquad \text{(D-2.2)}$$

In ***Equation D-2.2***, *y, I,* X, and μ are defined as in ***Equation D-2.1***. The variable *t* is linear time ranging from 1 to 3. The linear time trend in the comparison group is $\theta \bullet t$, whereas for IHP group beneficiaries (I=1) it is $(\theta + \lambda) * t$. Hence, λ measures the difference in linear trends and the *t*-statistic for this coefficient can be used to test the null hypothesis of equal trends (λ =0). In other words, rejecting the null hypothesis would suggest that the assumption of equal trends underlying our outcome models is not met.

The parameters of ***Equation D-2.2*** were estimated using weighted least-squares regression models for 3 key outcomes. The weights are a function of the eligibility fraction and

propensity scores. For each outcome, we report estimates and standard errors of the difference between the baseline trend in the IHP and the comparison groups (λ).

Tables D-2-13 show estimates of the baseline trend differences for the following outcomes:

- Probability of an acute inpatient stay
- Probability of an outpatient ED visit
- Probability of any primary care practice visit

Table D-2-13. Differences in average utilization outcomes during the baseline period, IHP beneficiaries and comparison group beneficiaries

Parameter estimate	Any inpatient	Any outpatient ED visit	Any primary care practice visit
IHP–CG trend difference	−0.001	−0.002	0.003**
	(0.001)	(0.002)	(0.001)

CG = comparison group; ED = emergency department; IHP = Integrated Health Partnerships.

Baseline is the period January 2010–December 2012. The trend (slope) is the year-to-year change in the outcome variable. Standard errors are given in parentheses. *p < 0.10; **p < 0.05; ***p < 0.01.

Relative to the comparison group, there was no statistically significant difference in the baseline trend for any inpatient admissions or any outpatient ED visits for IHP beneficiaries. There was a statistically significant difference in any primary care practice visits although that difference was relatively minor. Based on the overall results, we concluded that in general beneficiaries in the IHP group were on a similar trajectory with comparison beneficiaries prior to the start of the IHP model, and thus the parallel trend assumption of the D-in-D model was satisfied.

D-in-D regression model—The D-in-D model is shown in ***Equation D-2.3***. The model is an annual fixed effects model as shown in ***Equation D-2.1.*** As in ***Equation D-2.1***, Y_{ijt} is the outcome for individual i (IHP or comparison group) in state j in year t; I_{ij} (=0,1) is an indicator equal to 1 if the individual is in the IHP group and 0 if the individual is in its comparison group; Qn is a series of yearly dummies for the baseline period (years 1 to 3); and Q_t is a series of yearly dummies for the post years (years x to x). The interaction of the IHP group indicator and Q_t ($I_{ij} * Q_t$) measures the difference in the pre-post change between the IHP group and its comparison states.

$$Y_{ijt} = \alpha_0 + \beta_1 I_{ij} + \sum \beta_2 Q_n + \sum \alpha_2 Q_t + \sum \gamma I_{ijt} * Q_t + \lambda X_{ijt} + \varepsilon_{ijt} \qquad \text{(D-2.3)}$$

Table D-2-14 illustrates the interpretation of the D-in-D estimate from this model. The coefficient β_1 in ***Equation D-2.3*** is the difference in the measure between IHP beneficiaries and comparison beneficiaries at the start of the baseline period, holding constant other variables in the equation. The β_2 and α_2 coefficients are for the annual fixed effects and capture differences over time for each baseline and post year, respectively. The coefficient of the interaction term between Q_t and IHP (I) measures any differences for the IHP group relative to the comparison group in the post years relative to baseline years. Thus, in the post-period, the comparison group mean is captured by $\alpha_0 + \alpha_2$, whereas the IHP group mean is captured by $(\alpha_0 + \beta_1) + (\alpha_2 + \gamma)$. In other words, the between-group difference changes from β_1 during the baseline years to $\beta_1 + \gamma$ during the post-period. The D-in-D parameter, γ, shows whether the between-group difference increased ($\gamma>0$) or decreased ($\gamma<0$) after the IHP was implemented. Using the annual fixed effects model, we calculated overall estimates by taking linear combinations of the yearly estimates.

Table D-2-14. Difference-in-differences estimate

Group	Pre-period	Post-period	Pre-post difference
IHP	$\alpha_0 + \beta_1 + \beta_2$	$(\alpha_0 + \beta_1) + (\alpha_2 + \gamma)$	$\alpha_2 + \gamma$
Comparison	$\alpha_0 + \beta_2$	$\alpha_0 + \alpha_2$	α_2
Between group	β_1	$\beta_1 + \gamma$	Γ

IHP = Integrated Health Partnership.

Models for 30-day readmissions and mental health follow-ups were estimated at the annual-admission level. All other outcomes were estimated with the beneficiary year as the unit of analysis.

We converted utilization counts into binary outcomes (1 = any use) and used weighted logistic regression models. Count models are not appropriate because of the low occurrence of most types of utilization for individual beneficiaries in any year; however, we multiplied the marginal effect from the logistic regression models by 1,000 to obtain approximate rates of utilization per 1,000 beneficiaries. Multiplying the marginal effect by 1,000 does not produce an exact rate of utilization per 1,000 beneficiaries as it assumes no person has more than one visit or admission per year. However, we concluded that this is a reasonable approximation because only a small percentage of beneficiaries had counts exceeding 1 for any of the utilization measures. For expenditure outcomes, we used weighted generalized linear models with a normal distribution and identity link.

The models for inpatient admissions, ED visits, and readmissions were run separately for children and adults as well as among beneficiaries with a mental or behavioral health condition.

Control Variables. In all models we controlled for the following variables:

- Beneficiary type (child nondisabled, adult nondisabled or adult/child disabled)
- Age and age squared
- Gender
- Beneficiary's classification on the Chronic Illness and Disability Payment System
- Whether the beneficiary was continuously enrolled during the entire year
- Whether the beneficiary was enrolled for at least 9 months in the previous year
- Median age in the beneficiary's county of residence
- Percent of persons below the poverty line in the beneficiary's county of residence
- Number of hospital beds per 1,000 residents in the beneficiary's county of residence
- Classification of beneficiary's county of residence on the rural/urban continuum

Weighting and Clustering. All of the regression models were estimated using weighted regressions and weighted by the propensity score times the eligibility fraction. In addition, standard errors were clustered at the beneficiary level to account for repeated observations on the same beneficiaries over time.

Adjusted means. The regression-adjusted D-in-D estimate and the D-in-D calculated from regression-adjusted means will differ for one of two reasons. First, in nonlinear specifications the D-in-D calculated from the regression-adjusted means is known to be a biased estimator for the treatment effect. To address this bias, we use the nonlinear D-in-D approach described in Puhani (2012). In some cases the bias may be extreme, leading to substantial differences between the regression-adjusted D-in-D estimates versus the D-in-D calculated from regression-adjusted means.

Second, in linear specifications the D-in-D calculated from the overall regression-adjusted means may be substantially different than the overall regression-adjusted D-in-D estimate because we use different weights to obtain the overall figures. Specifically, the regression-adjusted D-in-D estimates are weighted using the number of treatment beneficiaries observed in each year relative to the total number of treatment beneficiaries ever observed during the test period. This is mathematically equivalent to weighting the test-period adjusted means for both groups with the same weights that are applied to the treatment group. However, the test-period adjusted means that are presented for the comparison group are weighted using the number of comparison beneficiaries observed in each year relative to the total number of comparison beneficiaries ever observed during the test period. The implication of this is that in cases where there are large differences in the rates of rolling entry or exit across the two groups,

we may observe large differences in the D-in-D calculated from the overall regression-adjusted means versus the overall regression-adjusted D-in-D estimate.

D-2.2 Methods for the Minnesota IHP Impact Analysis Using the Minnesota All Payer Claims Database

To estimate the impact of the IHPs in Minnesota, we used the Minnesota All Payer Claims Database (MN APCD) to conduct two D-in-D regression analyses. The first analysis examined the effects of IHPs on Medicaid beneficiaries' medical expenditures. This analysis complements our analyses of utilization, care coordination, and quality of care outcomes conducted with Medicaid claims and enrollment data provided by the Minnesota DHS. The second analysis examined the impact of IHPs on medical expenditures, health care utilization, and care coordination for commercially insured individuals. Although the IHP model was designed to serve Medicaid beneficiaries, this analysis focused on ascertaining whether there were any spillover effects of the IHP model on commercially insured members who also received care from IHP-participating providers. For both the Medicaid and commercial analyses, the baseline period is 2010 through 2012, and the intervention period is 2013 through 2015. This appendix details the methods we used for this analysis.

IHPs in the context of Minnesota Medicaid and commercial insurance. As noted in ***Section D-2.1***, IHPs are ACOs that serve Minnesota's Medicaid population.[26] The IHP concept was created in legislation from 2010, and the first IHPs launched in January 2013. The program has added new IHP organizations and newly attributed Medicaid beneficiaries in every year since 2013. Not all Medicaid beneficiaries are eligible to be attributed to IHPs. Individuals who have no health care home claims or no E&M claims, who have less than 6 months of continuous enrollment or less than 9 months of noncontinuous enrollment, who are dually enrolled in Medicare, or who only have partial Medicaid benefits cannot be attributed to IHPs. However, by 2017, 58 percent of Minnesota's Medicaid population—almost everyone eligible for attribution—was attributed to an IHP.

Because most IHPs also treat commercially insured and other publicly insured individuals, it is possible that provider participation in an IHP would produce a "spillover effect" for the commercial population. In other words, the practice changes produced by IHP affiliation could affect how providers care not only for Medicaid beneficiaries but also for commercially insured patients, who represent 59 percent of Minnesota's population.[27] Minnesota does not attribute commercially insured individuals to IHPs—these individuals are not the IHP program's target population—but it is possible to mimic the IHP attribution process in the commercial population. Applying the state's attribution methodology in MN APCD data, we attributed

[26] Minnesota's Medical Assistance program and its MinnesotaCare program are collectively called "Medicaid."
[27] Minnesota Department of Health, Health Economics Program. Chartbook Section 2: Trends and variation in health insurance coverage. http://www.health.state.mn.us/divs/hpsc/hep/chartbook/section2.pdf

between 262,000 to 572,000 commercially insured individuals to IHP-affiliated providers, with more commercially insured individuals attributed in 2015 than in 2013. The number of commercial plan members we attributed to IHPs exceeds the number of Medicaid beneficiaries attributed to IHPs because commercial coverage is more common than Medicaid coverage among Minnesota residents.[28]

Profile of IHP participating providers. As in the Medicaid claims-based analysis, we combined individuals attributed to either an integrated or a virtual IHP into a single treatment group. As noted in ***Section D-2.1*****,** the number of IHPs has grown since the program began in 2013. Six IHPs launched on January 1, 2013; three launched on January 1, 2014; seven started on January 1, 2015, and three started on January 1, 2016. IHPs serve both urban and rural areas, and some IHPs represent providers treating specific populations, such as children or individuals with disabilities. As in Minnesota Medicaid claims data, the MN APCD provides limited information on providers. Although we cannot directly control for specific practice-level characteristics, we do control for and balance on a number of county- and beneficiary-level characteristics across the IHP and comparison groups.

Study design. The MN APCD Medicaid and commercial analyses use similar designs. Both analyses compare the pre-period (2010–2012) and post-period (2013–2015) trends for the IHP and comparison groups using a longitudinal design with an unbalanced panel. That is, we used all available data for Medicaid or commercial plan members attributed to the IHP and comparison groups in any given year and did not restrict our analysis to individuals who maintain sufficient enrollment from year to year in Medicaid or in commercial coverage. This means that some individuals are included in the sample only during the post-period. The advantages and disadvantages of using an unbalanced panel are discussed in ***Section D-2.1***. In the MN APCD Medicaid analyses, 24 percent of the sample has a full panel of data (i.e., are observed in each year from 2010 through 2015).[29] In the MN APCD commercial analyses, 34 percent of the sample was observed in all study years.

For both the MN APCD Medicaid and commercial analyses, we conducted tests and examined visual output to understand if trends in key outcomes were parallel across the IHP and comparison groups prior to the start of the IHP program. These tests inform the validity of our study design and also indicate whether sample composition changed outcome trends in meaningful ways over time. ***Figures D-2-18*** through ***D-2-23*** show that the Medicaid analysis passes the parallel trends test in the baseline period. ***Figures D-2-24*** through ***D-2-29*** also show

[28] Minnesota Department of Health, Health Economics Program. Chartbook Section 2: Trends and variation in health insurance coverage. http://www.health.state.mn.us/divs/hpsc/hep/chartbook/section2.pdf

[29] This is similar to the finding from the Medicaid claims-based analysis that 21 percent of the sample was continuously enrolled in Medicaid from 2010 through 2016.

that baseline trends for commercial outcomes are similar between the IHP and comparison groups.

Year-specific balance tables (***Tables D-2-17*** through ***D-2-28***) show that average Medicaid beneficiary or commercial plan member characteristics do not differ substantially year to year within the IHP or comparison group, suggesting that even though some beneficiaries may not have baseline data, the characteristics of the sample are stable over time. Lastly, as in the Minnesota Medicaid claims analysis, if a Medicaid beneficiary or a commercial plan member was ever attributed to an IHP, they were excluded from the comparison group, regardless of attribution year. In both the MN APCD Medicaid and commercial analyses, the difference in the changes over time from the pre-period to the post-period between the IHP group and comparison group provides an estimate of the impact of the IHP in its first 3 years of implementation.

Profiles of IHP and comparison groups. The process for creating the IHP and comparison groups was identical across the MN APCD Medicaid and MN APCD commercial analyses. However, in contrast to the Medicaid claims analysis, we were not provided a roster of IHP-attributed Medicaid beneficiaries in the MN APCD. Therefore, we replicated Minnesota's attribution process to identify IHP-attributed Medicaid beneficiaries in MN APCD data. (The attribution process is described in ***Section D-2.2.3*** below.) As noted above, the IHP program was targeted toward Medicaid beneficiaries, so Minnesota did not attribute commercial plan members to IHPs. However, to examine the potential for spillover of the IHP model in the commercial population, we also replicated Minnesota's attribution methodology to assign commercial plan members to IHPs.

Subpopulations. In addition to the analysis on the overall population, we conducted two subpopulation analyses: (1) children and adults separately and (2) beneficiaries diagnosed with mental and behavioral health conditions. As noted in ***Section D-2.1***, we examine the child subpopulation because two IHPs focused on pediatric populations. We also examine the effects of IHPs on the health care for individuals with behavioral health conditions because behavioral health integration was an important component of Minnesota's SIM Initiative.

Balancing IHP and comparison groups. Because Medicaid and commercial plan members were not randomly assigned to IHPs or the comparison group, there may be observed sociodemographic and geographic differences between IHP-attributed and comparison group individuals that may influence results. To address this, we used propensity score weighting in both the MN APCD Medicaid and MN APCD commercial analyses to statistically adjust the Medicaid and commercial study samples to reduce these differences. We calculated separate propensity score models for the MN APCD Medicaid analysis and the MN APCD commercial analysis. To apply propensity score weighting, we first used logistic regression to predict an individual's likelihood of being attributed to an IHP based on select sociodemographic and geographic characteristics. This predicted likelihood is known as the propensity score. We then

took the inverse of the propensity score using the formula (propensity score/(1-propensity score)) to create what is known as the inverse probability of treatment weight. We then applied each comparison group member's inverse probability of treatment weight to our regression models. IHP-attributed Medicaid beneficiaries or commercial plan members receive an inverse probability of treatment weight of one. By applying these weights, comparison group individuals are made to look more like IHP group members. After propensity score weighting, the standardized differences between the weighted comparison group and IHP group means were all well under the standard 10 percent threshold for both the MN APCD Medicaid and commercial groups. More information on propensity score weighting is available in ***Section D-2.2.4.***

Regression-adjusted difference-in-differences outcomes. Analyses for expenditures used ordinary least squares models. Analyses for utilization and care coordination outcomes, which were represented as binary variables, used logistic regression. Following the Medicaid claims-based analysis, all MN APCD analyses used standard errors clustered at the beneficiary level.

The outcome models controlled for demographic, health status, and county-level characteristics. More information on outcomes is available in ***Section D-2.2.2.*** More information on the regression model is available in ***Section D-2.2.5***.

D-2.2.1 Data sources

MN APCD data. To derive eligibility information and claims-based outcomes for our analytic samples of Medicaid beneficiaries and commercial plan members, we used the MN APCD provided by the Minnesota Department of Health (MDH). The MN APCD is periodically updated by MDH; we used MN APCD extract number 20, version 1. Using these data, we were able to examine the 3 years before (2010 through 2012) and 3 years after (2013 through 2015) the launch of IHPs. The MN APCD is a relational database with separate, linkable files for different types of data. The key files we used for our analyses were (1) an enrollment file, containing individual characteristics, monthly enrollment indicators, and coverage information and (2) a medical claims file. These files included information for all Medicaid beneficiaries and all commercial plan members in Minnesota through 2015. Our analytic sample included the subset of Medicaid beneficiaries and commercial plan members who were attributed to IHP and non-IHP providers in years 2013–2015.

Provider file. We received a file from Minnesota's Department of Human Services (DHS) that included the list of provider NPIs associated with each IHP in each year of the intervention period. We used this file to identify the providers to which Medicaid beneficiaries or commercial plan members were attributed were affiliated with an IHP.

Area Health Resource File (AHRF). The AHRF comprises data collected by the Health Resources and Services Administration from more than 50 sources containing more than 6,000

variables related to health care access at the county level. We used 2010 and 2012 information on health professions supply, population characteristics, and economic data as covariates in the outcome model analyses.

D-2.2.2 Outcome measures

For the MN APCD Medicaid analysis, we only look at expenditure outcomes measured at the per beneficiary per month (PBPM) level. The MN APCD commercial analysis includes outcomes for care coordination, utilization, and expenditures. The expenditure outcomes for the MN APCD commercial analysis were equivalently defined as in the MN APCD Medicaid analysis.

Care coordination

To evaluate the impact of the IHP demonstration in Minnesota on care coordination for commercial plan members, we report the following care coordination measures. Each measure was calculated annually for all eligible commercial plan members in the IHP group and comparison group overall and separately for children, adults, and individuals with behavioral health conditions.

- **Percentage of beneficiaries with any visit to a primary care provider:** This is an indicator for whether a commercial plan member had at least one visit to a primary care provider reported in the MN APCD for the year, divided by the number of commercial plan members in the same year. Primary care physicians were identified using their primary taxonomy code, which was obtained from the National Plan and Provider Enumeration System (NPPES) file. A taxonomy code was considered primary where it was denoted in the NPPES file with a Y or an X. When searching for primary care visits, claims were restricted to those with Healthcare Common Procedure Coding System (HCPCS)/Current Procedural Terminology (CPT) codes indicating evaluation and management (E&M) visits associated with planned physician care (i.e., office visits). We used the same algorithm as in the Medicaid claims analysis to assign providers as primary care or specialty providers based on taxonomy code. ***Table D-2-1*** provides a cross-walk between provider taxonomy codes and the categorization of either primary care physician or specialty provider designation.
- **Percentage of beneficiaries with any visit to a specialty provider:** This is an indicator of whether a commercial plan member had at least one visit to a specialty provider reported in the MN APCD data for the year, divided by the number of commercial plan members in the same year. Specialty care physicians were identified using their primary taxonomy code, which was obtained from the NPPES file. A taxonomy code was considered primary where it was denoted in the NPPES file with a Y or an X. When searching for specialty care visits, claims were restricted to those with HCPCS/CPT codes indicating E&M visits associated with planned physician care (i.e., office visits).
- **Percentage of acute inpatient hospital admissions with a follow-up visit within 14 days:** This measure calculates the percentage of commercial plan members who had an acute care hospitalization and who had a qualifying evaluation and management

outpatient visit within 14 days of discharge. For an acute care hospitalization to be included in the denominator, the commercial plan member was required to (1) be eligible at both the time of admission and 14 days post-discharge, (2) be alive both at discharge and 14 days post-discharge, and (3) not have a readmission within 14 days post-discharge. We used the following CPT codes to identify a follow-up visit:

- 99201 through 99205; New Patient, Office/Other Outpatient Services
- 99211 through 99215; Established Patient, Office/Other Outpatient Services
- 99241 through 99245; Evaluation and Management Consultation Services
- 99341 through 99350; Home-based ambulatory care visits
- 99381 through 99387; New Patient, Preventive Medicine Services
- 99391 through 99397; Established patient, Preventive Medicine Services
- 99401 through 99412; New or established patient; Counseling Risk Factor Reduction and Behavior Change Intervention
- 99420 through 99429; Other preventive medicine services, Counseling Risk Factor Reduction and Behavior Change Intervention
- G0402, G0438, G0439, G0463; Preventive Physical Examination and Wellness Visits

Utilization

Utilization measures are reported as rates per 1,000 covered lives (or discharges for readmissions) for commercial plan members. For each measure, we first calculate the probability of any use. To calculate the probability, the numerator was an indicator of having had at least one event (an inpatient admission or an ED visit that did not lead to a hospitalization) and the denominator is the number of eligible commercial plan members (or discharges) in the state enrolled during the period. We multiplied the probability of use by 1,000 to obtain approximate rates of utilization per 1,000 commercial plan members. Multiplying the probability by 1,000 does not produce an exact rate of utilization per 1,000 commercial plan members because it assumes no person has more than one visit or admission per year. However, we concluded that this is a reasonable approximation because only a small percentage of commercial plan members had counts exceeding one for any of the utilization measures. Events are included in a period's total if the discharge or the service date on the claim was during the period.

- **Probability of having any inpatient use:** This is an indicator of whether a commercial plan member had at least one admission to an acute-care hospital reported in the inpatient file for the year, divided by the number of commercial plan members in the same year. For Minnesota, we identified all-cause acute care hospital admissions using the type of bill for inpatient services (11 or 12). Some records in the inpatient claims files may appear to be multiple admissions but are in fact transfers between different facilities; these records were counted as a single admission. To combine transfers into one acute admission, we identified claims that had no more than 1 elapsed day between the discharge date of the index claim and the admission

date of the subsequent claim. We combined the claims into one record by taking the earliest admission date and latest discharge date and summing all payment amounts. This approach was also taken for continuing care claims when these criteria were met and the facilities were the same.

- **Probability of having any ED visits that did not lead to a hospitalization (outpatient ED) use:** This is an indicator of whether the commercial plan member had at least one visit to the ED that did not result in an inpatient hospital admission in a year, divided by the number of commercial plan members in the same year. ED visits (including observation stays) are identified in MN APCD commercial claims as visits with a revenue center line item equal to 045X, 0981 (ED care) or 0762 (treatment or observation room, thus including observation stays in the overall count). If the procedure code on every line item of the ED claim equaled 70000–89999 and no line items had a revenue center code equal to 0762, that claim was excluded (thus excluding claims for which only radiology or pathology/laboratory services were provided unless they were observation stays). Multiple ED visits on a single day were counted as a single visit. If there was an inpatient ED visit on the same day, the outpatient encounter was excluded.
- **Probability of having a readmission within 30 days of hospital discharge:** This is a dichotomous variable indicating that a commercial plan member had at least one hospital readmission within 30 days of a live discharge. The denominator includes all acute care hospital discharges identified using the criteria described above. Additionally, we excluded discharges if an individual died during the hospitalization or was not enrolled in commercial coverage for the full 30 days post-discharge. The numerator includes readmissions to any acute care hospital within 30 days of discharge.

Expenditures

For the MN APCD Medicaid analysis, we calculated total medical PBPM spending, facility PBPM spending, professional PBPM spending, and behavioral health spending. For the MN APCD commercial analysis, we present only total medical (per member per month) spending. We used the same method to calculate spending for both the MN APCD Medicaid and commercial analyses.

- **Total medical PBPM expenditures.** To calculate spending in Medicaid data from the MN APCD, we first dropped claims for which quantity and charge were negative. We also dropped claims for months in which an individual was not enrolled in relevant coverage (either Medicaid or commercial). We assigned individual claims to years according to the year of the last date of service on the claim. We summed up payments by medical claim line to calculate total claim-level using rules that MDH developed. To produce total spending, we summed claim-level payments at the individual-year level. We then annualized spending for individuals not enrolled in Medicaid or commercial coverage for the entire year by dividing spending by an eligibility fraction. This eligibility fraction is the number of months an individual was enrolled in coverage divided by 12 months in a year. To produce per individual per month payments, we divided spending by 12. Payments were not risk adjusted or

price standardized across geographic areas. Claims were included in a period's total if discharge or service date on the claim was during the period. We were unable to include pharmacy expenditures in total expenditure calculations because of limitations in the pharmacy data.[30]

- **Facility PBPM expenditures**. We calculated facility PBPM spending in the same way that we calculated total medical PBPM expenditures. We categorized claims with a valid bill type code, which describes the type of facility providing health care services, as facility claims.[31]
- **Professional PBPM expenditures**. We calculated facility PBPM spending in the same way that we calculated total medical PBPM expenditures. We categorized claims with a valid place of service code, which describes the setting in which a health care professional provided health care services, as professional claims.[32]
- **Behavioral health PBPM expenditures.** We categorized behavioral health spending as payments from all medical claims for which the primary diagnosis code was related to a mental disorder, as defined by the International Classification of Diseases, versions 9 and 10. Specifically, these codes were:
  - ICD-9: 290xx through 319xx *(x can be any value or missing)*
  - ICD-10: F01xxx through F99xxx; G44209, H9325, R37, R451, R457, R480, Z87890 *(x can be any value or missing)*

D-2.2.3 Populations studied

Medicaid study population. The study population for the MN APCD Medicaid analysis includes Medicaid beneficiaries that we attributed to IHPs at any point between 2013 and 2015. We identified the IHP group by replicating Minnesota's attribution method in the Minnesota Medicaid population. Additional details of the methodology are presented in ***Section D-2.1***. In summary, 12 months of claims were analyzed for all eligible beneficiaries and they were attributed to IHPs if (1) the plurality of their primary care services (from either a primary or specialty care physician visits) were with an IHP-affiliated provider, or (2) they had health home visits with an IHP-affiliated provider. We identified IHP-affiliated providers using a list of such providers provided to us by the state. Eligibility was determined based on having continuous Medicaid enrollment for at least 6 months, or a total of 9 or more months of discontinuous enrollment and not being dually eligible for Medicare.

---

[30] MDH has worked to de-duplicate pharmacy claims in the MN APCD pharmacy claims. However, as of the time that we were planning our analysis, MDH had implemented its de-duplication algorithm only in pharmacy claims from 2013 onward. Because our study includes years prior to 2013, we decided against calculating pharmacy expenditures.

[31] More specifically, we categorized claims as facility claims if they had a valid bill type code and a missing place of service code in consultation with MDH. A marginal number of claims had both a valid bill type and a valid place of service code. We also classified those claims as facility claims.

[32] We categorized claims as professional claims if they had a missing bill type and a valid place of service code.

In addition to analyses that include all IHP or comparison Medicaid beneficiaries or commercial plan members, we conducted subpopulation-specific analyses of children, adults, and beneficiaries with behavioral health conditions. Children were defined as persons aged 18 or younger, and adults were defined as persons aged 19–64 years old. Individuals were identified as having mental or behavioral health conditions prior to their attribution to the IHP or comparison group. To be included in the subpopulation with mental or behavioral health conditions, a person had to have two or more outpatient claims with a primary mental health or chemical dependency diagnosis or at least one inpatient claim with a primary diagnosis of mental health or chemical dependence.

Prior to 2014, Minnesota provided Medicaid coverage to childless adults up to 75 percent of the federal poverty line (FPL). In January 2014, Minnesota expanded Medicaid eligibility through the Patient Protection and Affordable Care Act (ACA) to childless nondisabled adults whose household income does not exceed 138 percent of the FPL. In the MN APCD, we cannot identify which Medicaid beneficiaries were newly eligible for coverage through the ACA's Medicaid expansion. Although we cannot identify which individuals are newly enrolled in Medicaid because of the ACA, we have no reason to believe that the 2014 Medicaid expansion affected the IHP and comparison groups differently. We cannot identify which beneficiaries in our data are newly eligible as a result of this expansion, but we have no reason to believe the expansion affects the IHP and comparison groups differently. Because expansion of coverage to a new set of adults in the 75–138 percent FPL range occurred early in the post-period, it is possible that changes in outcomes in the post-period when compared to the baseline may have been because of new beneficiaries in the sample and possibly bias outcomes to show greater use of primary care because of preexisting demand for services. However, among all childless adults covered by Medicaid in 2013, fewer than 40 percent fall into the 75–138 percent FPL range,[33] and potentially even fewer will meet the basic eligibility requirements to be enrolled in an IHP (e.g., 6 months of continuous enrollment or 9 months of noncontinuous enrollment) and therefore will not be eligible for our sample in either the IHP or comparison groups. Additionally, adults newly eligible in 2014 would still need 6 months of continuous or 9 months of noncontinuous enrollment to be attributed to the IHP or comparison group—restricting the proportion of the sample affected by the change in eligibility policy in 2014 to an even smaller group (i.e., those who enrolled in the first 6 months of 2014). Our model uses an unbalanced panel longitudinal design to maximize utilization of available data; that is, all available data for beneficiaries attributed to the IHP and comparison group in any given year were used for analysis. More information on study design is presented later in this section, and more information on sample construction is available in ***Section D-2.2.4***.

[33] Estimates based on issue brief available here: http://www.mnbudgetproject.org/research-analysis/economic-security/health-care/covering-more-minnesotans-through-medicaid-would-improve-health-outcomes-and-reduce-state-costs

To check whether the IHP-attributed Medicaid beneficiaries in the MN APCD data resembled the IHP-attributed Medicaid beneficiaries in the Medicaid claims-based analysis, we compared sample characteristics for the sample observed in the last baseline year (2012). As ***Table D-2-15*** shows, the two IHP-attributed samples have similar characteristics:

Table D-2-15. Characteristics of IHP-attributed beneficiaries in the MN APCD and Medicaid DHS claims-based analyses, 2012

Characteristic	Minnesota Medicaid claims	MN APCD
N	294,923	239,245
Beneficiary-level sociodemographic characteristics		
Female (%)	56.2	56.3
Age < 1 year (%)	4.2	3.9
Age 1–18 years (%)	51.2	50.6
Age 19–64 years (%)	44.6	45.5
Age ≥ 65 years (%)	0.01	0.0
Characteristics of beneficiary county of residence		
Metropolitan status (%)	77.4	77.1
Uninsured rate (%)	9.7	9.7
Median age mean	37.4	37.5
Poverty rate (%)	12.4	12.7
Hospital beds per 1,000 population	3.5	3.5
Health care utilization/expenditures for beneficiaries		
Any inpatient admissions in the prior year (%)	10.3	8.8
Any ED visits in the prior year (%)	31.8	29.5

DHS = Department of Human Services; ED = emergency department; IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

Commercial study population. The study population for the MN APCD commercial analysis includes commercial plan members that we have attributed to IHPs at any point between 2013 and 2015. As in the MN APCD Medicaid analysis, we replicated the state's attribution method in identifying both the IHP and the comparison groups. Following the MN APCD Medicaid analysis, we also carried subpopulation-specific analyses of children, adults, and individuals with behavioral health conditions. We defined these subgroups for the MN APCD commercial analysis in the same way that we defined subgroups for the MN APCD Medicaid analysis.

D-2.2.4 Comparison group and propensity score weighting

For both the MN APCD Medicaid and MN APCD commercial impact analyses, we used a pre-post comparison group design, in which the comparison group provides an estimate of what would have happened in the IHP group absent the effect on the intervention. The difference

in the changes over time between the IHP group and its comparison group provides an estimate of the impact of the IHP. Ideally, the comparison group should be similar to the IHP group on all relevant dimensions (e.g., demographic, socioeconomic, political, regulatory, health, and health systems) except for the policy change being tested.

In the following section, we detail the procedures we used to select the comparison group for the IHP demonstration in Minnesota.

Selection of comparison group

We created two comparison groups, one for the MN APCD Medicaid analysis and one for the MN APCD commercial analysis. We used the same approach to selecting the comparison group in both of these analyses. As noted earlier, we received a list of all IHP-affiliated providers for each program year and used this list to attribute Medicaid beneficiaries or commercial plan members to IHPs. Attribution was determined by all professional medical claims (in which the place of service field is populated with a valid code). We then replicated Minnesota's IHP attribution methodology among potential comparison group members for each demonstration year (2013 through 2015) separately. Anyone who was ever attributed to the IHP group was excluded from the comparison group.

There were three steps in attribution:

1. **Health Home Claim Code Attribution.** If procedure code S0280 or S0281 occurred on a line item, we attributed that individual to the billing provider that had the most occurrences of those S codes for each year.

   We included Medicaid beneficiaries or commercial plan members who remained unattributed after this step in the second step in the attribution process, attribution to a primary care provider.

2. **Primary Care E&M Attribution.** Line items from the professional claims were selected if the following E&M codes occurred on the line item; 99201–99215, 99304–99350, 99381–99387, 99391–99397, G0402, G0438, or G0439. These claims were then subset to those provided by primary care providers (PCPs). A PCP was identified using the provider taxonomy crosswalk provided by DHS and included family medicine, internal medicine, obstetrics & gynecology, pediatrics, ambulatory health care facilities/clinics, midwives, clinical nurse specialist, nurse practitioner, physician assistant, and students in an organized health care training program. A Medicaid beneficiary or commercial plan member was attributed to the PCP billing provider that had the most E&M service claims within the attribution year. If an equal number of such claims were present for different providers, then the Medicaid beneficiary or commercial plan member was attributed to the provider with the most recent E&M date of service.

   We included Medicaid beneficiaries or commercial plan members who remained unattributed after this step in the third step in the attribution process, attribution to a specialty care provider.

3. **Specialist E&M Attribution.** Using the E&M line items selected in step 2, use the taxonomy crosswalk to select those provided by specialists in including surgery, mental and behavioral health, emergency medicine, oncology, neonatal critical care, allergy & immunology, dermatology, and ophthalmology. A Medicaid beneficiary or commercial plan member was attributed to the specialty billing provider that had the most E&M service claims within the attribution year. If an equal number of such claims were present for different providers, then the Medicaid beneficiary or commercial plan member was attributed to the provider with the most recent E&M date of service.

   Individuals who did not receive any of these services from any of these providers during the applicable year remain unattributed to either the IHP or comparison group during the applicable year. A full description of Minnesota's attribution methodology is available from the Minnesota Department of Human Services.[34]

Calculation of person-level weights

To balance the population characteristics for both the MN APCD Medicaid analysis and the MN APCD commercial analysis, we estimated propensity scores for all individuals from the Medicaid and commercial comparison groups. A propensity score is the probability that an individual is in the intervention group rather than the comparison group.

The objective of propensity score modeling is to create a weighted comparison group with characteristics equivalent to those for the IHP population. To the extent that these characteristics are correlated with outcomes, propensity weighting will help balance pre-intervention levels of the outcomes as well.

There are other methods to apply propensity scores to an analysis. Aside from weighting, one frequently used method is matching, whereby an intervention beneficiary is matched to a comparison group beneficiary who has a similar propensity score. Although we considered this method, we decided not to pursue matching for several reasons. First, propensity score weighting has been shown to produce less biased estimates, less modeling error (e.g., mean squared error, type 1 error), and more accurate variance estimation and confidence intervals when modeling dichotomous outcomes; and the MN APCD analyses include many dichotomous utilization and care coordination outcomes. Second, matching may exclude many comparison group members from the analysis if a good match cannot be found. Weighting has the advantage of preserving sample size.

Person-level characteristics

The initial step in the process was to select person-level characteristics to be used in the propensity score model for the MN APCD Medicaid and commercial analyses. We used slightly

---

34 http://www.dhs.state.mn.us/main/idcplg?IdcService=GET_FILE&RevisionSelectionMethod=LatestReleased&Rendition=Primary&allowInterrupt=1&noSaveAs=1&dDocName=dhs16_177106

different person-level characteristics in the MN APCD Medicaid and MN APCD commercial analyses. Specifically, the MN APCD commercial propensity score models include two commercial insurance-specific variables: an indicator for whether an individual had one or months of pharmaceutical coverage and a categorical variable that describes an individual's relationship to the primary commercial policy holder. ***Table D-2-16*** shows the characteristics included in these analyses, respectively, grouped by whether they control for demographics, enrollment, or health status. We considered also including county-level characteristics to control for geographic characteristics, such as physician supply and median income to account for potential differences in access to care or other geographic differences. However, we found that there was little variation in county-level characteristics, which made it difficult to balance on these variables. To optimize the balance and to avoid extreme weights, we therefore excluded county-level covariates from the propensity score model. However, we do control for county-level characteristics in the outcome model.

The variables included in the MN APCD Medicaid propensity score models are slightly different than those in the Medicaid claims-based analysis described in ***Section D-2.1***. The MN APCD, unlike Minnesota Medicaid claims data, does not include variables that describe race or disability status. The MN APCD Medicaid analysis measures person-level health risk with the Hierarchical Condition Category (HCC) score, rather than the Chronic Illness and Disability Payment System (CDPS) score used in the Medicaid claims-based analysis. The MN APCD analysis uses HCC scores rather than CDPS scores because disability status, incorporated into the CDPS algorithm, is not captured in MN APCD data.

Estimation and weighting procedures

Using the characteristics listed in ***Table D-2-16***, we used logistic regression models to estimate propensity models for the MN APCD Medicaid and commercial analyses. The outcome for these logistic regression models was equal to 1 for individuals attributed to an IHP-affiliated provider and equal to zero for individuals attributed to a non-IHP provider. Separate models were estimated for 2010, 2011, 2012, 2013, 2014, and 2015 data.

We set propensity weights to 1 for all individuals in the intervention group. The propensity weight for a comparison individual was a function of his or her predicted propensity score—where weight = p/(1-p), with p the predicted propensity. Our procedure includes trimming weights that are either less than 0.05 or greater than 20. In both the MN APCD Medicaid and MN APCD commercial analyses, there were no weights that had values less than 0.05. In the MN APCD Medicaid analysis, we trimmed four weights with values greater than 20. In the MN APCD commercial analysis, we trimmed three weights with values greater than 20.

Table D-2-16. Covariates for propensity score logistic regressions for the MN APCD Medicaid analysis

Characteristic	Variable type	MN APCD Medicaid analysis	MN APCD commercial analysis
Demographic characteristics			
Female	Dichotomous	X	X
Age (age and age squared)	Continuous	X	X
Enrollment			
Had 9+ months of eligibility in the prior calendar year [a]	Dichotomous	X	X
Continuously enrolled in the entire calendar year	Dichotomous	X	X
Has at least one month of pharmacy coverage	Dichotomous		X
Relationship to primary insured (self, child, spouse, other relationship to primary insured)	Categorical		X
Health status measures			
Total medical spending PBPM in the prior calendar year (except in the 2010 model)	Continuous	X	X
Had any ED visits in the prior calendar year (except in the 2010 model)	Dichotomous	X	X
Had any inpatient admissions in the prior calendar year (except in the 2010 model)	Dichotomous	X	X
Health status measures			
Hierarchical Condition Category (HCC) score	Continuous	X	X[b]

ED = emergency department; MN APCD = Minnesota All Payer Claims Database; PBPM = per beneficiary per month.

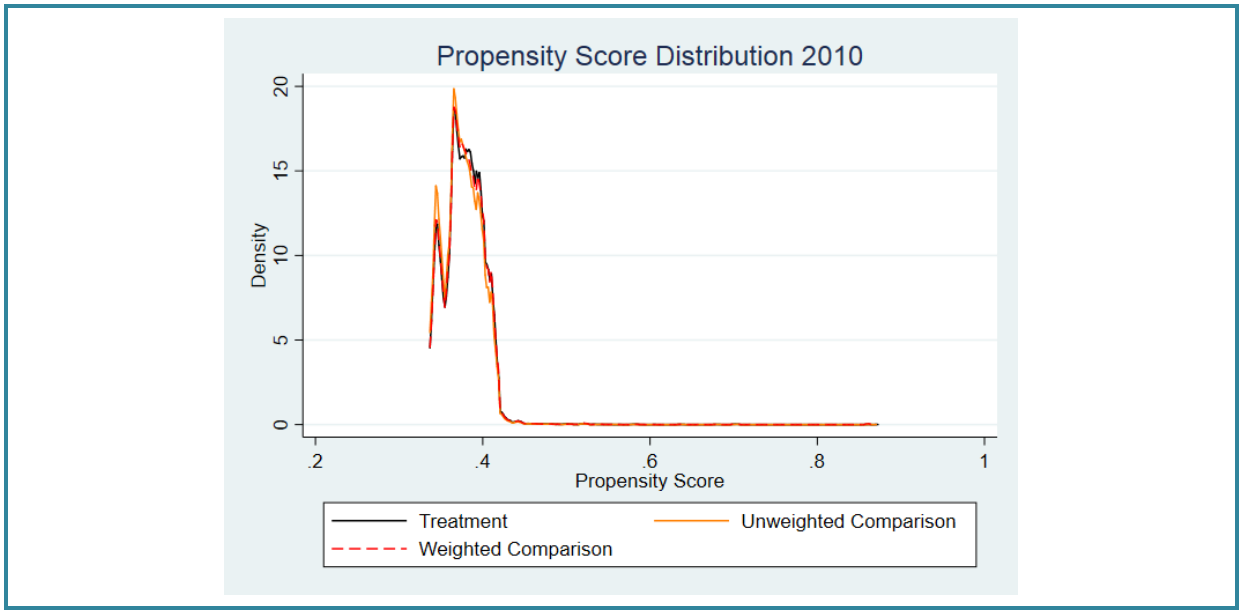
[a] Enrollment in previous year is counted if member was enrolled for 9 or more months.

[b] We use the logged value of the HCC score for the commercial analysis. The logit models for the binary outcomes for the MN APCD commercial analysis fail to converge when the HCC score is not logged.

Propensity model evaluation

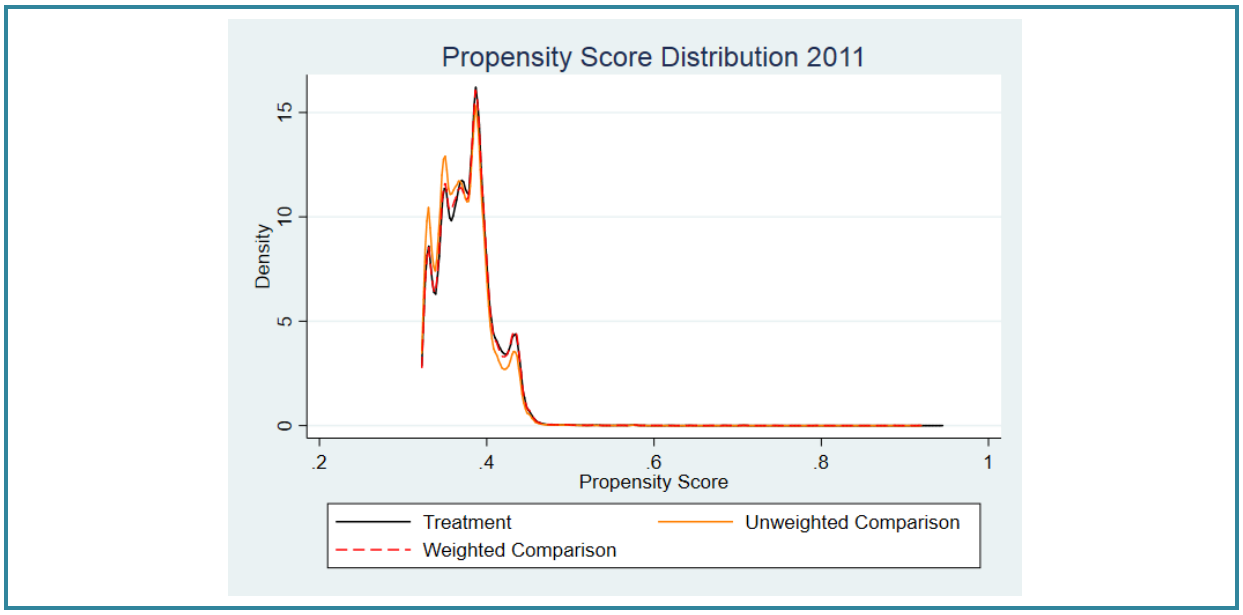
We evaluated several aspects of the propensity score models. First, we examined plots of predicted probabilities to ensure sufficient overlap in the distributions of the MN APCD Medicaid and commercial IHP and comparison groups. This feature, known as common support, is critical because it provides the basis for inferring effects from group comparisons. ***Figure D-2-18*** through ***Figure D-2-23*** present the common support plots for the MN APCD Medicaid analysis. These figures show a high level of overlap between MN APCD Medicaid IHP and comparison groups.

Figure D-2-18. Weighted and unweighted propensity score density plots for the Medicaid IHP and comparison groups, 2010, using MN APCD data[35]



IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

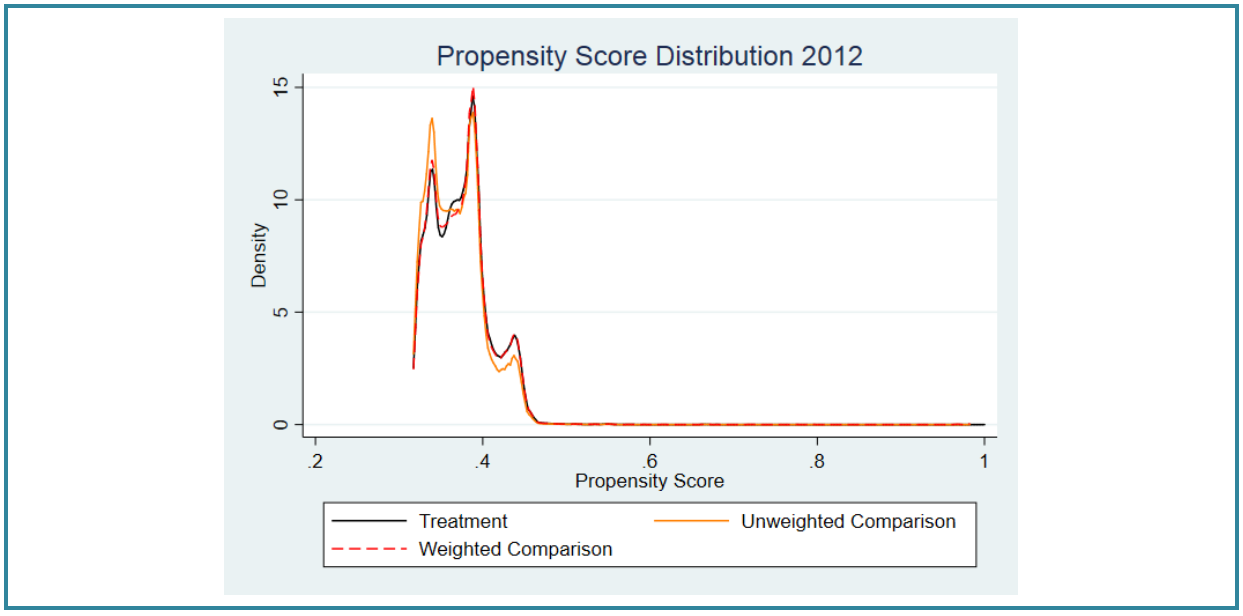
Figure D-2-19. Weighted and unweighted propensity score density plots for the Medicaid IHP and comparison groups, 2011, using MN APCD data



IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

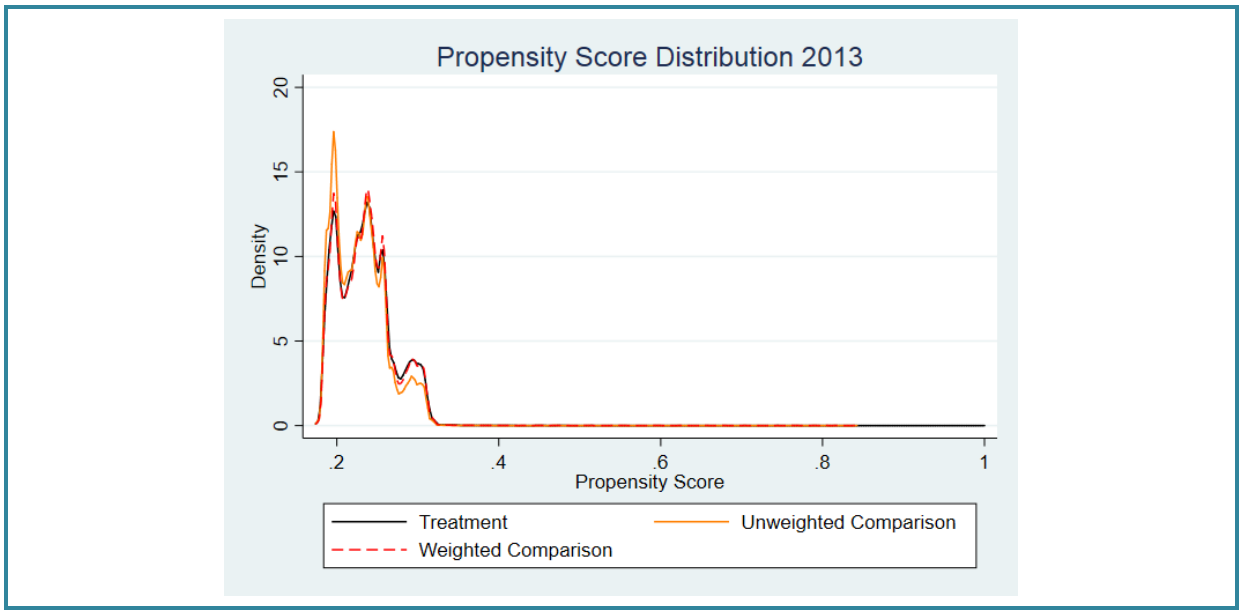
[35] In ***Figures D-2-18*** through ***D-2-29***, the IHP group is represented by the treatment line.

Figure D-2-20. Weighted and unweighted propensity score density plots for the Medicaid IHP and comparison groups, 2012, using MN APCD data



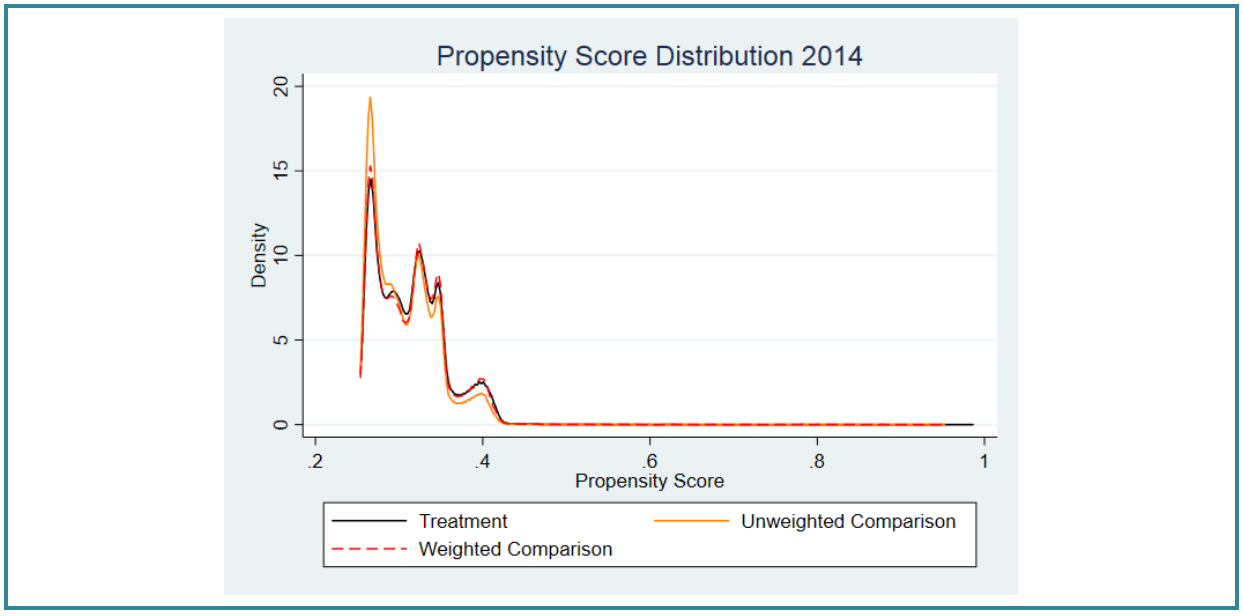
IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

Figure D-2-21. Weighted and unweighted propensity score density plots for the Medicaid IHP and comparison groups, 2013, using MN APCD data



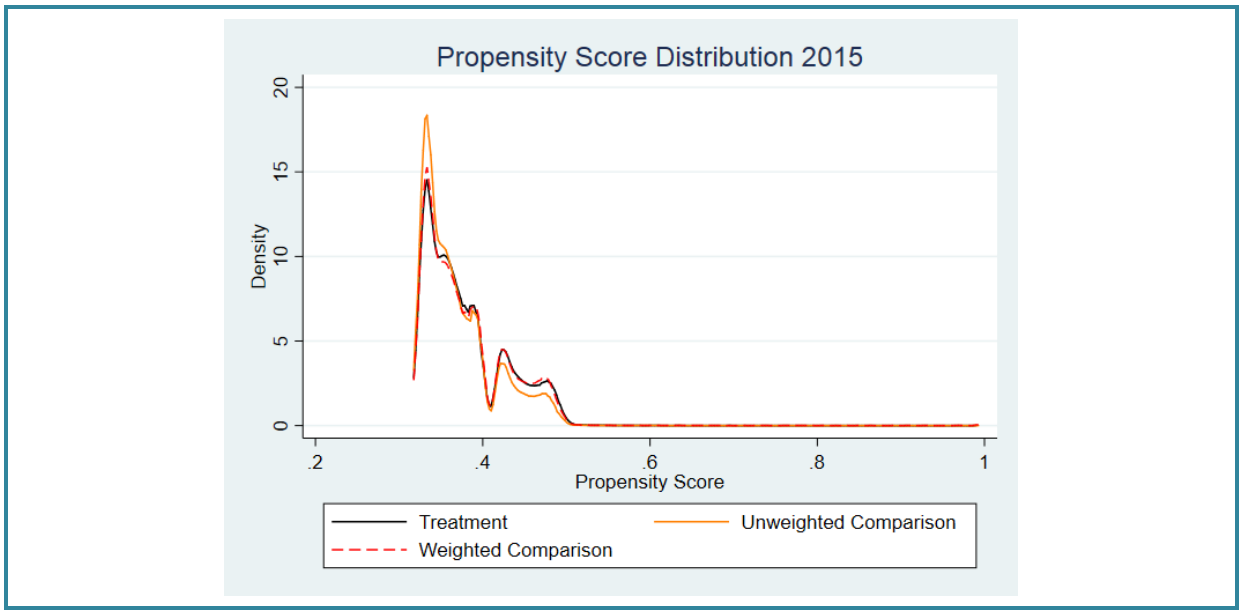
IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

Figure D-2-22. Weighted and unweighted propensity score density plots for the Medicaid IHP and comparison groups, 2014, using MN APCD data



IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

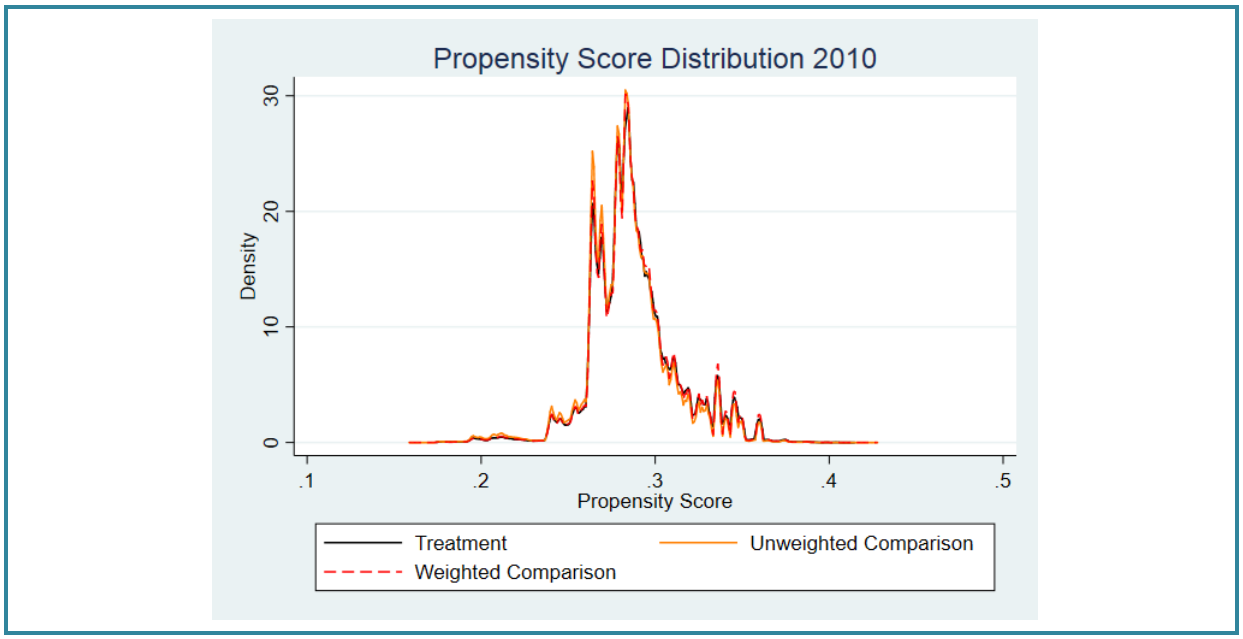
Figure D-2-23. Weighted and unweighted propensity score density plots for the Medicaid IHP and comparison groups, 2015, using MN APCD data



IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

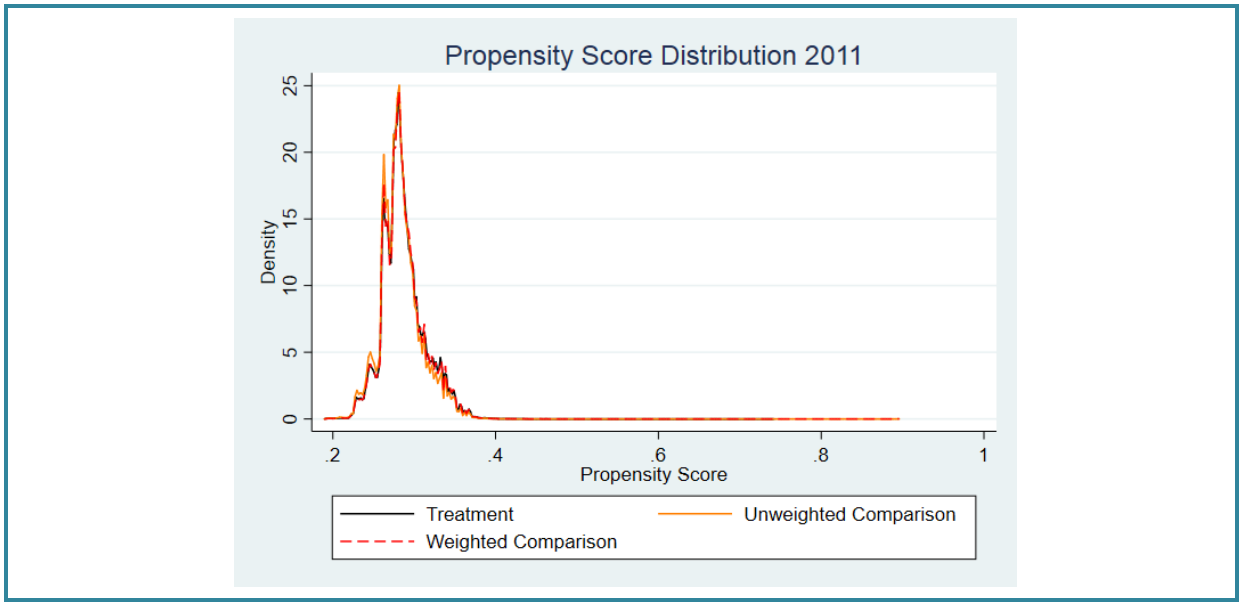
Figure D-2-24 through ***Figure D-2-29*** are the common support plots for the MN APCD-based commercial analysis. These figures also show a high level of overlap between the IHP and comparison groups.

Figure D-2-24. Weighted and unweighted propensity score density plots for the commercially insured IHP and comparison groups, 2010, using MN APCD data



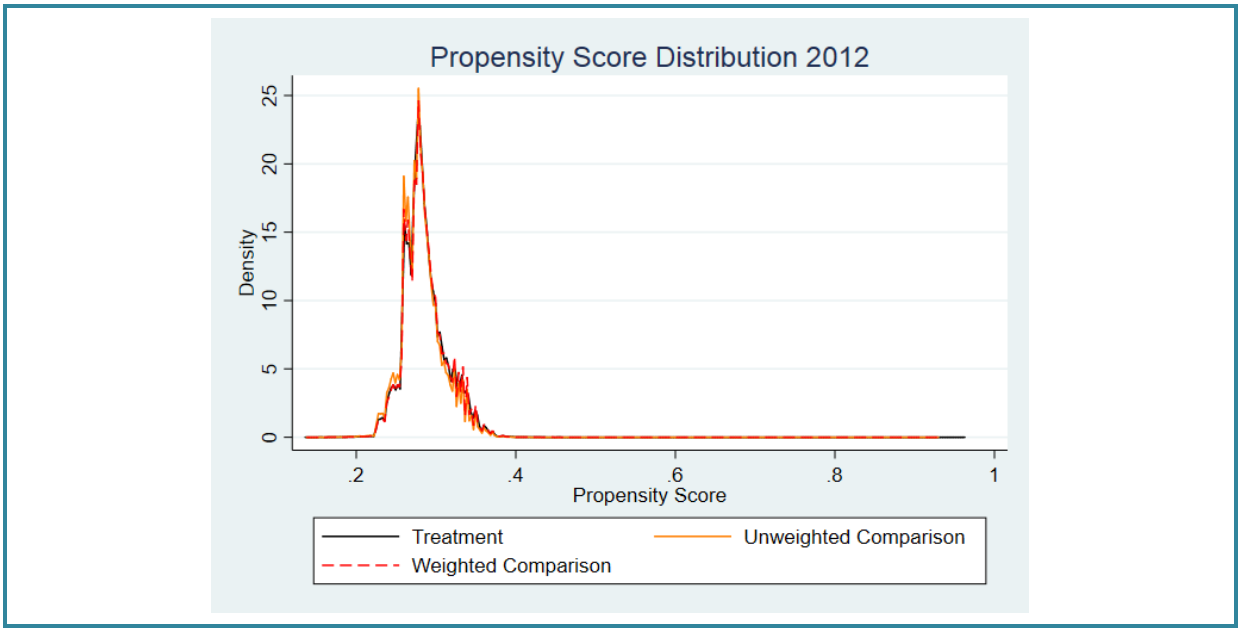
IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

Figure D-2-25. Weighted and unweighted propensity score density plots for the commercially insured IHP and comparison groups, 2011, using MN APCD data



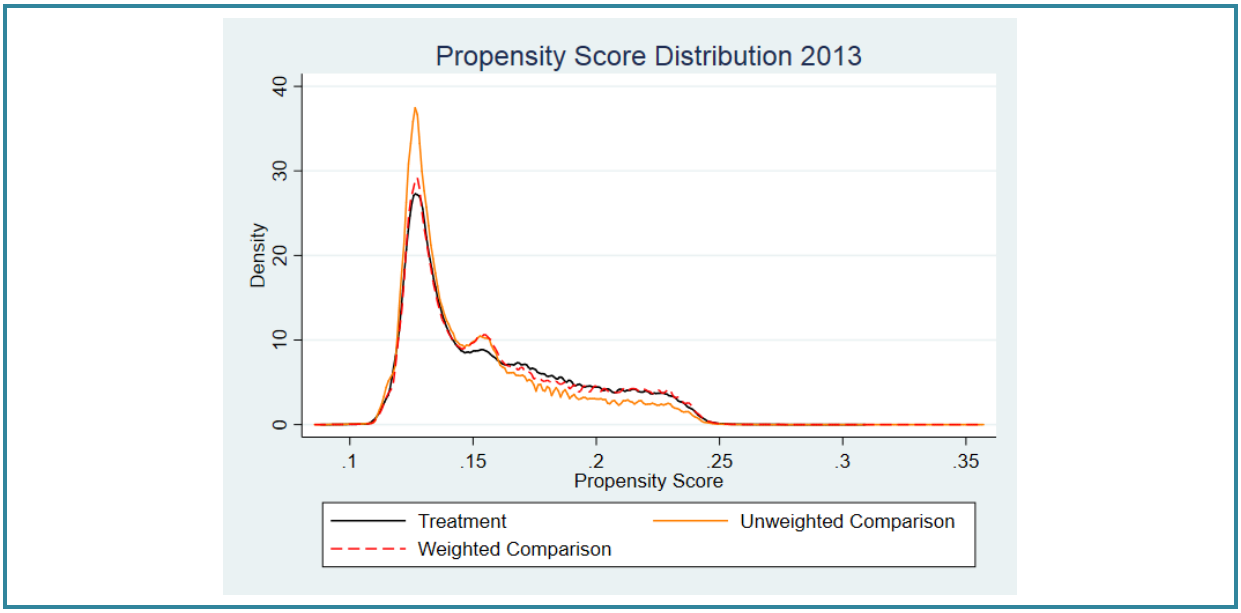
IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

Figure D-2-26. Weighted and unweighted propensity score density plots for the commercially insured IHP and comparison groups, 2012, using MN APCD data



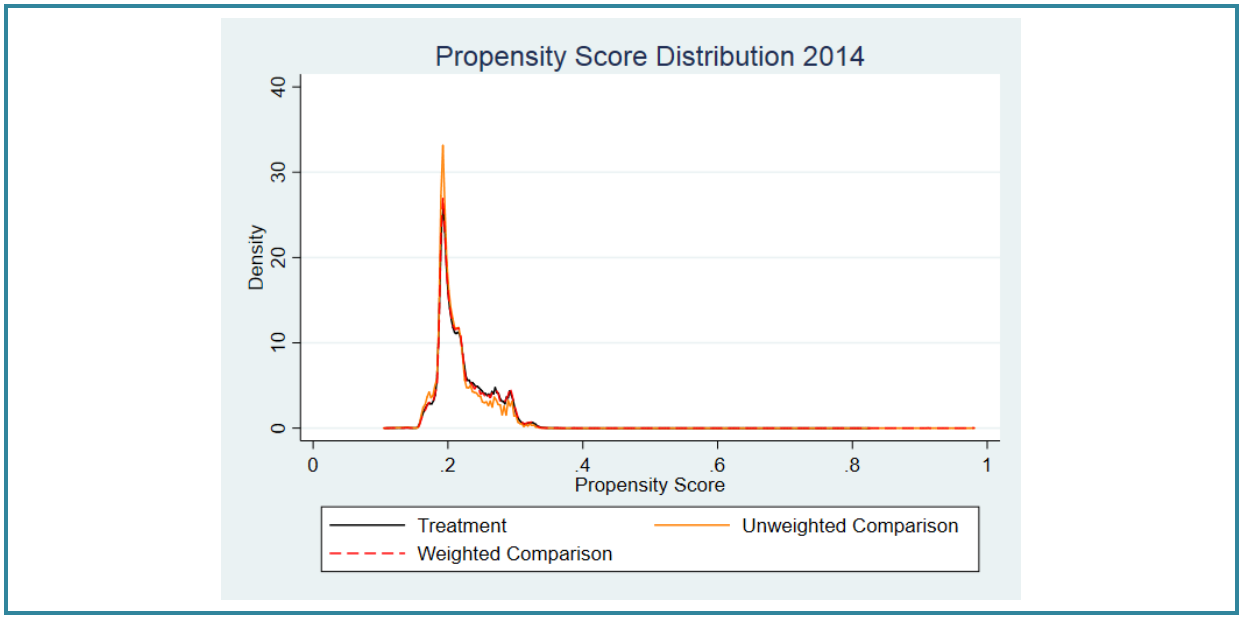
IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

Figure D-2-27. Weighted and unweighted propensity score density plots for the commercially insured IHP and comparison groups, 2013, using MN APCD data



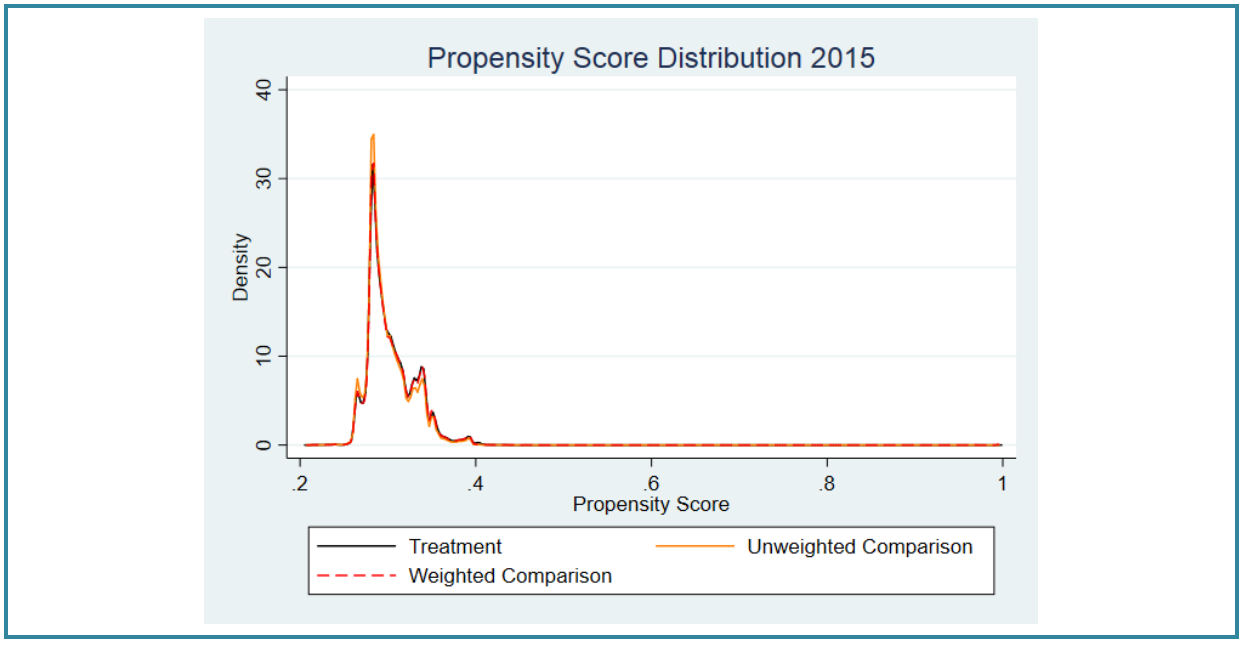
IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

Figure D-2-28. Weighted and unweighted propensity score density plots for the commercially insured IHP and comparison groups, 2014, using MN APCD data



IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

Figure D-2-29. Weighted and unweighted propensity score density plots for the commercially insured IHP and comparison groups, 2015, using MN APCD data



IHP = Integrated Health Partnership; MN APCD = Minnesota All Payer Claims Database.

In all years, we found that both the Medicaid and commercial comparison groups passed the common support assumption ($P(D = 1|X)>0$) for almost the entire range of the Medicaid and commercial IHP group's propensity scores, respectively. These plots provide ample evidence that the common support assumption is upheld.

Second, we compared the logistic results of the models to see which variables had the greatest impact on the propensity score weights. Overall, there were few substantial differences between the IHP and comparison groups in terms of person-level characteristics. Furthermore, most differences were not consistent across measurement years.

Finally, we compared unweighted and propensity-weighted means for the characteristics in the model. As expected, we found that, after weighting, the comparison group means were within a few percentage points of the values for the IHP group.

Tables D-2-17 to ***D-2-22*** show unweighted and propensity score weighted means/proportions for 2010–2015 for the MN APCD-based Medicaid analysis. ***Tables D-2-23*** to ***D-2-28*** show unweighted and propensity score weighted means/proportions for 2010 through 2015 for the MN APCD-based commercial analysis. With the exception of several county-level characteristics in select years, such as the number of hospital beds per 1,000 residents and the percent uninsured, all covariates were well under the commonly accepted threshold of less than 10 percent standardized difference after weighting.

Table D-2-17. Unweighted and weighted means and standardized differences, Medicaid IHP and comparison groups, 2010

Characteristic	Unweighted			Weighted			
	Comparison group	Medicaid IHP group	Standardized difference[a]	Comparison group	Medicaid IHP group	Standardized difference[a]	p-value
N	314,632	190,235		190,291	190,235		
Individual-level sociodemographic characteristics							
Age	21.1	20.2	4.9	20.2	20.2	0.03	0.92
Age (squared)	719.5	691.4	3.1	692.0	691.4	0.1	0.83
HCC Risk Score	1.8	2.0	3.7	2.1	2.0	1.5	0.00
Female (%)	56.7	56.7	0.2	56.7	56.7	0.02	0.95
Continuously enrolled in 2010 (%)	61.6	63.9	4.8	63.9	63.9	0.1	0.87
County-level characteristics							
Median age	37.6	37.6	0.5	37.6	37.6	0.4	0.20
Percent below the poverty line	12.2	13.0	20.3	12.2	13.0	20.2	0.00
Hospital beds per 1,000 residents	3.1	3.6	16.0	3.1	3.6	16.0	0.00
Percent without health insurance	10.7	10.9	10.3	10.7	10.9	10.2	0.00
Rural/urban continuum (%)	72.1	76.7	10.5	72.1	76.7	10.6	0.00

HCC = Hierarchical Condition Category; IHP = Integrated Health Partnership.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-18. Unweighted and weighted means and standardized differences, Medicaid IHP and comparison groups, 2011

Characteristic	Unweighted			Weighted			
	Comparison group	Medicaid IHP group	Standardized difference[a]	Comparison group	Medicaid IHP group	Standardized difference[a]	p-value
N	362,894	216,669		216,704	216,669		
Individual-level sociodemographic characteristics							
Age	21.8	20.9	5.5	20.9	20.9	0.1	0.85
HCC Risk Score	1.8	2.0	2.9	2.1	2.0	0.4	0.19
Female (%)	56.4	56.5	0.3	56.5	56.5	0.01	0.98
Continuously enrolled in 2011 (%)	53.3	55.3	4.0	55.3	55.3	0.1	0.81
Enrolled for at least 9 months in 2010 (%)	64.2	66.9	5.8	66.9	66.9	0.1	0.85
Any ED visits in 2010 (%)	24.3	28.1	8.6	28.1	28.1	0.01	0.96
Any inpatient admissions in 2010 (%)	8.2	9.2	3.7	9.2	9.2	0.02	0.95
Total spending in 2010 (PBPM)	331.4	382.4	3.0	414.3	382.4	1.2	0.00
County-level characteristics							
Median age	37.6	37.5	0.8	37.6	37.5	1.0	0.00
Percent below the poverty line	12.6	13.3	18.9	12.6	13.3	18.8	0.00
Hospital beds per 1,000 residents	3.1	3.6	16.5	3.1	3.6	16.6	0.00
Percent without health insurance	10.6	10.7	5.0	10.6	10.7	4.7	0.00
Rural/urban continuum (%)	72.5	77.0	10.5	72.3	77.0	10.7	0.00

HCC = Hierarchical Condition Category; ED = emergency department; IHP = Integrated Health Partnership; PBPM = per beneficiary per month.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-19. Unweighted and weighted means and standardized differences, Medicaid IHP and comparison groups, 2012

Characteristic	Unweighted			Weighted			
	Comparison group	Medicaid IHP group	Standardized difference[a]	Comparison group	Medicaid IHP group	Standardized difference[a]	p-value
N	406,203	239,245		239,288	239,245		
Individual-level sociodemographic characteristics							
Age	22.2	21.0	7.0	21.0	21.0	0.1	0.86
HCC Risk Score	1.9	2.1	2.5	2.1	2.1	0.2	0.44
Female (%)	56.1	56.3	0.3	56.3	56.3	0.0	1.00
Continuously enrolled in 2012 (%)	52.0	53.7	3.5	53.7	53.7	0.1	0.82
Enrolled for at least 9 months in 2011 (%)	62.5	64.9	4.9	64.8	64.9	0.1	0.81
Any ED visits in 2011 (%)	25.2	29.5	9.7	29.5	29.5	0.01	0.96
Any inpatient admissions in 2011 (%)	7.8	8.8	3.8	8.8	8.8	0.01	0.96
Total spending in 2011 (PBPM)	357.1	412.0	1.7	487.9	412.0	1.4	0.00
County-level characteristics							
Median age	37.5	37.5	0.9	37.5	37.5	1.0	0.00
Percent below the poverty line	12.1	12.7	17.2	12.1	12.7	16.9	0.00
Hospital beds per 1,000 residents	3.1	3.5	17.1	3.1	3.5	17.0	0.00
Percent without health insurance	9.6	9.7	3.9	9.6	9.7	3.5	0.00
Rural/Urban Continuum (%)	72.8	77.1	10.0	72.6	77.1	10.4	0.00

HCC = Hierarchical Condition Category; ED = emergency department; IHP = Integrated Health Partnership; PBPM = per beneficiary per month.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-20. Unweighted and weighted means and standardized differences, Medicaid IHP and comparison groups, 2013

Characteristic	Unweighted			Weighted			
	Comparison group	Medicaid IHP group	Standardized difference[a]	Comparison group	Medicaid IHP group	Standardized difference[a]	p-value
N	348,642	104,383		104,399	104,383		
Individual-level sociodemographic characteristics							
Age	23.3	21.4	10.6	21.4	21.4	0.1	0.91
HCC Risk Score	2.1	2.2	2.1	2.2	2.2	0.2	0.62
Female (%)	56.6	55.9	1.5	55.9	55.9	0.01	0.99
Continuously enrolled in 2013 (%)	60.1	62.3	4.7	62.3	62.3	0.04	0.93
Enrolled for at least 9 months in 2012 (%)	66.6	67.8	2.6	67.8	67.8	0.1	0.76
Any ED visits in 2012 (%)	28.3	33.3	10.9	33.3	33.3	0.1	0.86
Any inpatient admissions in 2012 (%)	9.3	9.9	2.0	9.9	9.9	0.0	0.99
Total spending in 2012 (PBPM)	432.8	513.4	2.8	538.5	513.4	0.7	0.12
County-level characteristics							
Median age	37.5	37.5	0.02	37.5	37.5	0.03	0.95
Percent below the poverty line	11.8	12.5	21.5	11.8	12.5	21.2	0.00
Hospital beds per 1,000 residents	3.1	3.5	15.7	3.1	3.5	15.8	0.00
Percent without health insurance	9.9	10.0	8.1	9.9	10.0	7.7	0.00
Rural/Urban Continuum (%)	72.8	82.6	23.8	72.6	82.6	24.2	0.00

HCC = Hierarchical Condition Category; ED = emergency department; IHP = Integrated Health Partnership; PBPM = per beneficiary per month.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-21. Unweighted and weighted means and standardized differences, Medicaid IHP and comparison groups, 2014

Characteristic	Unweighted			Weighted			
	Comparison group	Medicaid IHP group	Standardized difference[a]	Comparison group	Medicaid IHP group	Standardized difference[a]	p-value
N	419,829	186,670		186,739	186,670		
Individual-level sociodemographic characteristics							
Age	24.7	22.8	10.2	22.8	22.8	0.1	0.87
HCC Risk Score	1.6	1.8	3.3	1.8	1.8	0.5	0.10
Female (%)	56.5	55.4	2.3	55.4	55.4	0.1	0.87
Continuously enrolled in 2014 (%)	68.3	70.2	4.0	70.2	70.2	0.1	0.84
Enrolled for at least 9 months in 2013 (%)	58.2	60.3	4.3	60.3	60.3	0.1	0.84
Any ED visits in 2013 (%)	20.7	25.7	12.0	25.7	25.7	0.1	0.83
Any inpatient admissions in 2013 (%)	7.1	8.3	4.5	8.4	8.3	0.2	0.59
Total spending in 2013 (PBPM)	355.1	441.2	3.7	516.9	441.2	1.9	0.00
County-level characteristics							
Median age	37.5	37.4	2.5	37.5	37.5	2.7	0.00
Percent below the poverty line	11.9	12.4	16.7	11.9	12.4	16.4	0.00
Hospital beds per 1,000 residents	3.0	3.7	24.4	3.0	3.7	24.4	0.00
Percent without health insurance	7.0	7.1	6.5	7.0	7.1	5.9	0.00
Rural/urban continuum (%)	73.8	77.7	9.1	73.6	77.7	9.8	0.00

HCC = Hierarchical Condition Category; ED = emergency department; IHP = Integrated Health Partnership; PBPM = per beneficiary per month.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-22. Unweighted and weighted means and standardized differences, Medicaid IHP and comparison groups, 2015

Characteristic	Unweighted			Weighted			
	Comparison group	Medicaid IHP group	Standardized difference[a]	Comparison group	Medicaid IHP group	Standardized difference[a]	p-value
N	441,391	262,419		262,519	262,419		
Individual-level sociodemographic characteristics							
Age	25.3	24.1	6.4	24.1	24.1	0.1	0.70
HCC Risk Score	1.4	1.6	3.2	1.6	1.6	0.5	0.10
Female (%)	56.6	55.5	2.2	55.5	55.5	0.04	0.88
Continuously enrolled in 2015 (%)	68.4	68.9	1.2	68.9	68.9	0.1	0.85
Enrolled for at least 9 months in 2014 (%)	65.3	67.7	5.2	67.7	67.7	0.1	0.78
Any ED visits in 2014 (%)	17.7	24.1	15.8	24.1	24.1	0.04	0.89
Any inpatient admissions in 2014 (%)	6.4	7.9	5.8	8.0	7.9	0.2	0.42
Total spending in 2014 (PBPM)	331.0	422.9	3.8	548.5	422.9	2.3	0.00
County-level characteristics							
Median age	37.5	37.5	1.9	37.5	37.5	2.1	0.00
Percent below the poverty line	10.6	11.2	19.6	10.6	11.2	19.0	0.00
Hospital beds per 1,000 residents	3.0	3.5	21.0	3.0	3.5	21.0	0.00
Percent without health insurance	5.4	5.4	5.5	5.4	5.4	4.7	0.00
Rural/urban continuum (%)	73.9	76.4	5.8	73.6	76.4	6.4	0.00

HCC = Hierarchical Condition Category; ED = emergency department; IHP = Integrated Health Partnership; PBPM = per beneficiary per month.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-23. Unweighted and weighted means and standardized differences, commercially insured IHP and comparison groups, 2010

Characteristic	Unweighted			Weighted			
	Comparison group	Commercial IHP group	Standardized difference[a]	Comparison group	Commercial IHP group	Standardized difference[a]	p-value
N	1,421,408	568,829		568,827	568,829		
Individual-level sociodemographic characteristics							
Female (%)	51.7	53.3	3.1	53.3	53.3	0.03	0.86
Continuously enrolled in 2010 (%)	87.8	88.8	2.9	87.8	88.8	0.03	0.88
Age	32.9	31.8	6.0	31.8	31.8	0.1	0.65
Age (Squared)	1,408.9	1,358.5	4.4	1,359.70	1358.5	0.1	0.60
HCC Risk Score (logged)	−0.4	−0.4	0.4	−0.4	−0.4	0.1	0.74
One or more months of pharmacy coverage (%)	97.8	98.5	5.2	98.5	98.5	0.0	0.99
Individual is primary policyholder (%)	48.2	45.6	5.1	45.7	45.6	0.1	0.80
Spouse to primary policyholder (%)	18.9	18.5	0.9	18.6	18.5	0.1	0.77
Child to primary policyholder (%)	32.8	35.7	6.1	35.7	35.7	0.1	0.63
Other relationship to primary policyholder (%)	0.1	0.1	0.1	0.1	0.1	0.01	0.95
County-level characteristics							
Rural/urban continuum (%)	80.3	79.3	2.7	80.3	79.3	2.7	0.00
Percent without health insurance	10.1	10.3	14.5	10.1	10.3	14.7	0.00
Percent below the poverty line	10.8	11.4	16.2	10.8	11.4	16.5	0.00
Hospital beds per 1,000 residents	2.8	3.5	28.6	2.7	3.5	28.7	0.00
Median age	37.2	37.4	7.3	37.2	37.4	7.3	0.00

HCC = Hierarchical Condition Category; IHP = Integrated Health Partnership.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-24. Unweighted and weighted means and standardized differences, commercially insured IHP and comparison groups, 2011

Characteristic	Unweighted			Weighted			
	Comparison group	Commercial IHP group	Standardized difference[a]	Comparison group	Commercial IHP group	Standardized difference[a]	p-value
N	1,549,069	615,642		615,671	615,642		
Individual-level sociodemographic characteristics							
Female (%)	51.8	53.4	3.0	53.4	53.4	0.04	0.83
Continuously enrolled in 2011 (%)	86.1	87.4	3.9	87.4	87.4	0.01	0.96
Enrolled for at least 9 months in 2010 (%)	83.2	84.5	3.4	84.5	84.5	0.02	0.90
Age	33.2	32	6.2	32.1	32	0.1	0.56
Age (Squared)	1,433.7	1,381.1	4.4	1,382.60	1,381.1	0.1	0.50
HCC Risk Score (logged)	−0.4	−0.4	0.2	−0.4	−0.4	0.1	0.54
Total spending in 2010 (PMPM)	184.1	213.0	2.4	231.3	213	0.9	0.00
Any inpatient admissions in 2010 (%)	3.7	4.2	2.5	4.2	4.2	0.0	0.99
Any ED visits in 2010 (%)	8.5	9.5	3.6	9.5	9.5	0.0	0.99
One or more months of pharmacy coverage (%)	98.1	98.5	2.7	98.5	98.5	0.01	0.96
Individual is primary policyholder (%)	47.8	45.1	5.4	45.1	45.1	0.1	0.75
Spouse of primary policyholder (%)	18.6	18.2	1.0	18.3	18.2	0.1	0.78
Child of primary policyholder (%)	33.5	36.5	6.5	36.5	36.5	0.1	0.59
Other relationship to primary policyholder (%)	0.1	0.2	0.6	0.2	0.2	0.03	0.86
County-level characteristics							
Rural/urban continuum (%)	80.1	79.2	2.3	80.1	79.2	2.4	0.00
Percent without health insurance	9.9	10.1	10.2	9.9	10.1	10.5	0.00
Percent below the poverty line	11.1	11.8	17.1	11.1	11.8	17.4	0.00
Hospital beds per 1,000 residents	2.8	3.5	28.7	2.8	3.5	28.8	0.00
Median age	37.2	37.4	7.0	37.2	37.4	7.1	0.00

HCC = Hierarchical Condition Category; ED = emergency department; IHP = Integrated Health Partnership; PMPM = per member per month.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-25. Unweighted and weighted means and standardized differences, commercially insured IHP and comparison groups, 2012

Characteristic	Unweighted			Weighted			
	Comparison group	Commercial IHP group	Standardized difference[a]	Comparison group	Commercial IHP group	Standardized difference[a]	p-value
N	1,666,800	659,534		659,560	659,534		
Individual-level sociodemographic characteristics							
Female (%)	52	53.5	2.9	53.5	53.5	0.1	0.79
Continuously enrolled in 2012 (%)	86.2	87.4	3.6	86.2	87.4	0.04	0.82
Enrolled for at least 9 months in 2011 (%)	83.0	84.3	3.6	84.3	84.3	0.02	0.89
Age	33.5	32.3	6.5	32.3	32.3	0.1	0.56
Age (Squared)	1459.9	1404.1	4.6	1,405.60	1404.1	0.1	0.48
HCC Risk Score (logged)	−0.4	−0.4	0.3	−0.4	−0.4	0.1	0.52
Total spending in 2011 (PMPM)	196.7	228.4	2.2	241.9	228.4	0.6	0.00
Any inpatient admissions in 2011 (%)	4.0	4.4	2.2	4.4	4.4	0.02	0.92
Any ED visits in 2011 (%)	9.2	10.4	4.0	10.4	10.4	0.01	0.96
One or more months of pharmacy coverage (%)	98.5	98.7	1.8	98.7	98.7	0.0	0.98
Individual is primary policyholder (%)	47.8	45	5.7	45.0	45.0	0.1	0.75
Spouse of primary policyholder (%)	18.3	17.8	1.3	17.9	17.8	0.1	0.79
Child of primary policyholder (%)	33.7	37.1	7.1	37.1	37.1	0.1	0.59
Other relationship to primary policyholder (%)	0.2	0.1	1.4	0.1	0.1	0.0	0.98
County-level characteristics							
Rural/urban continuum (%)	80.1	79.1	2.5	80	79.1	2.3	0.00
Percent without health insurance	9	9.1	8.1	9	9.1	8.4	0.00
Percent below the poverty line	10.8	11.4	18.5	10.7	11.4	18.9	0.00
Hospital beds per 1,000 residents	2.8	3.5	28.8	2.8	3.5	29.0	0.00
Median age	37.2	37.4	7.2	37.2	37.4	7.1	0.00

HCC = Hierarchical Condition Category; ED = emergency department; IHP = Integrated Health Partnership; PMPM = per member per month.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-26. Unweighted and weighted means and standardized differences, commercially insured IHP and comparison groups, 2013

Characteristic	Unweighted			Weighted			
	Comparison group	Commercial IHP group	Standardized difference[a]	Comparison group	Commercial IHP group	Standardized difference[a]	p-value
N	1,464,910	260,620		260,629	260,620		
Individual-level sociodemographic characteristics							
Female (%)	54.4	54.3	0.2	54.3	54.3	0.1	0.86
Continuously enrolled in 2013 (%)	87.4	88.0	1.7	88.0	88.0	0.1	0.83
Enrolled for at least 9 months in 2012 (%)	86.2	85.6	1.7	85.5	85.6	0.1	0.69
Age	34.3	31.0	16.9	31.0	31.0	0.04	0.90
Age (Squared)	1529.1	1364.8	12.7	1,365.60	1364.8	0.1	0.83
HCC Risk Score (logged)	−0.3	−0.4	5.9	−0.4	−0.4	0.1	0.84
Total spending in 2012 (PMPM)	260	278	1.0	279.4	278	0.1	0.82
Any inpatient admissions in 2012 (%)	4.7	5.1	1.5	5.1	5.1	0.03	0.92
Any ED visits in 2012 (%)	10.7	11.3	2.1	11.4	11.3	0.1	0.87
One or more months of pharmacy coverage (%)	98.6	98.7	0.8	98.7	98.7	0.0	0.99
Individual is primary policyholder (%)	47.3	40.5	13.7	40.5	40.5	0.03	0.92
Spouse of primary policyholder (%)	18.6	16	7.1	16	16	0.02	0.95
Child of primary policyholder (%)	33.9	43.4	19.6	43.4	43.4	0.04	0.89
Other relationship to primary policyholder (%)	0.2	0.1	1.0	0.1	0.1	0.01	0.97
County-level characteristics							
Rural/urban continuum (%)	80.5	85.3	12.8	80.4	85.3	13.1	0.00
Percent without health insurance	9.2	9.4	11.5	9.2	9.4	12.2	0.00
Percent below the poverty line	10.6	11.2	18.0	10.6	11.2	18.6	0.00
Hospital beds per 1,000 residents	2.7	3.3	21.0	2.7	3.3	21.3	0.00
Median age	37.2	37.4	8.3	37.2	37.4	8.1	0.00

HCC = Hierarchical Condition Category; ED = emergency department; IHP = Integrated Health Partnership; PMPM = per member per month.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-27. Unweighted and weighted means and standardized differences, commercially insured IHP and comparison groups, 2014

Characteristic	Unweighted			Weighted			
	Comparison group	Commercial IHP group	Standardized difference[a]	Comparison group	Commercial IHP group	Standardized difference[a]	p-value
N	1,395,914	382,854		382,935	382,854		
Individual-level sociodemographic characteristics							
Female (%)	54.3	54.2	0.1	54.2	54.2	0.02	0.93
Continuously enrolled in 2014 (%)	86.9	88.3	4.4	88.3	88.3	0.02	0.92
Enrolled for at least 9 months in 2013 (%)	86.2	85.6	1.7	85.5	85.6	0.1	0.75
Age	34.3	31.7	13.1	31.7	31.7	0.02	0.92
Age (Squared)	1532.3	1405.5	9.8	1,405.60	1405.5	0.01	0.98
HCC Risk Score (logged)	−0.3	−0.4	4.3	−0.4	−0.4	0.2	0.50
Total spending in 2013 (PMPM)	227.6	280.7	2.5	381.7	280.7	1.7	0.00
Any inpatient admissions in 2013 (%)	4.2	4.7	2.7	4.7	4.7	0.04	0.88
Any ED visits in 2013 (%)	8.9	10.4	5.2	10.4	10.4	0.02	0.93
One or more months of pharmacy coverage (%)	98.9	98.9	0.5	98.9	98.9	0.01	0.95
Individual is primary policyholder (%)	47.2	42.2	10.2	42.2	42.2	0.0	0.99
Spouse of primary policyholder (%)	18.7	16.9	4.8	16.9	16.9	0.0	0.99
Child of primary policyholder (%)	33.9	40.9	14.4	40.9	40.9	0.0	1.00
Other relationship to primary policyholder (%)	0.2	0.1	1.6	0.1	0.1	0.01	0.97
County-level characteristics							
Rural/urban continuum (%)	80.9	76.9	9.8	80.8	76.9	9.7	0.00
Percent without health insurance	6.6	6.6	7.3	6.5	6.6	8.0	0.00
Percent below the poverty line	10.7	11.2	13.8	10.7	11.2	14.4	0.00
Hospital beds per 1,000 residents	2.7	3.8	40.7	2.7	3.8	41.1	0.00
Median age	37.2	37.6	12.8	37.2	37.6	12.6	0.00

HCC = Hierarchical Condition Category; ED = emergency department; IHP = Integrated Health Partnership; PMPM = per member per month.

[a] Absolute standardized differences are expressed as percentages.

Table D-2-28. Unweighted and weighted means and standardized differences, commercially insured IHP and comparison groups, 2015

Characteristic	Unweighted			Weighted			
	Comparison group	Commercial IHP group	Standardized difference[a]	Comparison group	Commercial IHP group	Standardized difference[a]	p-value
N	1,317,973	567,624		567,761	567,624		
Individual-level sociodemographic characteristics							
Female (%)	54.5	54.2	0.5	54.2	54.2	0.02	0.90
Continuously enrolled in 2015 (%)	73.2	73.7	1.1	73.2	73.7	0.01	0.94
Enrolled for at least 9 months in 2014 (%)	84.3	85.2	2.5	84.3	85.2	0.02	0.92
Age	34.4	33.6	4.5	33.6	33.6	0.02	0.90
Age (Squared)	1,545.60	1,514.10	2.4	1,514.10	1,514.10	0.0	1.0
HCC Risk Score (logged)	−0.4	−0.4	0.7	−0.4	−0.4	0.2	0.22
Total spending in 2014 (PMPM)	216.8	278.5	2.5	385.6	278.5	1.9	0.00
Any inpatient admissions in 2014 (%)	4.0	4.5	2.3	4.5	4.5	0.04	0.85
Any ED visits in 2014 (%)	8.2	10.1	6.6	10.1	10.1	0.04	0.83
One or more months of pharmacy coverage (%)	98.4	98.5	0.6	98.5	98.5	0.02	0.91
Individual is primary policyholder (%)	47.5	45.7	3.7	45.7	45.7	0.1	0.97
Spouse of primary policyholder (%)	18.5	17.5	2.4	17.5	17.5	0.02	0.91
Child of primary policyholder (%)	33.8	36.6	5.9	36.7	0.4	0.02	0.90
Other relationship to primary policyholder (%)	0.2	0.1	0.9	0.1	0.1	0.01	0.95
County-level characteristics							
Rural/urban continuum (%)	81.1	78.7	6.02	81.1	78.7	5.8	0.00
Percent without health insurance	5.0	5.1	7.3	5.0	5.0	7.5	0.00
Percent below the poverty line	9.5	10.3	25.8	9.5	9.5	25.9	0.00
Hospital beds per 1,000 residents	2.7	3.5	31.3	2.7	2.7	31.4	0.00
Median age	37.2	37.4	7.1	37,2	37.2	6.8	0.00

HCC = Hierarchical Condition Category; ED = emergency department; IHP = Integrated Health Partnership; PMPM = per member per month.

[a] Absolute standardized differences are expressed as percentages.

Propensity model evaluation for subpopulation

In addition to the overall model, we evaluated common support graphs and standardized differences of the propensity score models for the MN APCD Medicaid and MN APCD commercial subpopulation analyses. We found that we maintained balance across the IHP-attributed and comparison group beneficiaries among the subpopulations. Common support overlap looked very similar to plots presented for the overall Medicaid and commercial populations.

D-2.2.5 Statistical analysis

Regression model

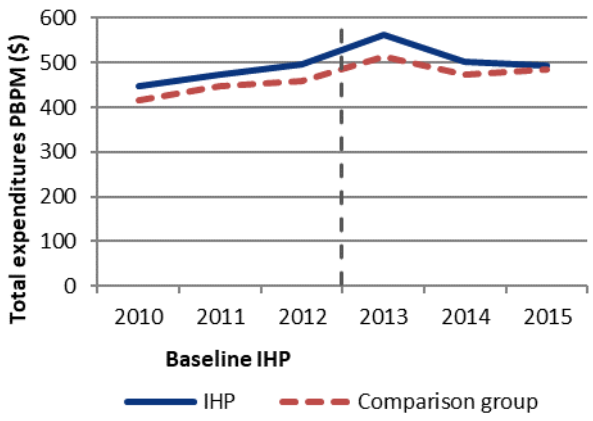
The underlying assumption in D-in-D models estimating the impact of IHPs on the Medicaid and commercial populations is that trends in the test group would be similar to that of the comparison group in the absence of the model (i.e., that the two were on "parallel paths" prior to the start of the IHPs).

To assess the parallel assumption's validity empirically for the MN APCD Medicaid and commercial analyses, we modeled core outcomes during the baseline period with a linear time trend interacted with a dichotomous variable indicating the beneficiary was attributed to an IHP provider (i.e., the "test" group).

To examine descriptively whether the trends in the IHP and comparison group for the MN APCD Medicaid and MN APCD commercial analyses are parallel, we present graphs of annual unadjusted averages for (1) IHP-attributed Medicaid beneficiaries and the Medicaid comparison group and (2) IHP-attributed commercial plan members and the commercial comparison group for the baseline period (2010–2012) and the first 3 years of the IHP implementation (2013–2015).

- For the MN APCD Medicaid analysis, the baseline trends were similar for total, facility, and professional expenditures.
- For the MN APCD commercial analysis, the baseline trends were similar for total medical spending PMPM, the ED visit rate per 1,000 population, the inpatient admission rate per 1,000 population, and the primary care visits per 1,000 population. Baseline trends for 14-day follow up visits per 1,000 discharges appeared parallel at baseline, but trends in 30-day readmissions per 1,000 discharges did not.

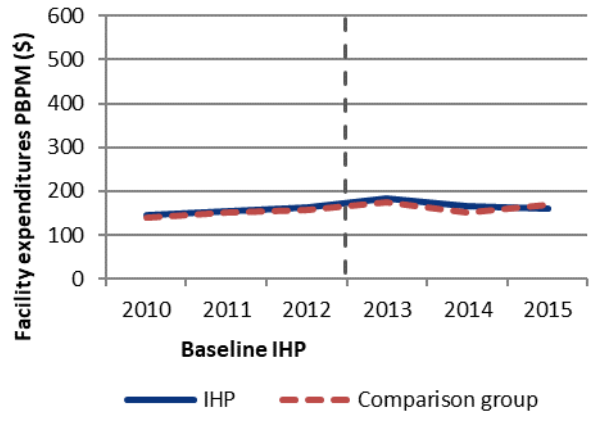
Figure D-2-30. Total medical expenditures PBPM, 2010 through 2015, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



Trends in **total medical expenditures PBPM** among Medicaid IHP-attributed beneficiaries and comparison group beneficiaries were similar throughout the baseline period (***Figure D-2-30***). Trends diverged between the IHP-attributed group and the comparison group from 2014 through 2015, the second and third years of the intervention period. Medicaid comparison group spending increased from 2014 through 2015, while spending for the Medicaid IHP group declined during that period.

IHP = Integrated Health Partnership; PBPM = per beneficiary per month.

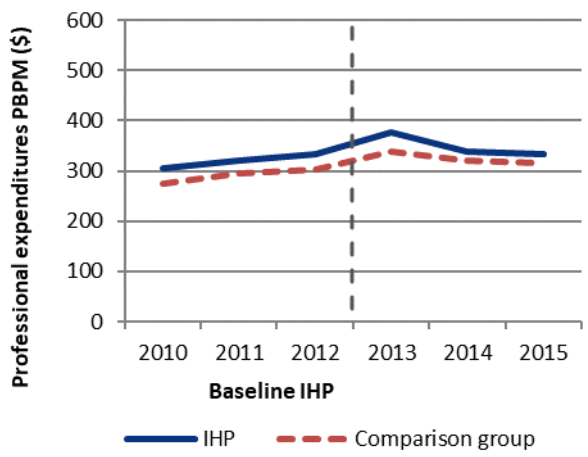
Figure D-2-31. Facility expenditures PBPM, 2010 through 2015, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



Trends in **facility expenditures PBPM** among Medicaid IHP-attributed beneficiaries and comparison group beneficiaries were similar throughout the baseline period (***Figure D-2-31***). However, trends diverged during the intervention period. Facility expenditures PBPM increased from 2014 to 2015 in the Medicaid comparison group but declined over the same period for the Medicaid IHP group.

IHP = Integrated Health Partnership; PBPM = per beneficiary per month.

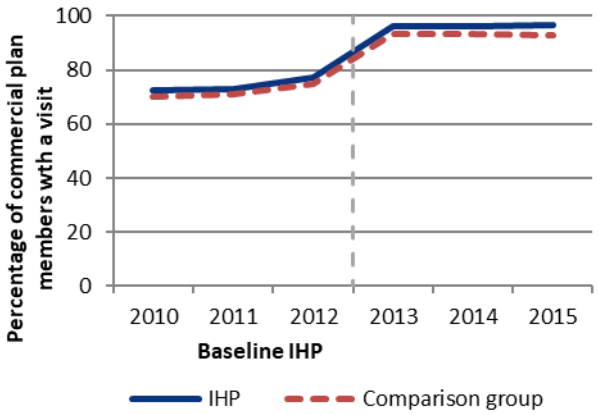
Figure D-2-32. Professional expenditures PBPM, 2010 through 2015, Minnesota Medicaid IHP-attributed beneficiaries and comparison group



Trends in **professional expenditures PBPM** among Medicaid IHP-attributed beneficiaries and comparison group beneficiaries were similar throughout the baseline period (***Figure D-2-32***). Unlike for total expenditures PBPM and facility expenditures PBPM, there is no difference in expenditure trends for the Medicaid IHP and comparison groups during the first three years of the intervention period.

IHP = Integrated Health Partnership; PBPM = per beneficiary per month.

Figure D-2-33. Percentage of commercial plan members with a visit to a primary care provider, 2010 through 2015, IHP-attributed commercial plan members and comparison group

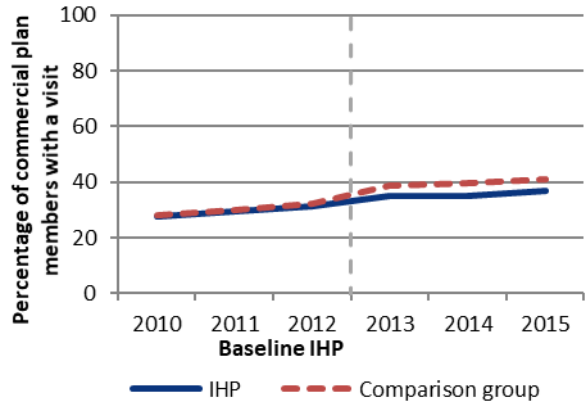


Trends in **the percent of commercial plan members with a primary care visit** were similar for the IHP group and the comparison group throughout the baseline period (***Figure D-2-33***). In both the IHP and comparison groups, the percentage of beneficiaries with a primary care visit increased markedly from 2013. This increase from 2012 to 2013 is due to our sample construction method. Having an E&M visit to a primary care provider is the most common way that an individual becomes attributed to an IHP or comparison group provider in a specific year. Therefore, individuals who are part of the sample from 2013 to 2015 are likely to have visited a primary care provider. On the other hand, we do not apply the attribution methodology to years prior to the IHP model launch in 2013. Because "attribution" is not a requirement for inclusion in the baseline sample, individuals are less likely to have a primary care visit during baseline.

The percentage of beneficiaries with a primary care visit increased slightly in the IHP group between 2013 and 2015 but declined over the same period in the comparison group.

IHP = Integrated Health Partnership; E&M = evaluation and management.

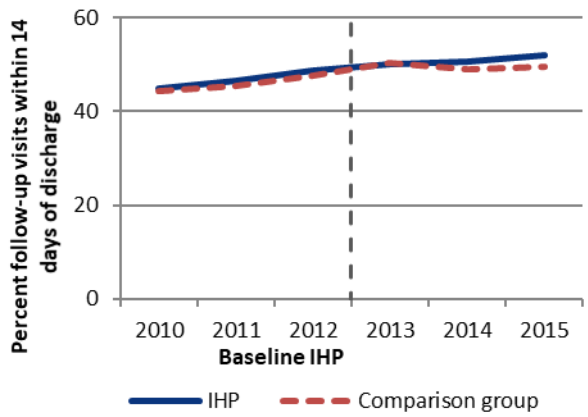
Figure D-2-34. Percentage of commercial plan members with a visit to a specialty care provider, 2010 through 2015, IHP-attributed commercial plan members and comparison group



The **percentages of commercial plan members with a visit to a specialty care provider** were similar in the IHP and comparison groups throughout the baseline period (***Figure D-2-34***). During the IHP implementation period, the rate increased more rapidly for the comparison group than for the IHP group.

IHP = Integrated Health Partnership.

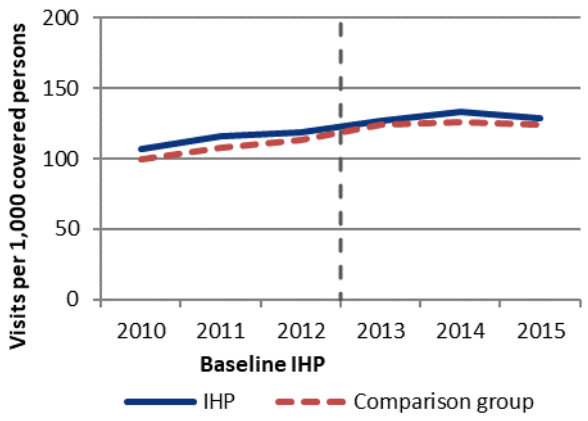
Figure D-2-35. Percentage of Medicaid beneficiaries with a follow-up visit within 14 days of discharge, 2010 through 2016, Medicaid IHP-attributed beneficiaries and comparison group



The percentage of hospitalizations with a **follow-up visit within 14 days of discharge** among IHP-attributed commercial plan members and the comparison group was similar throughout the baseline and IHP implementation period (***Figure D-2-35*** The percentage of hospitalizations in the IHP group that had a follow-up visit within 14 days of discharge increased more than the IHP group across the latter years of the IHP implementation period.

IHP = Integrated Health Partnership.

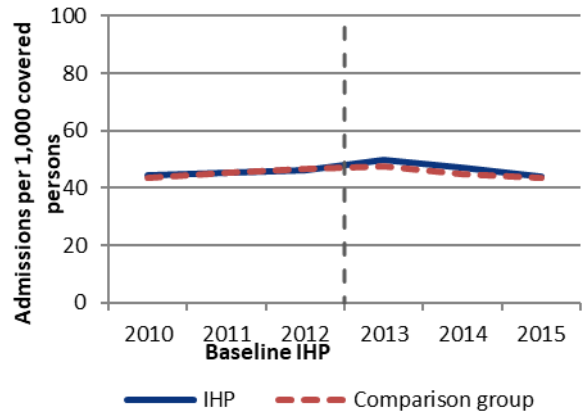
Figure D-2-36. ED visits per 1,000 covered persons, 2010 through 2015, IHP-attributed commercial plan members and comparison group



Trends in **ED visits per 1,000 covered persons** among IHP-attributed commercial plan members and commercial plan members in the comparison group were similar throughout the baseline period (***Figure D-2-36***). In both the IHP and comparison groups, the ED visit rate increased between 2010 and 2014 and declined slightly between 2014 and 2015.

ED = emergency department; IHP = Integrated Health Partnership.

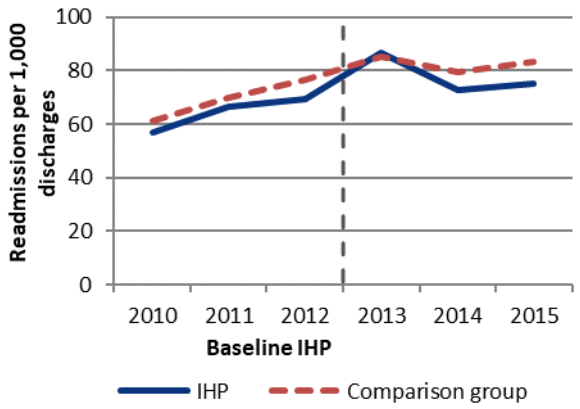
Figure D-2-37. Inpatient admissions per 1,000 covered persons, 2010 through 2015, IHP-attributed commercial plan members and comparison group



Trends in **inpatient admissions per 1,000 covered persons** among IHP-attributed commercial plan members and commercial plan members in the comparison group were similar throughout the baseline period (***Figure D-2-37).*** In both the IHP and comparison groups, the inpatient admissions increased from 2010 to 2013 and declined from 2013 to 2015.

IHP = Integrated Health Partnership.

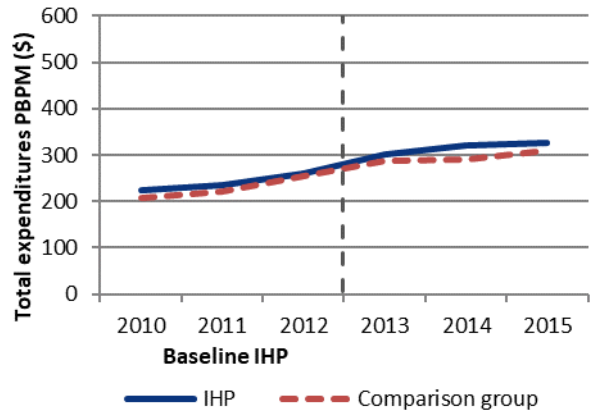
Figure D-2-38. Readmissions per 1,000 discharges, 2010 through 2015, IHP-attributed commercial plan members and comparison group



The rates of **readmissions per 1,000 covered persons** among IHP-attributed commercial plan members and commercial plan members in the comparison group trended upwards during the baseline period. (***Figure D-2-38***). The readmission rates for both the IHP and comparison groups peaked in 2013, declined from 2013 to 2014, and increased slightly from 2014 to 2015.

IHP = Integrated Health Partnership.

Figure D-2-39. Total medical expenditures PMPM, 2010 through 2015, IHP-attributed commercial plan members and comparison group



Trends in **total medical expenditures PMPM** among IHP-attributed commercial plan members and commercial plan members in the comparison group were similar throughout the baseline period (***Figure D-2-39***). Total medical expenditures PMPM increased consistently in both the IHP and comparison groups from the baseline period through the end of the IHP period.

IHP = Integrated Health Partnership; PMPM = per member per month.

To examine the effects of IHPs on expenditure outcomes for the Minnesota Medicaid population and expenditure, utilization, and care coordination outcomes for the Minnesota commercial population, we used an annual fixed-effects model as shown in ***Equation D-2.4***:

$$\gamma = \alpha_0 + \alpha_1 I + \sum \beta_n Q_{n,b} + \sum \phi_t Q_{t,p} \bullet I + \delta X + \mu \qquad \text{(D-2.4)}$$

where

y = a performance measure (e.g., total PBPM cost per year) for the i-th individual in the j-th group (test or comparison), in period t (i,j,t subscripts suppressed).

I = a 0,1 indicator (1 = individual is currently attributed to an IHP, 0 = individual is not currently attributed to an IHP or beneficiary is part of the comparison group).

X = a vector of patient and demographic characteristics.

$Q_{n,b}$, $Q_{t,d}$ = 0,1 indicator of the *n*-th or *t*-th calendar year in the base (*b*) or post (*p*) period (*n* starts counting at first baseline period, whereas *t* starts with first IHP year).

μ = error term.

The model in ***Equation D-2.4*** assumes that, except for an intercept difference α_1, the outcomes for individuals in the IHP group and individuals in the comparison groups followed a similar growth trend during the baseline period. For both the MN APCD Medicaid and MN APCD commercial analyses, we investigated whether the baseline period before the IHP implementation satisfied the baseline trend assumptions of the D-in-D model in ***Equation D-2.4***—that is, whether the outcome trends for individuals in IHP group and in the comparison group were similar during this period.

To test the similarity of baseline trends in both the MN APCD Medicaid and MN APCD commercial populations, we used a model with a linear trend during the baseline period. We tested whether this trend differed for IHP-attributed individuals relative to comparison group individuals. Specifically, the model for the outcomes may be written as follows.

$$y = \alpha_0 + \alpha_1 I + \theta \bullet t + \lambda I \bullet t + \delta X + \mu. \qquad \text{(D-2.5)}$$

In ***Equation D-2.5***, *y, I,* X, and μ are defined as in ***Equation D-2.4***. The variable *t* is linear time ranging from 1 to 3. The linear time trend in the comparison group is $\theta \bullet t$, whereas for IHP group beneficiaries ($I = 1$) it is $(\theta + \lambda) * t$. Hence, λ measures the difference in linear trends and the *t*-statistic for this coefficient can be used to test the null hypothesis of equal trends ($\lambda = 0$). In other words, rejecting the null hypothesis would suggest that the assumption of equal trends underlying our outcome models is not met.

The parameters of ***Equation D-2.5*** were estimated using weighted least-squares regression models for key outcomes for both the MN APCD Medicaid and commercial analyses. The weights are a function of the eligibility fraction and propensity scores. For each outcome, we report estimates and standard errors of the difference between the baseline trend in the IHP and the comparison groups (λ) separately for the MN APCD Medicaid and MN APCD commercial analyses.

Table D-2-29 shows estimates of the baseline trend differences for the following MN APCD Medicaid outcomes:

- Total medical expenditures PBPM
- Facility expenditures PBPM
- Professional expenditures PBPM

Table D-2-29. Differences in average expenditure outcomes during the baseline period, IHP-attributed beneficiaries and comparison group beneficiaries

Parameter estimate	Total PBPM ($)	Facility PBPM ($)	Professional PBPM ($)
Test–CG trend difference	−0.861	−1.055	0.194
	(2.685)	(2.272)	(1.101)

CG = comparison group; IHP = Integrated Health Partnership; PBPM = per beneficiary per month.

Baseline is the period January 2010–December 2012. The trend (slope) is the quarter-to-quarter change in the outcome variable. Standard errors are given in parentheses. *p < 0.10; **p < 0.05; ***p < 0.01.

Relative to the Medicaid comparison group, there were no differences in the baseline trends for total, facility, and professional PBPM expenditures in the Medicaid IHP group. Based on the overall results, we concluded that in general beneficiaries in the Medicaid IHP were on a similar trajectory with comparison beneficiaries prior to January 2013, and thus the parallel trend assumption of the D-in-D model was satisfied.

Table D-2-30 shows estimates of the baseline trend differences for the following MN APCD commercial analysis outcomes:

- Total medical expenditures PBPM
- Probability of an acute inpatient stay
- Probability of an outpatient ED visit

Table D-2-30. Differences in average expenditure and utilization outcomes during the baseline period, IHP-attributed commercial plan members and comparison group commercial plan members

Parameter estimate	Total medical expenditures PMPM ($)	Any inpatient admission	Any outpatient ED visit
Test–CG trend difference	−5.14	−0.0002	−0.0009
	(1.27)***	(0.0002)	(0.0003)***

CG = comparison group; ED = emergency department; IHP = Integrated Health Partnership; PMPM = per member per month.

Baseline is the period January 2010–December 2012. The trend (slope) is the quarter-to-quarter change in the outcome variable. Standard errors are given in parentheses. *p < 0.10; **p < 0.05; ***p < 0.01.

Relative to the comparison group, there were differences in the baseline trends for total medical spending and the probability of an ED visit but not for the probability of any inpatient admission. Because the baseline trends in the commercial IHP and comparison groups were visually similar for total expenditures and the ED visit rates and because the coefficient on the trend for the probability of any ED visit was small in magnitude, we concluded that it would be acceptable to construct the D-in-D models for the MN APCD commercial analysis assuming a parallel trend.

D-in-D regression model—We used the same basic D-in-D model for both the MN APCD Medicaid and MN APCD commercial analyses. The D-in-D model is shown in ***Equation D-2.6***. The model is an annual fixed effects model as shown in ***Equation D-2.4.*** As in ***Equation D-2.4***, Y_{ijt} is the outcome for individual i (IHP or comparison group) in state j in year t; I_{ij} (=0,1) is an indicator equal to 1 if the individual is in the IHP group and 0 if the individual is in its comparison group; Q_n is a series of yearly dummies for the baseline period (years 1 to 3); and Q_t is a series of yearly dummies for the post years (years x to x). The interaction of the IHP group indicator and Q_t (I_{ij}* Q_t) measures the difference in the pre-post change between the IHP group and its comparison states.

$$Y_{ijt} = \alpha_0 + \beta_1 I_{ij} + \sum \beta_2 Q_n + \sum \alpha_2 Q_t + \sum \gamma I_{ijt} * Q_t + \lambda X_{ijt} + \varepsilon_{ijt} \quad \text{(D-2.6)}$$

Table D-2-31 illustrates the interpretation of the D-in-D estimate from this model. The coefficient β_1 in ***Equation D-2.6*** is the difference in the measure between IHP-attributed individuals and comparison individuals at the start of the baseline period, holding constant other variables in the equation. The β_2 and α_2 coefficients are for the annual fixed effects and capture differences over time for each baseline and post year, respectively. The coefficient of the interaction term between Q_t and IHP (I) measures any differences for the IHP group relative to the comparison group in the post years relative to baseline years. Thus, in the post-period, the comparison group mean is captured by $\alpha_0 + \alpha_2$, whereas the IHP group mean is captured by $(\alpha_0 + \beta_1) + (\alpha_2 + \gamma)$. In other words, the between-group difference changes from β_1 during the baseline years to $\beta_1 + \gamma$ during the post-period. The D-in-D parameter, γ, shows whether the between-group difference increased ($\gamma>0$) or decreased ($\gamma<0$) after the IHP was implemented. Using the annual fixed effects model, we calculated overall estimates by taking linear combinations of the yearly estimates.

Table D-2-31. Difference-in-differences estimate

Group	Pre-period	Post-period	Pre-post difference
IHP	$\alpha_0 + \beta_1 + \beta_2$	$(\alpha_0 + \beta_1) + (\alpha_2 + \gamma)$	$\alpha_2 + \gamma$
Comparison	$\alpha_0 + \beta_2$	$\alpha_0 + \alpha_2$	α_2
Between group	β_1	$\beta_1 + \gamma$	γ

IHP = Integrated Health Partnership.

All expenditure outcomes in the MN APCD Medicaid analysis are estimated with the person-year as the unit of analysis. For the MN APCD commercial analysis, total medical expenditures PMPM, the probability of an acute inpatient stay, the probability of an outpatient ED visit, the probability of any primary care visit in a year, and the probability of any specialist visit in a year also are observed at the person-year level. However, two outcomes included in the MN APCD commercial analysis—the probability of a hospital readmission within 30 days of discharge and the probability of a follow-up visit within 14 days of discharge—use admissions as the unit of analysis.

We converted utilization counts into binary outcomes (1 = any use) and used weighted logistic regression models. Count models are not appropriate because of the low occurrence of most types of utilization for individual beneficiaries in any year; however, we multiplied the marginal effect from the logistic regression models by 1,000 to obtain approximate rates of utilization per 1,000 beneficiaries. Multiplying the marginal effect by 1,000 does not produce an exact rate of utilization per 1,000 beneficiaries as it assumes no person has more than one visit or admission per year. However, we concluded that this is a reasonable approximation because only a small percentage of beneficiaries had counts exceeding 1 for any of the utilization measures. For expenditure outcomes, we used weighted generalized linear models with a normal distribution and identity link.

The models for inpatient admissions, ED visits, and readmissions were run separately for children and adults as well as among individuals with a mental or behavioral health condition.

Control variables. In the MN APCD Medicaid models we controlled for the following variables:

- Age (and age squared)
- Gender
- HCC score
- If beneficiary was continuously enrolled in Medicaid in the year
- If beneficiary was enrolled in Medicaid at least 9 months in previous year
- Median age for the beneficiary's county of residence
- Percent below the poverty line for the beneficiary's county of residence
- Hospital beds per 1,000 residents for the beneficiary's county of residence
- Percent without health insurance for the beneficiary's county of residence
- Classification of beneficiary's county of residence on the rural/urban continuum

We controlled not only for these same variables in the MN APCD commercial models but also for:

- At least 1 month of pharmaceutical coverage in a year
- Relationship to primary subscriber (self, spouse, dependent, or other)

Weighting and clustering. All regression models were estimated using weighted regressions and weighted by the propensity score times the eligibility fraction. In addition, standard errors were clustered at the individual level to account for repeated observations.

D-2.3 Methods for Qualitative Data Collection and Analysis

The Minnesota SIM Initiative Round 1 Evaluation team collected and analyzed a wide range of qualitative data in the fifth year of the federal SIM Initiative evaluation. These data sources included interviews with key informants and focus groups conducted during in-person site visits in previous evaluation years, a review of relevant documents, and regular evaluation calls with the state officials leading the state's SIM Initiative. This report draws from past evaluation reports, where further detail is provided on previously conducted site visit interviews and focus groups.

D-2.3.1 Document review

The evaluation team used Minnesota's quarterly and annual reports, operational plans, state-evaluation reports, and other state documents to obtain updated information on its implementation progress during the SIM Initiative test period. To supplement these documents, we collected relevant news articles on the Minnesota SIM Initiative activities and related initiatives, and we searched reform-oriented websites that the state maintains.

In addition, we obtained numbers of providers participating in and populations served by the different innovation models from quarterly reports Minnesota submits to CMS. We provide Minnesota's reported numbers in ***Appendix D***. Sources for provider and population data as of March 2017 are detailed in the Year Four Annual Report (RTI International, 2018). Counts of providers and populations reached as of December 2017 are state reported numbers (CMS, 2017). Denominators used to compute percentage of population reached are Kaiser Family Foundation population estimates based on the Census Bureau's March 2017 Current Population Survey (Kaiser Family Foundation, 2018).

D-2.3.2 State evaluation calls

We conducted monthly federal evaluation-specific calls beginning in April 2014 and continued through the end of the SIM Initiative test period. The RTI//NASHP evaluation team for Minnesota, the state officials leading Minnesota's SIM team, and the state's Innovation Center project officer typically attended each state evaluation call. The purpose of the calls was

to review interim evaluation findings with the state (as available), discuss any outstanding federal evaluation data or other needs, review and discuss state implementation and self-evaluation updates, and gather more in-depth information on select topics of interest for the evaluation.

For each meeting used to collect additional information and perspective from state officials leading the SIM Initiative in Minnesota, the evaluation team prepared a list of state-specific questions—including the status of related policy levers and implementation successes, challenges, and lessons learned. We first reviewed relevant state documents for answers to our questions. When we did not find answers in the document or needed further clarification, we sent the questions to the state ahead of the call and asked the state to have knowledgeable state officials available to answer the questions during the call.

D-2.3.3 Analysis

The RTI/NASHP evaluation team conducted thematic analysis of each source of qualitative data and then synthesized across information gleaned from site visit key informant interviews, site visit focus groups, document review, and state evaluation calls. Site visit interviews and focus groups were conducted in previous years of the evaluation. For more detail on site visit and focus group methods, see past evaluation reports.

D-2.4 References

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates?. *The Quarterly Journal of Economics*, *119*(1), 249–275.

Centers for Medicare & Medicaid Services (CMS). (2017). *Minnesota State Innovation Model progress report, Quarter 4 2017*. Supplied by CMS.

Kaiser Family Foundation. (2018). *Health insurance coverage of the total population*. Available at http://kff.org/other/state-indicator/total-population/

Puhani, P. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters, 115*, 85–87.

RTI International. (2018, March). *State Innovation Models (SIM) Initiative evaluation: Model Test Year four annual report.* Prepared for the Centers for Medicare & Medicaid Services (CMS). Retrieved from https://downloads.cms.gov/files/cmmi/sim-rd1-mt-fourthannrpt.pdf

Sub-appendix E-2. Methods for Oregon Analyses

E-2.1 Methods for the Impact Analysis of Oregon's Coordinated Care Model Implemented in State Health Employee Health Plans

To estimate the impact of the Coordinated Care Model (CCM) in Oregon, we conducted difference-in-differences (D-in-D) regression analyses using the Oregon All-Payer All-Claims (APAC) data and logistic regression analyses using consumer survey data collected by RTI. In ***Appendix E*, *Section E.2***, we present analyses for outcomes across four evaluation domains: (1) physician visits, (2) utilization and expenditures, (3) quality of care, and (4) patient perceptions. This sub-appendix details the methods we used for this analysis.

The CCM in the context of Oregon's health care delivery system. The CCM was developed originally as the model for coordinated care organizations (CCOs), the provider networks that have served Oregon's Medicaid population since 2012. Beginning in 2015, the CCM was implemented by the Public Employees Benefit Board (PEBB), which manages health and retirement plans used by employees of the state government and state universities. Beginning in the fall of 2017, the CCM was implemented by the Oregon Educators Benefit Board (OEBB), which manages the health and retirement plans used by public educators throughout the state. The key elements of the CCM include best practices to manage and coordinate care, shared responsibility for health, performance measurement, paying for outcomes and health, transparency and clear information, and maintaining costs at a sustainable rate of growth (Oregon Health Authority, n.d.).

Profiles of state employee plans after CCM implementation. The expansion of the CCM to state employees in 2015 was achieved through the procurement process for the health plans offered through PEBB. Each plan was required to incorporate elements of the CCM into the plan, including agreement to participate in the state's quality incentive program for member use of patient-centered primary care homes (PCPCHs). The result of the procurement process was the inclusion of two new health plans and the continuation of the four plans previously available to members (Statewide PPO, Providence Choice, Kaiser HMO, and Kaiser Deductible). The two new plans included a CCO (AllCare,[36] which serves Medicaid beneficiaries in the southwest corner of the state), and one commercial plan (Moda Synergy/Summit).

Although all plans are required to include some elements of the CCM, plan premiums, cost sharing, benefits, and coverage areas vary. Two plans (Statewide PPO, Kaiser HMO) are designated by PEBB as higher cost, and in addition to having a higher total premium than the other plans, members are required to contribute a larger share (5 vs. 1 percent) to those

[36] AllCare dropped out of the PEBB market in 2018.

premiums. Statewide PPO, the largest plan by membership, has the least restricted network and is available in all counties. All other plans have limited geographic availability.

Our analysis compares pre- and post-periods for the CCM (PEBB) treatment group and the (OEBB) comparison group using a longitudinal design with an unbalanced panel. This means we included beneficiaries who were first observed to be covered by a PEBB or OEBB health plan for at least 6 months in 2011 and included these individuals in the sample for any subsequent calendar year in which they were covered for at least 6 months. If an individual had coverage by both an OEBB and PEBB plan in a given year they were excluded from that year. As shown in the balance tables by year, average member characteristics do not differ substantially year to year within the PEBB or OEBB groups, suggesting that even though some beneficiaries may not have baseline data, the characteristics of the sample are not changing over time. We conducted separate analyses of children and adults because the recommended schedule of primary care visits, the types of specialty care required, and appropriate quality measures all vary by age.

Balancing CCM and comparison groups. For the claims-based analyses, using the enrollment files of the APAC for each year 2011–2016 we selected as the CCM treatment group persons who were enrolled in a PEBB plan for at least 6 months. For the comparison group, we selected persons who were enrolled in an OEBB plan for at least 6 months. We then estimated propensity score weights to balance the treatment and comparison groups on the available individual characteristics (age, sex, and condition diagnoses as summarized by a Chronic Illness and Disability Payment System risk score, and physician [primary and specialty care] utilization as measured in 2011), year of observation, and metropolitan statistical area (MSA) dummies to control for health care market characteristics. Propensity score weighting rather than matching was used to avoid dropping any beneficiaries from the final analysis. After propensity score weighting, the standardized differences between the weighted comparison group means and Medicaid SSP group means were all well under the standard 10 percent threshold.

Survey analyses used survey sampling weights to make the estimation sample representative of the whole PEBB and OEBB populations and controlled for age, sex, education, marital status, and self-rated health in logistic regression models. For more details see ***Section E-2.1.1.***

Study design. Our analysis uses a repeated cross-section design with an unbalanced sample of intervention and comparison beneficiaries. The samples are unbalanced because beneficiaries are only included in our models in periods when they are enrolled in a PEBB or OEBB plan for at least 6 months out of a calendar year. Some beneficiaries are therefore missing in parts of the pre- and post-period. We model intervention effects using the traditional D-in-D framework where estimates represent the pre-post difference in outcome trends among intervention beneficiaries in the absence of treatment.

For the survey analysis, we include separate samples of PEBB and OEBB members in early 2015 and late 2017. Survey questions ask about care received in the preceding 12 months. Thus, the baseline (2015) survey includes data on care received in the pre-implementation period and the 2017 survey collected data on care received in the post-implementation period for PEBB members.

Statistical approach. Claims analyses used D-in-D ordinary least squares models and allowed for clustered standard errors at the individual level to account for repeated observations of beneficiaries over time. Survey analyses used logistic regression to estimate differential changes between treatment and control groups on a variety of indicators of care coordination, access to care, and quality of care. More information on the regression model is available in ***Section E-2.5***.

E-2.1.1 Data sources

Oregon APAC data. The Oregon SIM evaluation team at the Urban Institute used medical claims and enrollment data covering calendar years 2011 to 2016 from Oregon's APAC database. The data used in this analysis covered two populations: (1) PEBB members and (2) OEBB members.

APAC data are collected and processed by Milliman, Inc. with oversight from the Oregon Health Authority (OHA). In addition to validating the data collected from submitters, Milliman applies its Health Cost Guidelines (HCG) Grouper, which sorts medical claims into benefit service categories.

Our APAC data files included the following files: (1) medical claims, (2) enrollment data, and (3) provider crosswalk.

Experience of Care Survey. The survey analysis is based on data from two rounds of the Experience of Care Survey of Oregon state employees and public educators. The first round was administered by RTI in early 2015 to measure perceptions of care in 2014, before the CCM was implemented for either of these populations. The second round was administered in the fall of 2017, 2.5 years after CCM-compliant plans began serving members of PEBB and less than 3 months after CCM-compliant plans began serving members of OEBB. The survey was adapted from the Patient Perceptions of Integrated Care Survey (Singer, Friedberg, Kiang, Dunn, & Kuhn, 2013). It contained questions asking respondents how they perceived the level of coordination and patient-centeredness of health care they received in the prior 12 months. It also included questions from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) Clinician & Group Survey concerning goals for care, opportunities for shared decision-making, having access to care after usual business hours, ease in getting appointments with specialists, and getting help from a provider in coordinating care across multiple providers. In addition, the survey asked respondents about health care status, recent health care use, the

kind of place they went to when they were sick or needed advice about their health care (henceforth referred to as the respondent's "usual source of care"), whether they typically saw a specific person at their usual source of care, the number of times they had visited their usual source of care in the past year, and how long they had been going to their usual source of care. Finally, the survey had questions about basic demographic characteristics (e.g., age, gender, educational attainment, race and ethnicity, and marital status) which were used as control variables in the model. A full description of the instrument and survey administration is available in the SIM Initiative Evaluation Year Three Annual Report (RTI International, 2017).

For the 2015 round of the survey, the State of Oregon provided RTI random samples of 24,000 state employees and 24,000 educators who had been continuously covered by one of health plans that the state offers these populations for the 12 months preceding September 1, 2014. The samples constituted just over half of all active members in each group. Sampled members with e-mail addresses received an e-mail invitation to participate in the web-based survey. The 5 percent of the sample without e-mail addresses were sent an invitation by U.S. Postal Service. The questionnaire was available and accepted responses between January and March 2015. We received 11,930 complete and partial surveys, a 27 percent response rate. Returned surveys were classified as either (1) ineligible, (2) complete, (3) partial, or (4) insufficient. Ineligible surveys included respondents who did not indicate the kind of place they go to most often to see a health care provider. In contrast, surveys classified as complete had valid responses for a section-specific minimum number of questions in each of the eight sections of the survey. Surveys assessed as partial met the response count for at least one of the eight sections, whereas insufficient surveys did not meet the minimum response count for any section. We further limited our analysis sample to respondents who reported going to their usual source of care at least once in the past 12 months. Our final sample totaled 9,981 individuals (5,309 state employees and 4,672 public educators).

For the 2017 round of the survey, the state provided RTI random samples of 25,700 state employees and 28,800 educators covered by PEBB and OEBB health plans, respectively. To avoid conflict with PEBB and OEBB open enrollment periods, the two samples were drawn and surveyed at different times. The PEBB sample included primary subscribers employed by the state of Oregon or university employees as of August 1, 2017, who had been continuously covered by a PEBB plan for at least 12 months. This sample was interviewed during late August and early September 2017. The OEBB sample included primary subscribers enrolled for at least 12 months as of September 1, 2017. The OEBB sample was surveyed during the month of November 2017. Using the same sample restrictions as in 2015, our final sample totaled 11,716 individuals (6,158 state employees and 5,558 public educators).

To produce estimates representative of the study population, we calculated weights to adjust for survey sampling design and nonresponse. The weights reflect the inverse selection probabilities and differential response rates for sample members. We expect weighted estimation

to reduce bias in the sample estimates at the cost of inflating estimated variances above what would be obtained from a simple random sample of the same size.

E-2.1.2 Outcome measures

Claims and enrollment data covering calendar years 2011–2016 were used to create 32 person-month level measures of expenditures, utilization, and quality of care. Several of these correspond closely with Oregon's performance metrics that are shared by CCOs and PEBB and OEBB plans. Specifications for each measure are provided in this section. For most measures, specifications were sourced from publicly available OHA performance metric specifications (Oregon Health Authority, n.d.). However, in some cases measures specifications were not available from OHA and in other cases they needed to be cross-referenced against other sources to ensure that they were generalizable to all-payer data. For these reasons we consulted several other sources of measure specifications published by payers and quality measurement organizations (CMS, 2016; "Measure #134", 2017; Molina Health Care, 2017; WellCare Health Plans, 2017).

Utilization and expenditures

Utilization and spending measures are reported as rates per person-year. For utilization measures, the numerator is the number of events among eligible individuals in the month. For both, the denominator is the number of eligible individuals during the month. Events are included in a month's total if admission or service date on the claim was during the period. Expenditures were defined as payments for claims with a status of "paid" or "managed care encounter." Averages include all individuals enrolled during the period, meaning that they reflect the presence of individuals with zero medical costs. Negative payments on claims were included in total expenditures because we were advised by OHA that summing across negative and positive payments would be approximately accurate in aggregate. Payments were included in a month's total if the admission or service date on the claim occurred during the month.

- **Total spending:** Payments were summed across all inpatient and outpatient (facility and professional) claims. This measure excludes member cost-sharing and pharmacy component expenditures.
- **Primary care visits:** Visits to primary care providers (PCPs) were selected based on Milliman's HCG Grouper. A primary care visit was counted when a claim line had one of the following HCG codes:
  - P32c: Professional Office/Home Visits—PCP
  - P42: Professional Preventive Well Baby Exams
  - P43: Professional Preventive Physical Exams

  Multiple primary care visits on a single day were counted as a single visit.

- **Specialist visits:** Visits to specialty providers were counted when the claim line had an HCG code of "P32d" (Professional Office/Home Visit—Specialist). Multiple specialty provider visits on a single day were counted as a single visit.
- **ED utilization:** ED visits (including observation stays) were counted when a claim line had a revenue code of 0450, 0451, 0452, 0456, 0459, or 0981. Multiple ED visits on a single day were counted as a single visit.
- **Inpatient hospitalizations:** Inpatient hospitalizations were identified with the following HCG codes:
  - I11a: Facility Inpatient Medical—General
  - I12: Facility Inpatient—Surgical
  - I13a: Facility Inpatient Psychiatric—Hospital
  - I14a: Facility Inpatient Alcohol and Drug Abuse—Hospital
  - I21a: Facility Inpatient Maternal Normal Delivery
  - I21b: Facility Inpatient Maternal Normal Delivery—Mom\Baby Combined
  - I22a: Facility Inpatient Maternal C-Section Delivery
  - I22b: Facility Inpatient Maternal C-Section Delivery—Mom\Baby Combined
  - I23: Facility Inpatient Well Newborn
  - I24: Facility Inpatient Other Newborn
  - I25: Facility Inpatient Maternity Non-Delivery

We did not include facility inpatient claims with HCG codes for rehabilitation facility, psychiatric residential facility, alcohol and drug abuse residential facility, and skilled nursing facility.

Some claims that appear to indicate multiple admissions are in fact transfers between facilities. These records were counted as a single admission. To combine transfers into one acute admission, we identified claims that had no more than 1 elapsed day between discharge date of the index claim and admission date of the subsequent claim. We combined the claims into one record by taking the earliest admission date and latest discharge date.

Quality of care

To evaluate the impact on quality of care, we report the following quality measures. Measures were calculated on a PMPM basis, and we calculated both the probability of the event occurring in a person-month and the number of events per 1,000 person-months. Inpatient discharges were calculated on a per admission per month basis.

- **Thirty-day readmissions per hospital discharge:** This is the total number of unplanned hospital readmissions within 30 days of discharge, divided by the total number of index admissions in the month. An index hospital discharge is identified as

an inpatient stay with a discharge date within the given month. We excluded admissions if the beneficiary died during the hospitalization.

- Depression screenings (ages 12 and older): Depression screenings were counted when a visit included a claim line with the Healthcare Common Procedure Coding System (HCPCS) code G8431 or G8510. The denominator for this measure included enrollment months for individuals at least 12 years old as of December 31 of the measurement year that had at least one eligible outpatient encounter during the month as indicated by any of Current Procedural Terminology (CPT) codes 90791, 90792, 90832, 90834, 90837, 90839, 92625, 96116, 96118, 96150, 96151, 97003, 99201, 99202, 99203, 99204, 99205, 99212, 99213, 99214, or 99215; or HCPCS codes G0101, G0402, G0438, G0439, or G0444 and without telehealth modifiers GQ or GT.
- Screening, Brief Intervention and Referral to Treatment (SBIRT) screenings (ages 18 and older): SBIRT screenings were counted when a visit included a claim line with CPT code 99408 or 99409, HCPCS code G0443, G0396, or G0397, or diagnosis code V82.9. We also included visits that included both CPT code 99420 and ICD-9 diagnosis code V82.9. The denominator for this measure included enrollment months for individuals at least 18 years old as of December 31 of the measurement year and had at least one eligible outpatient service as indicated by CPT codes 99201–99205, 99211–99215, 99241–99245, 99341–99345, 99347–99350, 99383–99384, 99385–99387, 99393–99394, 99395–99397, 99401–99404, 99408, 99409, 99411, 99412, 99420, or 99429; HCPCS codes G0396, G0397, G0402, G0442, G0443, or T1015, or diagnosis code V20.2.
- **Cervical cancer screenings (females ages 24–64):** Cervical cancer screenings were counted when a visit included a claim line with CPT code 88141, 88142, 88143, 88147, 88148, 88150, 88152, 88153, 88154, 88164, 88165, 88166, 88167, 88174, or 88175; or HCPCS code G0123, G0124, G0141, G0143, G0144, G0145, G0147, G0148, P3000, P3001, or Q0091; or revenue code 0923. The denominator for this measure included enrollment months for women ages 24–64 as of December 31 of the measurement year.
- For the purposes of determining the population eligible for inclusion in depression and SBIRT screenings, inpatient and outpatient encounters were identified with the HCG codes in ***Table E-2-1***.

Table E-2-1. Codes for identifying inpatient and outpatient visits

Description	HCG code
Inpatient visit	I11A, P31a, P31b
Outpatient visit	P32c, P32d, P33, P43, P40a, P51, P52, O11a, O11b

HCG = Health Cost Guidelines.

Survey-based measures

For the survey analysis, we grouped the survey's measures of patient experience over the past 12 months into four domains: (1) overall quality and manageability, (2) care coordination, (3) patient-centeredness, and (4) accessibility. Overall quality and manageability is the respondent's perception of the quality of care they received in the past 12 months as a whole and how easy it was to manage that care. Care coordination is the coordination of care across providers and settings. Patient centeredness, as described by the Agency for Healthcare Research and Quality (AHRQ, 2016) is the orientation of care toward the whole person. This domain includes measures of whether the respondent's primary care provider(s) had comprehensive knowledge of the respondent's health care needs and whether their interactions with the respondent were respectful and attentive to the respondents' goals. Finally, accessibility is the respondent's ability to access primary care after office hours and make appointments with specialists.

All 16 measures used in the analysis were dichotomous patient experience outcomes. Each measure was constructed from a single survey item. In many cases, we collapsed survey responses to construct a binary measure. For example, respondents rated the quality of all the care they received in the past 12 months and the ease of managing that care on a scale of 1 to 10. We classified responses of 7 to 10 as high quality and easily managed care, respectively. Many other questions had a four-point response scale of "never," "sometimes," "usually," and "always," which we collapsed such that "never" and "sometimes" were coded as a negative response (given a score of 0) and "usually" and "always" were coded as a positive response (given a score of 1). Each measure was then used as the dependent variable in a separate logistic regression.

E-2.1.3 Population studied

In the report, for claims analyses the PEBB and OEBB populations were defined as those enrolled in a PEBB or OEBB plan for at least 6 out of 12 months in a calendar year, 2011–2016. We also limited the sample to individuals first observed in 2011. We estimated outcome models separately for two age groups:

- Adults (age > 18)
- Children (age <= 18)

For the survey analyses, PEBB and OEBB populations were those responding to the survey of adult subscribers (not spouses or children).

E-2.1.4 Comparison group and propensity score weighting

For the APAC analysis, we are using a pre-post comparison group design in which the comparison group provides an estimate of what would have happened among PEBB members absent the CCM implementation. The difference in the changes over time from the pre-period to

the intervention period between CCM (PEBB) beneficiaries and their comparison group provides an estimate of the impact of the CCM. The comparison group should be similar to the CCM group on all relevant dimensions (e.g., demographic, regulatory, health status, and health systems) except for the policy change being tested.

In the following section, we detail the procedures we used to select the comparison group for the CCM treatment group in Oregon.

Selection of comparison group

Because CCM requirements were not implemented in the OEBB population until the fall of 2017, its population of public sector employees was selected as a natural comparison for the state government employees who experienced the CCM beginning in 2015.

Calculation of person-level weights

To balance the population characteristics for the claims-based analyses, we estimated propensity scores for all individuals from the comparison group. A propensity score is the probability that an individual is in the CCM treatment group rather than the comparison group.

The objective of propensity score modeling is to create a weighted comparison group with characteristics equivalent to those for the CCM population. To the extent that these characteristics are correlated with expenditure, utilization, and quality outcomes, propensity weighting will help balance pre-CCM levels of the outcomes as well.

There are other methods to apply propensity scores to an analysis. Aside from weighting, one frequently used method is matching, whereby a CCM beneficiary is matched to a comparison group beneficiary who has a similar propensity score. Although we considered this method, we decided not to pursue matching for several reasons. First, propensity score weighting has been shown to produce less biased estimates, less modeling error (e.g., mean squared error, type 1 error), and more accurate variance estimation and confidence intervals when modeling dichotomous outcomes, and this analysis includes many dichotomous utilization and quality of care outcomes. Second, matching may exclude many comparison group beneficiaries from the analysis if a good match cannot be found. Weighting has the additional advantage of preserving sample size.

Person-level characteristics

The initial step in the process was to select person-level characteristics to be used in each propensity score model. ***Table E-2-2*** shows the characteristics we used grouped by whether they control for demographics, enrollment, beneficiary health status, geography, or year.

Table E-2-2. Covariates for propensity score logistic regressions

Demographic characteristics
Gender
Age
Gender*Age interaction
Enrollment
Number of months enrolled in plan during current year (dummies for 7, 8, 9, 10, 11, 12 months)
Health status
CDPS
CDPS*Gender interaction
CDPS*Age interaction
CDPS*Age*Gender interaction
Primary care visits in 2011
Specialist visits in 2011
Geography
Dummies for residence MSA: Bend (reference), Corvallis, Eugene, Medford, Portland, Salem, rest of state
Year
Dummies for 2012–2016

CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

Estimation and weighting procedures

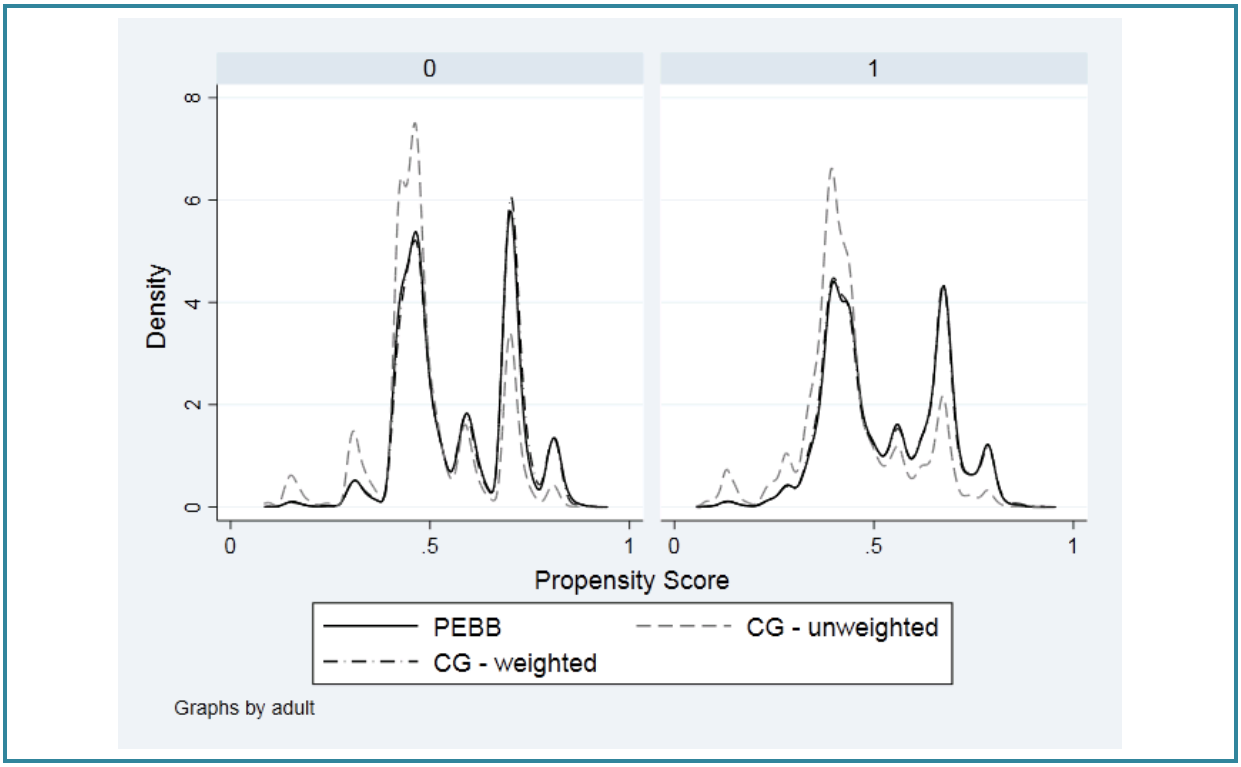
Using the characteristics listed in Table E-2-2, we estimated propensity models by logistic regression, in which the outcome was 1 for beneficiaries in a PEBB (CCM) health plan and 0 for the comparison group. Separate models were estimated for adults and children.

We set propensity weights to 1 for all individuals in the CCM group. The propensity weight for a comparison individual was a function of his or her predicted propensity score—where weight = p/(1-p), with p the predicted propensity. Our procedure typically includes trimming weights that are either less than 0.05 or greater than 20, although in this analysis no weights needed trimming.

F-2.1.5 Propensity model evaluation

We evaluated several aspects of the propensity score models. First, we examined plots of predicted probabilities to ensure sufficient overlap in the distributions of the CCM and comparison groups. This feature, known as common support, is critical because it provides the basis for inferring effects from group comparisons (***Figure E-2-1***).

Figure E-2-1. Weighted and unweighted propensity score density plots for the PEBB and comparison groups



CG = comparison group; PEBB = Public Employees Benefit Board.

In all years, we found that the comparison group passed the common support assumption ($P(D = 1|X)>0$) for the entire range of the CCM group's propensity scores. Weighting the comparison group eliminated any visible difference in the distribution of the CCM and comparison group.

Second, we compared the logistic results of the models to see which variables had the greatest impact on the propensity score weights. We found that the major differences between the groups were in their MSA of residence. Overall, we found that CCM (PEBB) beneficiaries were much more likely to live in the Corvallis, Eugene, and Salem areas and less likely to live in Bend, Medford, Portland, or the rest of the state.

Finally, we compared unweighted and propensity-weighted means for the characteristics in the model. As expected, we found that, after weighting, the comparison group means were well within the 10 percent threshold for standardized difference.

Tables E-2-3 to ***E-2-14*** show unweighted and (propensity score) weighted means/proportions for 2011–2016. The only notable group differences in the unweighted samples—geographic distribution—are substantially mitigated post-weighting as evidenced by the minimized standardized differences.

Table E-2-3. Unweighted and weighted means and standardized differences, CCM and comparison groups, adults, 2011

	Unweighted			Weighted	
Characteristic	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.7	11.6	−3.1	11.7	3.1
Age	43.7	44.8	5.4	43.9	1.0
Female	0.5	0.6	3.1	0.5	−0.7
CDPS score	0.2	0.2	−2.3	0.2	0.6
MSA of Residence					
Bend	0.02	0.06	12.2	0.02	−0.7
Corvallis	0.06	0.02	−16.0	0.07	1.8
Eugene	0.11	0.09	−5.3	0.11	0.5
Medford	0.03	0.04	0.6	0.03	0.3
Portland	0.25	0.37	18.7	0.25	−0.8
Salem	0.29	0.14	−26.9	0.29	−0.5
Rest of state	0.23	0.29	10.2	0.23	0.1

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

Table E-2-4. Unweighted and weighted means and standardized differences, CCM and comparison groups, adults, 2012

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.8	11.7	−5.9	11.8	0.1
Age	43.8	45.0	5.8	44.4	2.6
Female	0.5	0.6	3.4	0.5	−0.5
CDPS score	0.2	0.2	−1.4	0.2	1.1
MSA of Residence					
Bend	0.02	0.06	12.5	0.02	−0.5
Corvallis	0.06	0.02	−16.5	0.07	0.9
Eugene	0.11	0.09	−5.9	0.11	0.2
Medford	0.03	0.04	0.3	0.03	−0.1
Portland	0.24	0.37	19.6	0.24	−0.2
Salem	0.29	0.14	−27.1	0.29	−0.5
Rest of state	0.23	0.29	10.1	0.23	0.2

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

Table E-2-5. Unweighted and weighted means and standardized differences, CCM and comparison groups, adults, 2013

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.8	11.7	−7.4	11.8	−0.9
Age	44.3	45.5	5.7	44.8	2.4
Female	0.5	0.6	3.6	0.5	−0.3
CDPS score	0.2	0.2	−2.3	0.2	−0.2
MSA of Residence					
Bend	0.03	0.06	12.3	0.02	−0.6
Corvallis	0.07	0.02	−16.9	0.07	0.4
Eugene	0.11	0.09	−5.9	0.11	0.2
Medford	0.03	0.03	0.1	0.03	−0.2
Portland	0.24	0.37	20.3	0.24	0.4
Salem	0.29	0.14	−27.5	0.29	−0.8
Rest of state	0.23	0.29	10.2	0.23	0.4

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

Table E-2-6. Unweighted and weighted means and standardized differences, CCM and comparison groups, adults, 2014

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.8	11.7	−7.5	11.8	−2.5
Age	44.6	45.9	6.4	45.1	2.5
Female	0.5	0.6	4.0	0.5	−0.2
CDPS score	0.3	0.2	−2.1	0.3	−0.3
MSA of Residence					
Bend	0.02	0.06	13.0	0.02	−0.1
Corvallis	0.07	0.02	−17.2	0.07	0.0
Eugene	0.11	0.08	−6.3	0.11	−0.3
Medford	0.03	0.03	0.4	0.03	0.0
Portland	0.24	0.37	20.3	0.24	0.5
Salem	0.30	0.14	−28.1	0.29	−1.3
Rest of state	0.23	0.29	10.9	0.23	1.2

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

Table E-2-7. Unweighted and weighted means and standardized differences, CCM and comparison groups, adults, 2015

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.8	11.7	−9.5	11.8	−4.1
Age	44.9	46.2	6.3	45.3	1.8
Female	0.5	0.6	4.5	0.5	0.4
CDPS score	0.3	0.2	−2.2	0.3	−0.4
MSA of Residence					
Bend	0.02	0.06	12.8	0.02	−0.2
Corvallis	0.07	0.02	−17.6	0.06	−1.3
Eugene	0.11	0.08	−6.3	0.11	−0.3
Medford	0.03	0.03	0.0	0.03	−0.4
Portland	0.23	0.37	20.8	0.24	0.8
Salem	0.30	0.14	−28.4	0.30	−1.0
Rest of state	0.23	0.30	11.1	0.23	1.4

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

Table E-2-8. Unweighted and weighted means and standardized differences, CCM and comparison groups, adults, 2016

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.8	11.7	−8.5	11.8	−2.6
Age	44.9	46.3	6.2	45.3	1.4
Female	0.5	0.6	5.1	0.5	1.1
CDPS score	0.2	0.2	−2.8	0.2	−1.4
MSA of Residence					
Bend	0.02	0.06	13.6	0.03	0.5
Corvallis	0.07	0.02	−17.8	0.06	−1.2
Eugene	0.11	0.08	−6.5	0.11	−0.3
Medford	0.03	0.03	0.1	0.03	−0.2
Portland	0.23	0.37	20.7	0.24	0.7
Salem	0.31	0.14	−28.4	0.30	−1.1
Rest of state	0.23	0.30	11.2	0.24	1.4

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

Table E-2-9. Unweighted and weighted means and standardized differences, CCM and comparison groups, children, 2011

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.6	11.6	−2.3	11.7	2.7
Age	9.9	10.2	3.8	9.9	0.5
Female	0.5	0.5	0.1	0.5	1.0
CDPS score	0.4	0.4	−2.7	0.4	1.4
MSA of Residence					
Bend	0.03	0.07	13.9	0.03	0.2
Corvallis	0.06	0.02	−15.6	0.06	0.8
Eugene	0.10	0.08	−5.2	0.10	−0.2
Medford	0.03	0.03	0.8	0.03	0.2
Portland	0.24	0.37	18.9	0.24	−0.8
Salem	0.30	0.16	−24.0	0.33	3.9
Rest of state	0.24	0.28	6.3	0.22	−4.0

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

Table E-2-10. Unweighted and weighted means and standardized differences, CCM and comparison groups, children, 2012

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.8	11.8	−4.8	11.9	3.0
Age	10.1	10.4	2.9	10.4	3.1
Female	0.5	0.5	0.0	0.5	0.5
CDPS score	0.4	0.4	−2.2	0.4	0.2
MSA of Residence					
Bend	0.03	0.07	13.9	0.03	0.0
Corvallis	0.06	0.02	−16.6	0.06	−1.1
Eugene	0.10	0.08	−6.4	0.10	−0.7
Medford	0.03	0.03	1.3	0.03	0.4
Portland	0.24	0.36	19.3	0.24	−0.5
Salem	0.30	0.16	−23.7	0.33	5.0
Rest of state	0.24	0.28	6.3	0.22	−4.1

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

Table E-2-11. Unweighted and weighted means and standardized differences, CCM and comparison groups, children, 2013

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.8	11.8	−6.2	11.9	3.0
Age	10.6	10.8	2.1	10.8	2.6
Female	0.5	0.5	−0.4	0.5	−0.1
CDPS score	0.4	0.4	−2.1	0.4	−0.4
MSA of Residence					
Bend	0.03	0.07	14.6	0.03	0.5
Corvallis	0.06	0.01	−17.5	0.05	−2.6
Eugene	0.10	0.08	−6.5	0.10	−1.1
Medford	0.03	0.03	1.0	0.03	0.2
Portland	0.24	0.36	19.5	0.24	0.1
Salem	0.29	0.16	−23.6	0.33	5.1
Rest of state	0.25	0.28	6.1	0.22	−3.8

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

Table E-2-12. Unweighted and weighted means and standardized differences, CCM and comparison groups, children, 2014

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.9	11.8	−3.7	11.9	1.6
Age	11.1	11.1	1.3	11.2	1.6
Female	0.5	0.5	−0.3	0.5	0.1
CDPS score	0.4	0.4	−2.4	0.4	−0.2
MSA of Residence					
Bend	0.03	0.07	15.2	0.03	0.7
Corvallis	0.06	0.01	−18.1	0.05	−2.9
Eugene	0.10	0.08	−6.6	0.10	−1.2
Medford	0.03	0.03	0.4	0.03	−0.5
Portland	0.23	0.36	20.0	0.24	0.5
Salem	0.30	0.16	−24.1	0.33	4.8
Rest of state	0.24	0.28	6.4	0.22	−3.3

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.).

Table E-2-13. Unweighted and weighted means and standardized differences, CCM and comparison groups, children, 2015

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.9	11.8	−4.8	11.9	1.5
Age	11.5	11.5	0.5	11.5	0.3
Female	0.5	0.5	−0.3	0.5	−0.3
CDPS score	0.4	0.4	−1.4	0.4	0.5
MSA of Residence					
Bend	0.03	0.08	15.8	0.03	1.1
Corvallis	0.06	0.01	−18.8	0.05	−4.4
Eugene	0.10	0.07	−7.5	0.10	−1.8
Medford	0.03	0.03	−0.2	0.03	−0.8
Portland	0.23	0.36	20.4	0.24	1.0
Salem	0.30	0.16	−24.2	0.33	4.8
Rest of state	0.24	0.29	6.9	0.23	−2.8

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

Table E-2-14. Unweighted and weighted means and standardized differences, CCM and comparison groups, children, 2016

Characteristic	Unweighted			Weighted	
	CCM group	Comparison group	Standardized difference[a]	Comparison group	Standardized difference[a]
Months enrolled	11.8	11.8	−3.0	11.9	5.8
Age	11.6	11.6	−0.2	11.6	−0.6
Female	0.5	0.5	−0.6	0.5	−0.7
CDPS score	0.4	0.3	−3.1	0.3	−2.1
MSA of Residence					
Bend	0.03	0.08	15.9	0.03	1.3
Corvallis	0.07	0.01	−19.2	0.05	−4.5
Eugene	0.10	0.07	−8.1	0.09	−2.2
Medford	0.03	0.03	−0.3	0.03	−1.0
Portland	0.23	0.36	20.3	0.24	0.9
Salem	0.30	0.16	−23.8	0.33	5.0
Rest of state	0.24	0.29	7.1	0.23	−2.5

CCM = Coordinated Care Model; CDPS = Chronic Illness and Disability Payment System (CDPS score is a risk-adjustment score calculated from ICD9 and ICD10 diagnosis codes included on hospital and outpatient claims, with larger CDPS scores corresponding to a larger number of comorbidities or a more severe set of comorbidities); MSA = Metropolitan Statistical Area.

[a] Absolute standardized differences are expressed as percentages.
Note: CCM group = PEBB. CG = OEBB.

E-2.1.6 Descriptive analyses

APAC claims analysis

Tables E-2-15 through ***E-2-26*** present, for adults and children, respectively, annual descriptive statistics on average monthly utilization and expenditures, and quality of care.

Table E-2-15. Summary statistics for PEBB and OEBB populations, adults (18+), 2011

Characteristic	CCM group (PEBB)	Comparison group (OEBB) unweighted	Comparison group weighted
Percent with spending in year	88.4%	81.7%	82.1%*
Total spending per month enrolled	$402	$346	$391
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$364	$311	$343
Percent with primary care visit	75.1%	71.5%	73.2%
Number of primary care visits	2.37	2.04	2.30
Percent with specialist visit	50.0%	46.8%	47.7%
Number of specialist visits	1.60	1.49	1.64
Percent with inpatient stay	3.7%	3.8%	4.0%
Number of inpatient stays	0.04	0.04	0.05
Percent with emergency department visit	12.1%	10.0%	10.8%
Number of emergency department visits	0.18	0.14	0.16
Percent with 30-day readmissions, among those with discharge	6.8%	4.9%	5.6%
Number of 30-day readmissions, per 1,000 discharges	92.43	64.26	80.41
Percent with SBIRT screening, ages 18+ and 1+ outpatient encounter	0.0%	0.0%	0.0%
Number of SBIRT screenings, ages 18+	0.00	0.00	0.00
Percent with cervical cancer screening, women ages 24–64	39.2%	38.9%	39.9%
Number of cervical cancer screenings, women ages 24–64	0.41	0.41	0.42
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.00%	0.00%	0.00%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database

Notes: Ages 18+. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10.

Table E-2-16. Summary statistics for PEBB and OEBB populations, adults (18+), 2012

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	86.5%	80.4%	80.8%*
Total spending per month enrolled	$380	$352	$375
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$336	$314	$333
Percent with primary care visit	74.7%	70.9%	72.3%
Number of primary care visits	2.30	2.01	2.18
Percent with specialist visit	47.7%	46.0%	46.4%
Number of specialist visits	1.49	1.46	1.54
Percent with inpatient stay	3.6%	3.6%	3.7%
Number of inpatient stays	0.04	0.04	0.04
Percent with emergency department visit	11.8%	10.2%	10.8%
Number of emergency department visits	0.17	0.14	0.15
Percent with 30-day readmissions, among those with discharge	5.9%	4.4%	4.7%
Number of 30-day readmissions, per 1,000 discharges	77.10	60.88	72.84
Percent with SBIRT screening, ages 18+ and 1+ outpatient encounter	0.0%	0.0%	0.0%
Number of SBIRT screenings, ages 18+	0.00	0.00	0.00
Percent with cervical cancer screening, women ages 24–64	34.7%	34.4%	34.5%
Number of cervical cancer screenings, women ages 24–64	0.37	0.36	0.36
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.00%	0.00%	0.00%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 18+. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10.

Table E-2-17. Summary statistics for PEBB and OEBB populations, adults (18+), 2013

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	86.4%	80.4%	80.8%*
Total spending per month enrolled	$399	$362	$374
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$354	$310	$320
Percent with primary care visit	71.4%	70.7%	72.2%
Number of primary care visits	2.21	1.98	2.15
Percent with specialist visit	50.2%	46.5%	46.5%
Number of specialist visits	1.61	1.37	1.42
Percent with inpatient stay	3.5%	3.7%	3.7%
Number of inpatient stays	0.04	0.04	0.04
Percent with emergency department visit	11.8%	9.7%	10.1%
Number of emergency department visits	0.17	0.13	0.14
Percent with 30-day readmissions, among those with discharge	6.0%	4.9%	5.5%
Number of 30-day readmissions, per 1,000 discharges	89.80	67.01	81.63
Percent with SBIRT screening, ages 18+ and 1+ outpatient encounter	1.3%	0.0%	0.1%*
Number of SBIRT screenings, ages 18+	0.01	0.00	0.00*
Percent with cervical cancer screening, women ages 24–64	27.8%	30.7%	30.6%
Number of cervical cancer screenings, women ages 24–64	0.29	0.32	0.32
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.00%	0.00%	0.00%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 18+. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10.

Table E-2-18. Summary statistics for PEBB and OEBB populations, adults (18+), 2014

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	86.9%	85.8%	85.9%
Total spending per month enrolled	$430	$382	$402
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$383	$337	$348
Percent with primary care visit	76.5%	72.4%	73.2%
Number of primary care visits	2.42	2.05	2.19
Percent with specialist visit	47.5%	44.6%	44.6%
Number of specialist visits	1.50	1.30	1.35
Percent with inpatient stay	3.5%	3.8%	4.0%
Number of inpatient stays	0.04	0.05	0.05
Percent with emergency department visit	12.4%	10.4%	10.7%
Number of emergency department visits	0.18	0.15	0.16
Percent with 30-day readmissions, among those with discharge	5.6%	6.2%	6.9%
Number of 30-day readmissions, per 1,000 discharges	81.55	79.52	90.21
Percent with SBIRT screening, ages 18+ and 1+ outpatient encounter	3.3%	2.1%	2.8%
Number of SBIRT screenings, ages 18+	0.03	0.02	0.03
Percent with cervical cancer screening, women ages 24–64	25.3%	25.3%	24.8%
Number of cervical cancer screenings, women ages 24–64	0.27	0.27	0.26
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.01%	0.02%	0.02%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 18+. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10 (not applicable in this table).

Table E-2-19. Summary statistics for PEBB and OEBB populations, adults (18+), 2015

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	86.8%	86.2%	86.3%
Total spending per month enrolled	$456	$392	$404
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$410	$343	$352
Percent with primary care visit	76.6%	72.7%	73.6%
Number of primary care visits	2.48	2.10	2.23
Percent with specialist visit	47.6%	44.3%	44.4%
Number of specialist visits	1.50	1.28	1.32
Percent with inpatient stay	3.3%	3.7%	3.8%
Number of inpatient stays	0.04	0.04	0.04
Percent with emergency department visit	12.7%	10.5%	11.0%
Number of emergency department visits	0.18	0.15	0.16
Percent with 30-day readmissions, among those with discharge	6.5%	6.1%	6.3%
Number of 30-day readmissions, per 1,000 discharges	87.59	92.05	102.09
Percent with SBIRT screening, ages 18+ and 1+ outpatient encounter	6.2%	3.7%	5.0%
Number of SBIRT screenings, ages 18+	0.07	0.04	0.05
Percent with cervical cancer screening, women ages 24–64	23.4%	23.6%	23.9%
Number of cervical cancer screenings, women ages 24–64	0.24	0.25	0.25
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.07%	0.11%	0.09%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 18+. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10 (not applicable in this table).

Table E-2-20. Summary statistics for PEBB and OEBB populations, adults (18+), 2016

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	86.1%	85.6%	85.6%
Total spending per month enrolled	$482	$411	$413
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$436	$357	$359
Percent with primary care visit	76.5%	72.7%	73.2%
Number of primary care visits	2.50	2.12	2.24
Percent with specialist visit	47.2%	43.9%	43.9%
Number of specialist visits	1.46	1.25	1.28
Percent with inpatient stay	3.4%	3.5%	3.6%
Number of inpatient stays	0.04	0.04	0.04
Percent with emergency department visit	13.1%	10.8%	11.0%
Number of emergency department visits	0.19	0.16	0.16
Percent with 30-day readmissions, among those with discharge	6.0%	6.7%	7.4%
Number of 30-day readmissions, per 1,000 discharges	75.77	90.22	100.15
Percent with SBIRT screening, ages 18+ and 1+ outpatient encounter	7.8%	6.9%	7.7%
Number of SBIRT screenings, ages 18+	0.08	0.07	0.08
Percent with cervical cancer screening, women ages 24–64	22.1%	23.0%	22.9%
Number of cervical cancer screenings, women ages 24–64	0.23	0.24	0.24
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.05%	0.07%	0.05%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 18+. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10 (not applicable in this table).

Table E-2-21. Summary statistics for PEBB and OEBB populations, children (0–17), 2011

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	85.2%	78.1%	79.8%*
Total spending per month enrolled	$144	$134	$160
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$127	$123	$138
Percent with primary care visit	71.8%	74.3%	76.7%
Number of primary care visits	2.11	2.09	2.35
Percent with specialist visit	25.6%	23.7%	25.0%
Number of specialist visits	0.55	0.53	0.59
Percent with inpatient stay	1.3%	1.0%	1.2%
Number of inpatient stays	0.02	0.01	0.02
Percent with emergency department visit	12.9%	10.5%	11.1%
Number of emergency department visits	0.17	0.13	0.14
Percent with SBIRT screening, ages 12–17 and 1+ outpatient encounter	0.0%	0.0%	0.0%
Number of SBIRT screenings, ages 12–17	0.00	0.00	0.00
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.00%	0.00%	0.00%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 0–17. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10.

Table E-2-22. Summary statistics for PEBB and OEBB populations, children (0–17), 2012

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	83.8%	75.9%	76.5%*
Total spending per month enrolled	$124	$127	$139
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$112	$117	$123
Percent with primary care visit	71.1%	74.3%	75.3%
Number of primary care visits	1.93	1.95	2.01
Percent with specialist visit	23.5%	22.1%	22.7%
Number of specialist visits	0.50	0.50	0.54
Percent with inpatient stay	1.1%	1.0%	1.1%
Number of inpatient stays	0.01	0.01	0.01
Percent with emergency department visit	11.8%	10.2%	10.4%
Number of emergency department visits	0.15	0.13	0.13
Percent with SBIRT screening, ages 12–17 and 1+ outpatient encounter	0.0%	0.0%	0.0%
Number of SBIRT screenings, ages 12–17	0.00	0.00	0.00
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.00%	0.00%	0.00%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 0–17. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10.

Table E-2-23. Summary statistics for PEBB and OEBB populations, children (0–17), 2013

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	83.7%	76.5%	77.3%*
Total spending per month enrolled	$117	$121	$124
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$108	$103	$108
Percent with primary care visit	70.5%	73.3%	74.3%
Number of primary care visits	1.76	1.75	1.81
Percent with specialist visit	27.8%	23.9%	24.5%
Number of specialist visits	0.62	0.51	0.55
Percent with inpatient stay	1.0%	1.0%	1.1%
Number of inpatient stays	0.01	0.01	0.01
Percent with emergency department visit	11.7%	10.0%	9.8%
Number of emergency department visits	0.15	0.12	0.12
Percent with SBIRT screening, ages 12–17 and 1+ outpatient encounter	1.8%	0.0%	0.0%*
Number of SBIRT screenings, ages 12–17	0.02	0.00	0.00*
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.01%	0.00%	0.00%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 0–17. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10.

Table E-2-24. Summary statistics for PEBB and OEBB populations, children (0–17), 2014

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	84.4%	83.8%	84.1%
Total spending per month enrolled	$124	$116	$115
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$115	$108	$108
Percent with primary care visit	76.0%	73.4%	74.1%
Number of primary care visits	1.94	1.69	1.76
Percent with specialist visit	24.9%	23.0%	22.9%
Number of specialist visits	0.54	0.50	0.51
Percent with inpatient stay	0.9%	0.9%	0.9%
Number of inpatient stays	0.01	0.01	0.01
Percent with emergency department visit	11.3%	9.7%	9.7%
Number of emergency department visits	0.15	0.12	0.12
Percent with SBIRT screening, ages 12–17 and 1+ outpatient encounter	1.3%	0.3%	0.2%
Number of SBIRT screenings, ages 12–17	0.01	0.00	0.00
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.00%	0.00%	0.00%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 0–17. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10 (not applicable in this table).

Table E-2-25. Summary statistics for PEBB and OEBB populations, children (0–17), 2015

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	84.7%	84.3%	84.4%
Total spending per month enrolled	$133	$128	$132
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$127	$119	$122
Percent with primary care visit	76.9%	74.6%	75.2%
Number of primary care visits	1.98	1.76	1.83
Percent with specialist visit	25.2%	23.3%	23.6%
Number of specialist visits	0.56	0.53	0.53
Percent with inpatient stay	0.8%	1.0%	1.0%
Number of inpatient stays	0.01	0.02	0.02
Percent with emergency department visit	11.5%	9.7%	9.2%
Number of emergency department visits	0.15	0.12	0.12
Percent with SBIRT screening, ages 12–17 and 1+ outpatient encounter	3.7%	8.4%	7.7%*
Number of SBIRT screenings, ages 12–17	0.04	0.08	0.08*
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.07%	0.04%	0.03%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 0–17. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10.

Table E-2-26. Summary statistics for PEBB and OEBB populations, children (0–17), 2016

Characteristic	CCM group	Comparison group unweighted	Comparison group weighted
Percent with spending in year	83.5%	84.1%	83.6%
Total spending per month enrolled	$144	$122	$125
Total spending per month enrolled (bottom/top coded at $0/99.5th percentile)	$139	$115	$117
Percent with primary care visit	76.3%	74.3%	74.7%
Number of primary care visits	1.95	1.71	1.76
Percent with specialist visit	25.4%	22.5%	22.1%
Number of specialist visits	0.57	0.50	0.50
Percent with inpatient stay	1.0%	0.9%	0.9%
Number of inpatient stays	0.02	0.01	0.01
Percent with emergency department visit	10.7%	8.8%	8.9%
Number of emergency department visits	0.15	0.12	0.11
Percent with SBIRT screening, ages 12–17 and 1+ outpatient encounter	0.0%	18.5%	14.1%*
Number of SBIRT screenings, ages 12–17	0.00	0.19	0.14*
Percent with depression screening, ages 12+ and 1+ outpatient encounter	0.14%	0.09%	0.10%
Number of depression screenings, ages 12+	0.00	0.00	0.00

CCM = Coordinated Care Model; OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Source: All Payer All Claims database.

Notes: Ages 0–17. Samples exclude individuals with both PEBB and OEBB insurance in a given year. CCM group = PEBB. CG = OEBB.

*Standardized difference between weighted comparison group mean and CCM group greater than 0.10.

Consumer survey analysis

Table E-2-27 provides descriptive statistics from the two rounds of consumer surveys on the demographic, socioeconomic, and health care characteristics of respondents in each of the four analysis groups. The only notable differences between the state employee and educator samples is that the OEBB sample has higher fractions of female and college educated.

Table E-2-27. Individual characteristics, by group (PEBB/OEBB), by year (2015/2017)

Characteristic	PEBB, 2015 (N=5,309)	PEBB, 2017 (N=6,158)	OEBB, 2015 (N=4,672)	OEBB, 2017 (N=5,558)
Age				
Less than 35	12%	14%	10%	13%
35–44	26%	26%	23%*	24%*
45–54	30%	28%	32%*	31%*
55–64	28%	26%	31%*	28%*
65 and older	4%	6%	3%*	4%*
Female	57%	57%	76%*	73%*
At least 4-year college degree	58%	59%	68%*	67%*
Non-white or Hispanic	18%	22%	13%*	18%*
Married or living with a significant other	78%	77%	80%*	79%*
Health and health care characteristics				
Health status				
Excellent	15%	12%	16%	12%
Very Good	42%	41%	45%*	44%*
Good	35%	36%	32%*	36%
Fair	8%	9%	6%*	8%*
Poor	1%	1%	1%	1%
Has visited usual source of care two or more times in the past 12 months	68%	68%	65%*	64%*
Has been going to usual source of care for at least 1 year	90%	88%	89%	87%

OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board.
Note: *OEBB and PEBB significantly different (P = 0.1) within year (2015 or 2017).

Table E-2-28 provides weighted, sample-wide means for each of the 16 patient experience measures we examined.[37] On balance, respondents reported positive perceptions of the health care they had received over the past year. For example, more than 85 percent reported that, overall, the health care they received in the past 12 months was high quality and easy to manage. Furthermore, a majority of respondents responded positively (i.e., "usually" or "always") on most of the indicators of care coordination, patient-centeredness, and accessibility. Over 90 percent of respondents, for example, reported that the provider at their usual source of care showed respect for what they had to say. On the other hand, less than half of respondents reported that their provider(s) seemed informed that they had received behavioral health care or that their provider(s) coordinated care among other providers. Only 10 percent reported that their provider(s) checked in with them between visits.

[37] The sample size is smaller than the full sample (9,981) for some measures (e.g., provider seemed informed that patient had received care from a specialist includes only those who reported having seen a specialist in the past 12 months).

Table E-2-28. Quality, care coordination, patient-centeredness, and access to care measures for PEBB and OEBB samples, by year

Indicator	PEBB				OEBB			
	2015		2017		2015		2017	
	Mean	Sample size	Mean	Sample size	Mean	Sample size	Mean	Sample size
Overall quality and manageability								
Overall, health care received in past 12 months was high quality	86%	5,309	86%	6,158	88%*	4,672	87%	5,558
Health care received in past 12 months was easy to manage	85%	5,309	84%	6,158	86%*	4,672	84%	5,558
Care Coordination								
Provider(s) followed up with patient's test results	88%	4,548	86%	5,329	87%	3,901	85%*	4,707
Provider(s) seemed informed that patient received care from a behavioral health provider	43%	793	42%	1,024	48%*	616	42%	757
Provider(s) seemed informed that patient received care from a specialist	66%	2,949	67%	3,333	67%	2,338	67%	2,750
Provider(s) helped coordinate patient's care among different providers	51%	2,935	53%	3,706	47%*	2,417	51%*	3,152
Provider(s) seemed to know important information about patient's hospital stay	85%	361	85%	426	85%	293	87%	339
Patient-centeredness								
Provider(s) seemed to know important information about medical history	85%	5,309	84%	6,158	86%	4,672	83%	5,558
Provider(s) asked about things in work or life at home that affect health	63%	5,309	63%	6,158	65%*	4,672	65%	5,558
Provider(s) showed respect for what patient had to say	92%	5,309	93%	6,158	94%*	4,672	94%*	5,558
Provider(s) talked with patient about specific goals for their health	60%	5,309	58%	6,158	56%*	4,672	55%*	5,558
Provider(s) checked in with patient between visits	10%	5,309	12%	6,158	10%	4,672	12%	5,558
One provider knew about all of patient's medical needs	82%	5,309	82%	6,158	83%	4,672	82%	5,558
Accessibility								
Able to get care from provider(s)'s office when it was closed	71%	2,032	73%	2,605	70%	1,821	71%	2,223
Easy to get appointment with behavioral health provider	78%	793	70%	1,024	75%	616	69%	757
Easy to get appointment with specialist	80%	2,949	76%	3,333	78%	2,338	75%	2,750

OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board.
Note: *OEBB and PEBB significantly different (P = 0.1) within year (2015 or 2017).

E-2.1.7 Statistical analysis

Regression model

APAC analysis

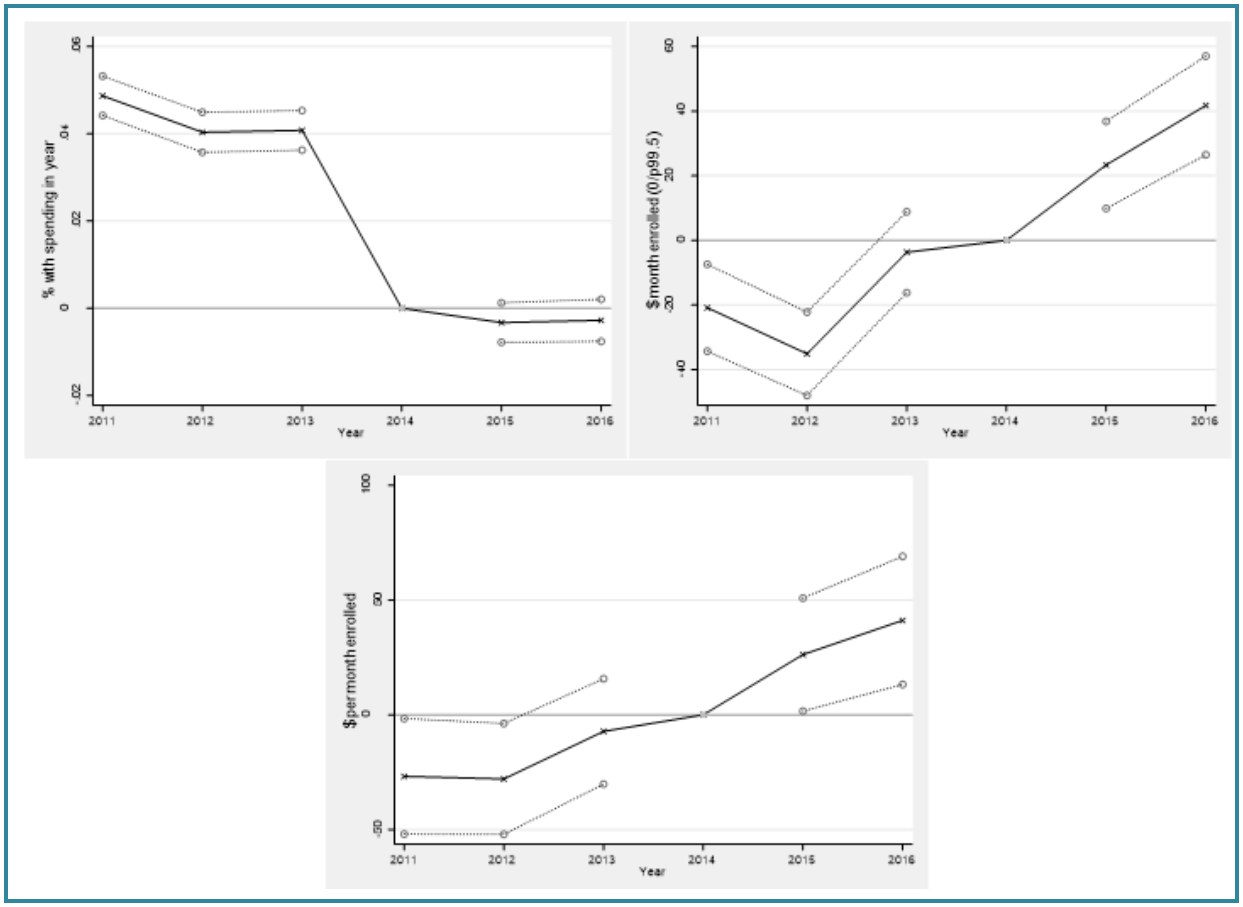
The underlying assumption in D-in-D estimation is that trends in the CCM treatment group would be similar to that of the comparison group in the absence of the initiative (i.e., that the two were on "parallel paths" prior to the start of the CCM in PEBB). Given that we only have one baseline observation period in the survey analysis, this assumption can only be tested for the claims analysis.

To assess the parallel assumption's validity in the APAC analyses empirically, we modeled core expenditure and utilization outcomes during the baseline period with year dummies interacted with a dichotomous variable indicating the beneficiary was enrolled in a PEBB (CCM) plan (i.e., the "test" group). This specification allows for nonlinear paths in the baseline period. The following section describes the baseline analysis we conducted to inform the D-in-D model.

We present "event history" graphs of annual unadjusted averages for the differences between PEBB and OEBB enrollees relative to the difference observed in the last year of the baseline period (2014) and their associated confidence intervals. If baseline trends are exactly parallel between PEBB and OEBB enrollees, the graphed values for 2011–2013 would be zero. If the confidence intervals around these points do not include zero we can reject the hypothesis that an outcome is following a parallel path for PEBB and OEBB. Such a finding at least suggests caution in interpreting significant D-in-D results as causal impact estimates.

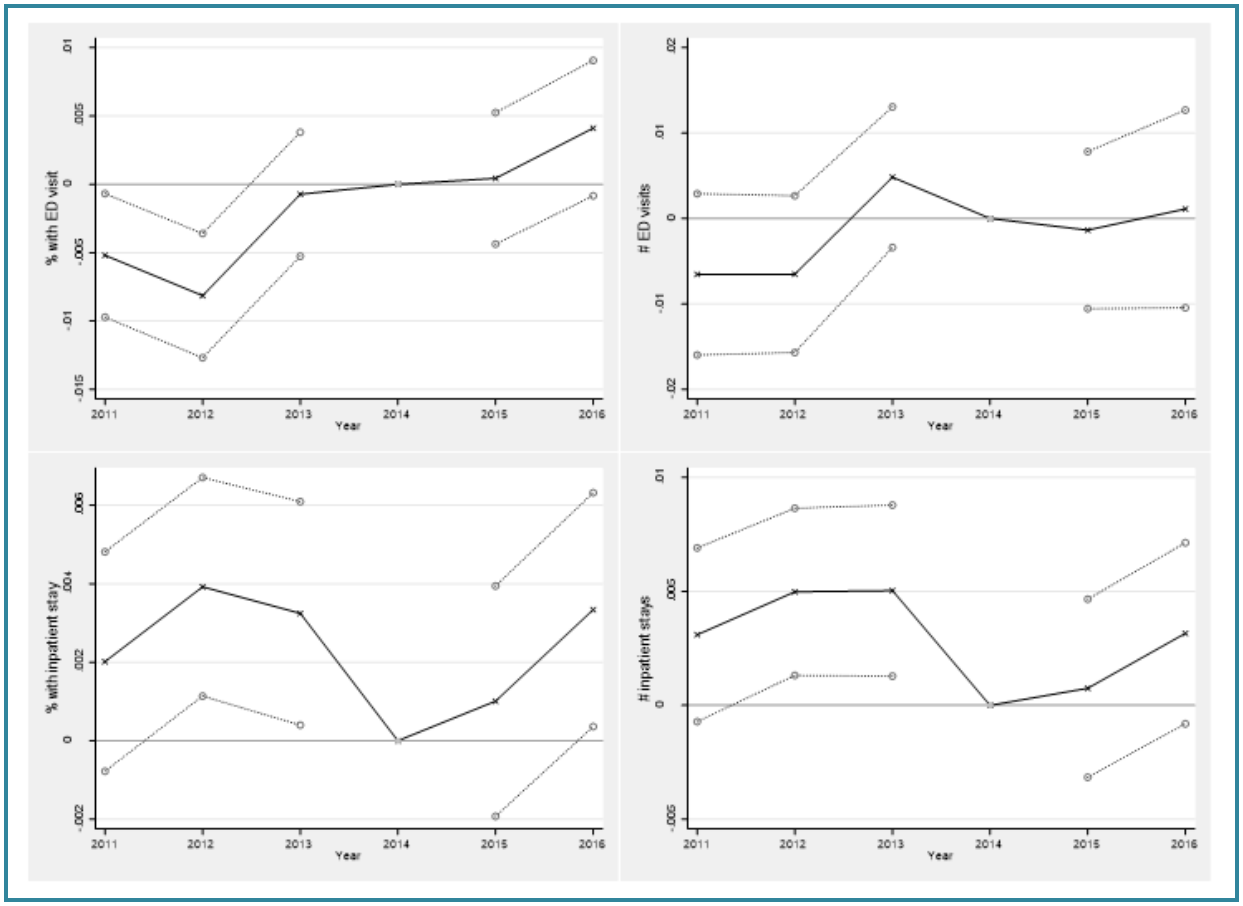
Figures E-2-2 to ***E-2-6*** provide event history graphs for the utilization and quality outcomes studied. With rare exception, we can reject the hypothesis that outcomes in the PEBB and OEBB populations are following the same patterns. It is possible that the differences in trend are not substantively meaningful but that the large number of observations produces very narrow confidence intervals making small deviations statistically significant. However, these plots also include single year values in the post-CCM implementation years, and there is little evidence that the post-period differences are of notably greater magnitude than the pre-period differences. **Thus, even though on average, several outcomes show significant changes between the average pre-implementation value and the average post-implementation for the CCM treatment group relative to the comparison group, these findings should be interpreted with caution**.

Figure E-2-2. Event history plots for adults' expenditures, PEBB vs. OEBB (reference year = 2014)



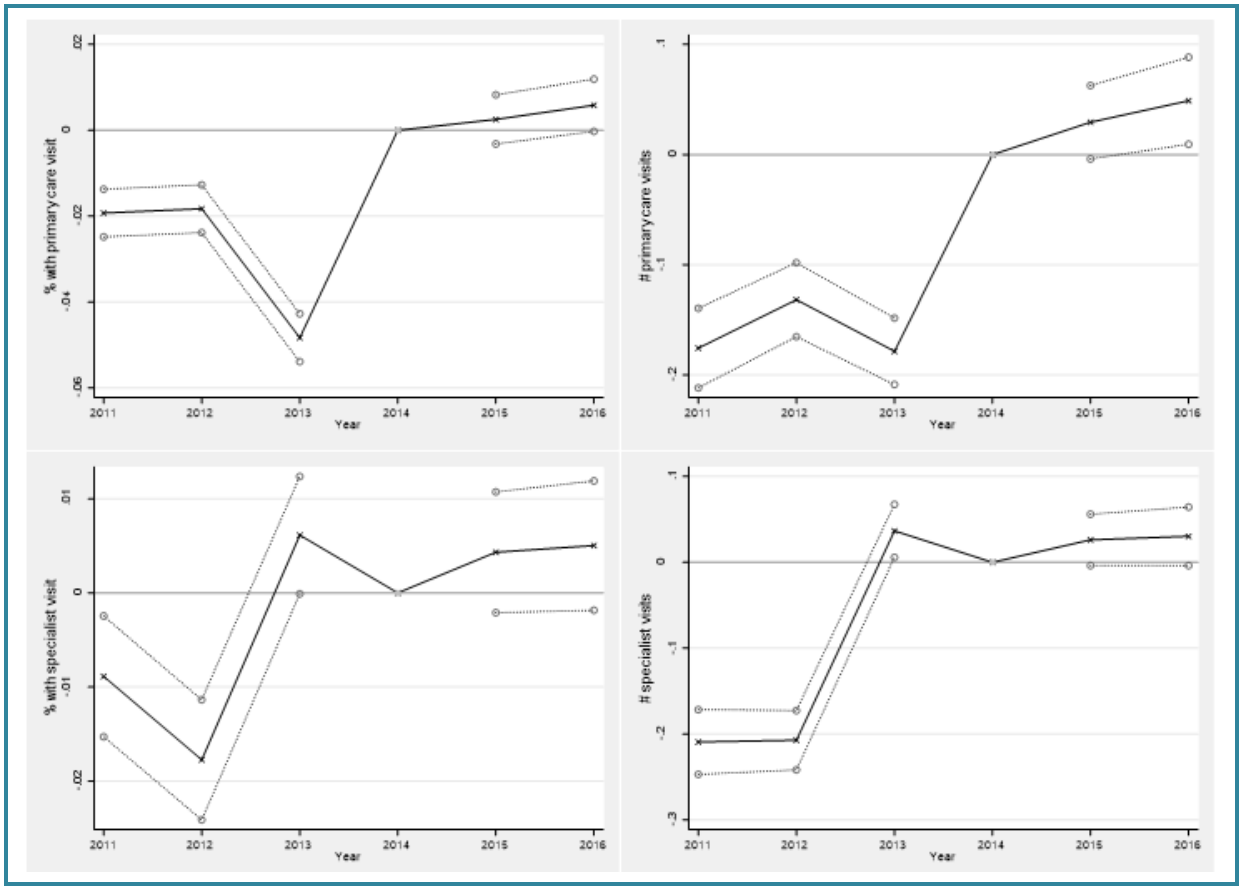
OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board.
Note: Event history plots show the difference in each outcome by year, relative to the final baseline year (2014). Values greater than zero in a given year indicate that the difference between the PEBB value and the OEBB value in that year was greater than the difference in 2014. Dotted lines indicate 90% confidence intervals.

Figure E-2-3. Event history plots for adults' utilization, PEBB vs. OEBB (reference year = 2014)



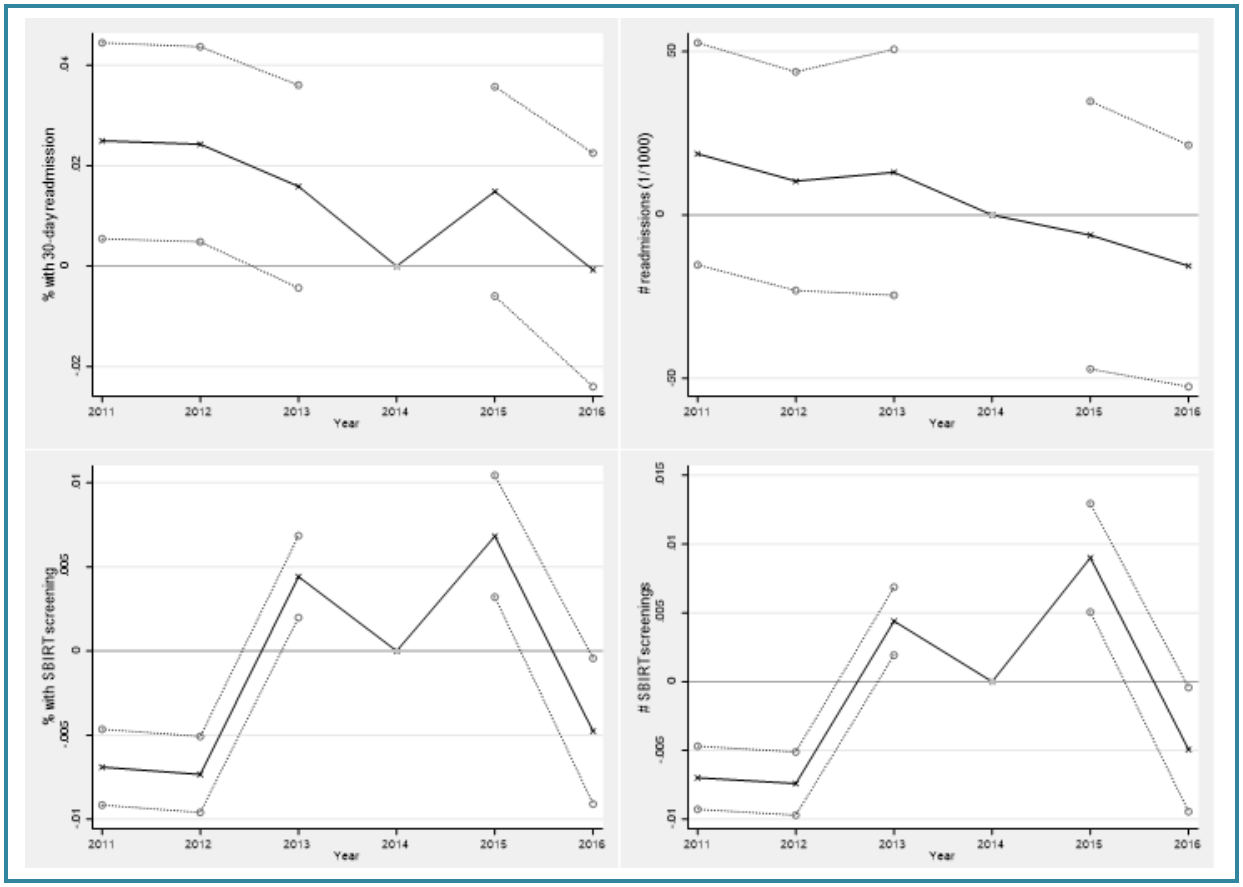
OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board.
Note: Event history plots show the difference in each outcome by year, relative to the final baseline year (2014). Values greater than zero in a given year indicate that the difference between the PEBB value and the OEBB value in that year was greater than the difference in 2014. Dotted lines indicate 90% confidence intervals.

Figure E-2-4. Event history plots for adults' physician visit utilization, PEBB vs. OEBB (reference year = 2014)



OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board.
Note: Event history plots show the difference in each outcome by year, relative to the final baseline year (2014). Values greater than zero in a given year indicate that the difference between the PEBB value and the OEBB value in that year was greater than the difference in 2014. Dotted lines indicate 90% confidence intervals.

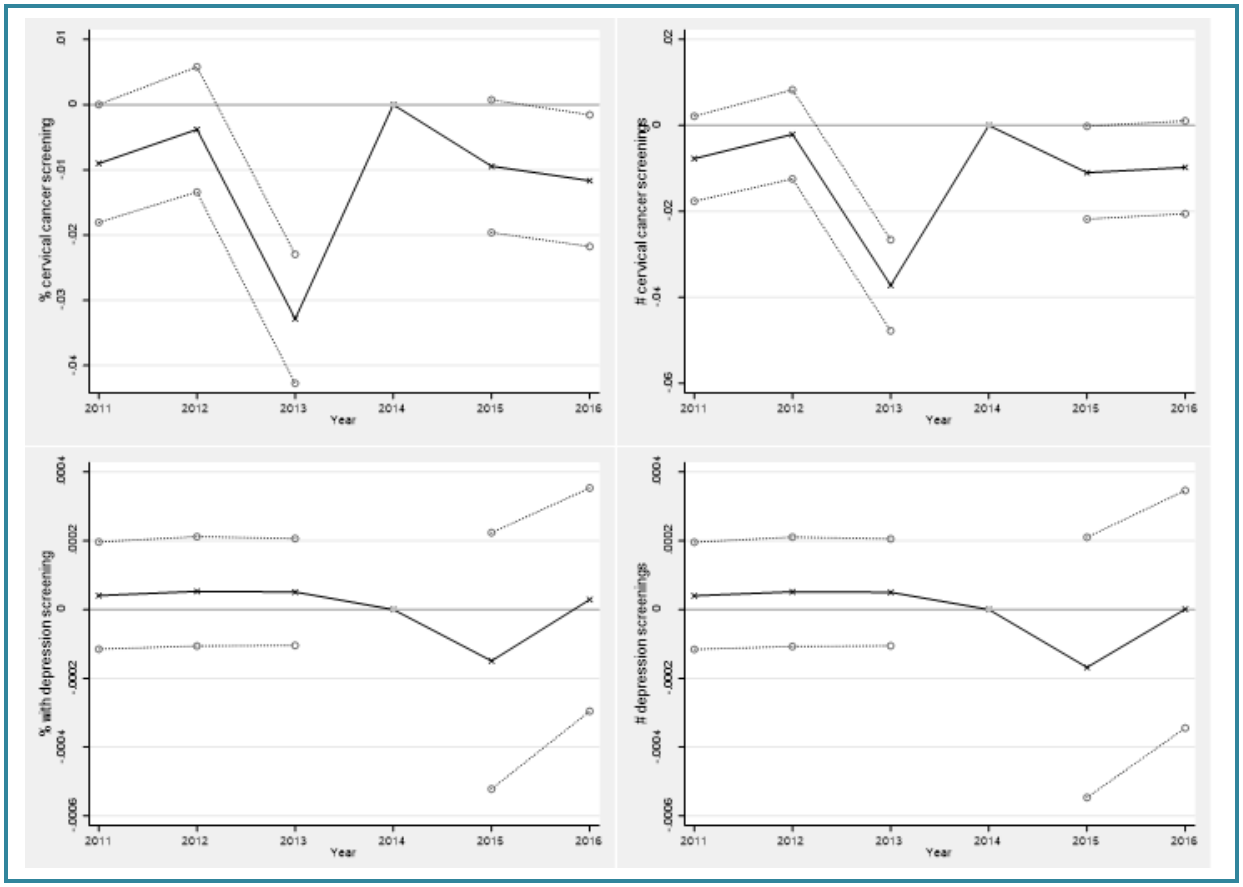
Figure E-2-5. Event history plots for adults' readmissions and SBIRT screenings, PEBB vs. OEBB (reference year = 2014)



OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board; SBIRT = Screening, Brief Intervention and Referral to Treatment.

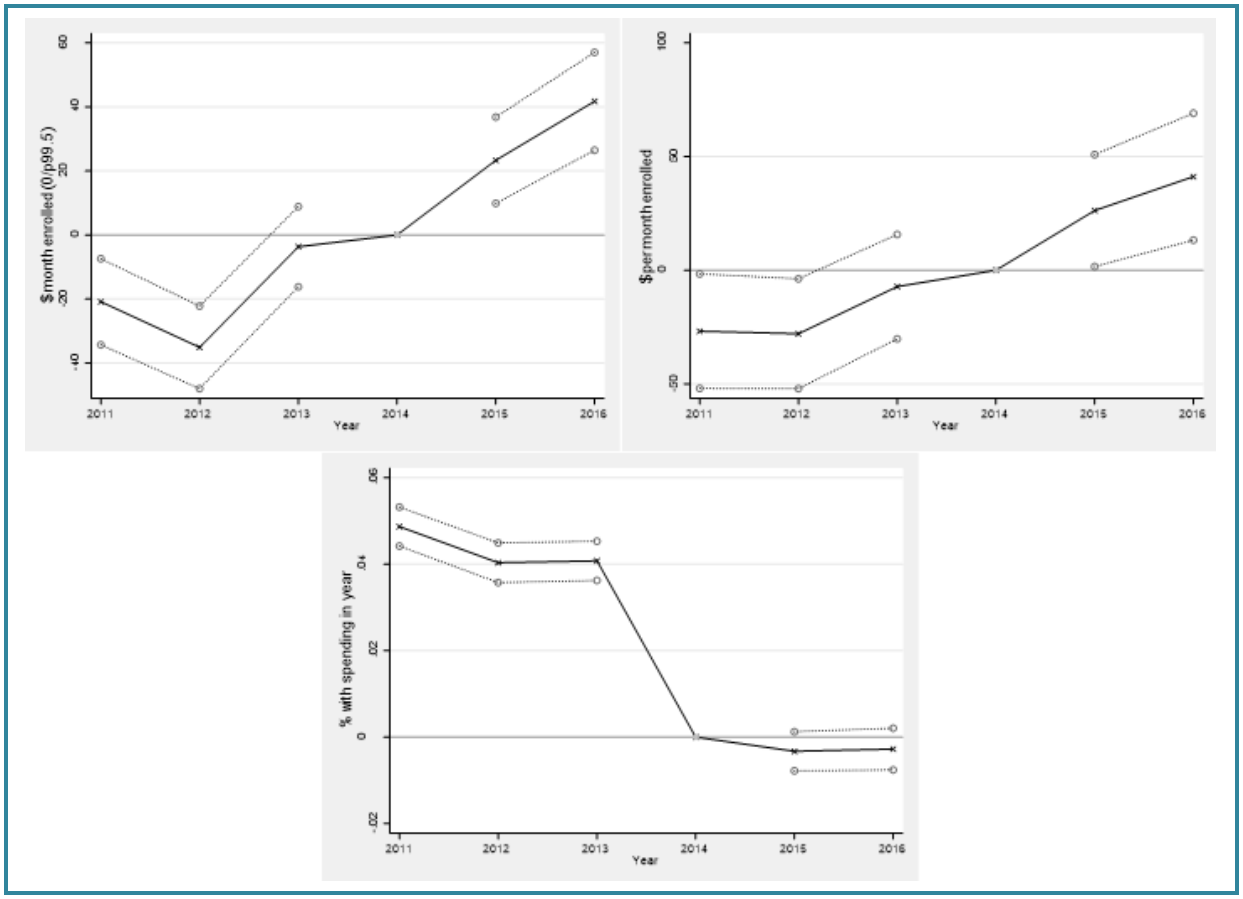
Note: Event history plots show the difference in each outcome by year, relative to the final baseline year (2014). Values greater than zero in a given year indicate that the difference between the PEBB value and the OEBB value in that year was greater than the difference in 2014. Dotted lines indicate 90% confidence intervals.

Figure E-2-6. Event history plots for adults' cervical cancer screenings and depression screenings, PEBB vs. OEBB (reference year = 2014)



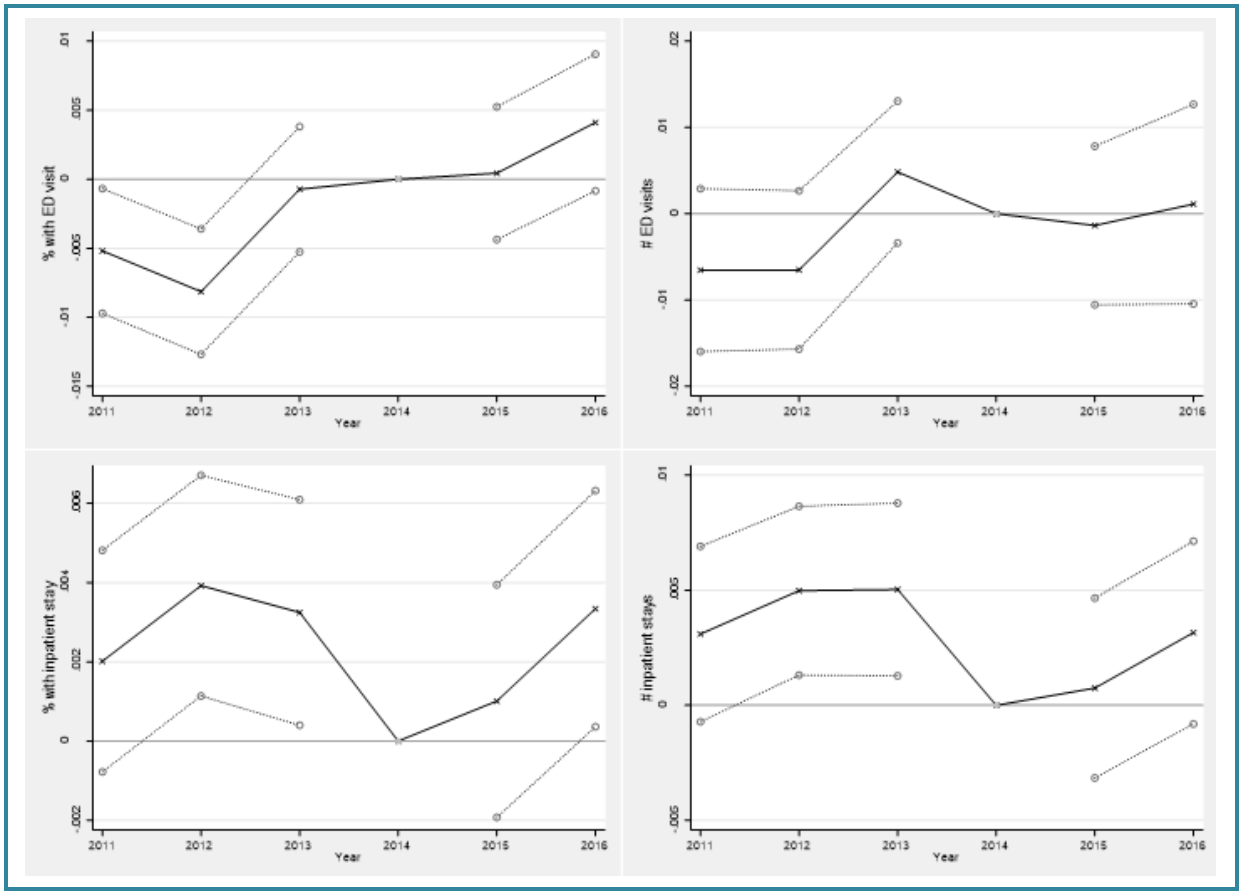
OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board.
Note: Event history plots show the difference in each outcome by year, relative to the final baseline year (2014). Values greater than zero in a given year indicate that the difference between the PEBB value and the OEBB value in that year was greater than the difference in 2014. Dotted lines indicate 90% confidence intervals.

Figure E-2-7. Event history plots for children's expenditures, PEBB vs. OEBB (reference year = 2014)



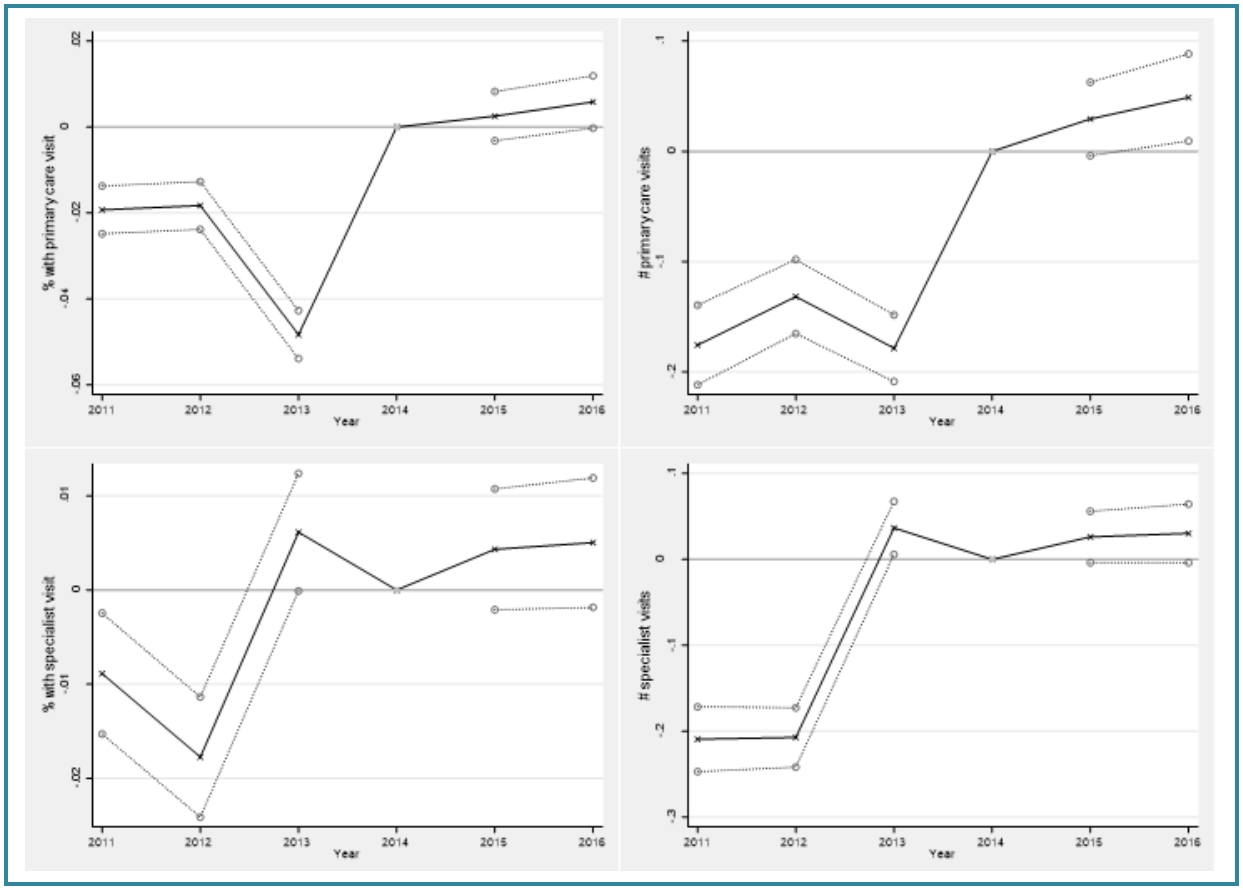
OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board.
Note: Event history plots show the difference in each outcome by year, relative to the final baseline year (2014). Values greater than zero in a given year indicate that the difference between the PEBB value and the OEBB value in that year was greater than the difference in 2014. Dotted lines indicate 90% confidence intervals.

Figure E-2-8. Event history plots for children's utilization, PEBB vs. OEBB (reference year = 2014)



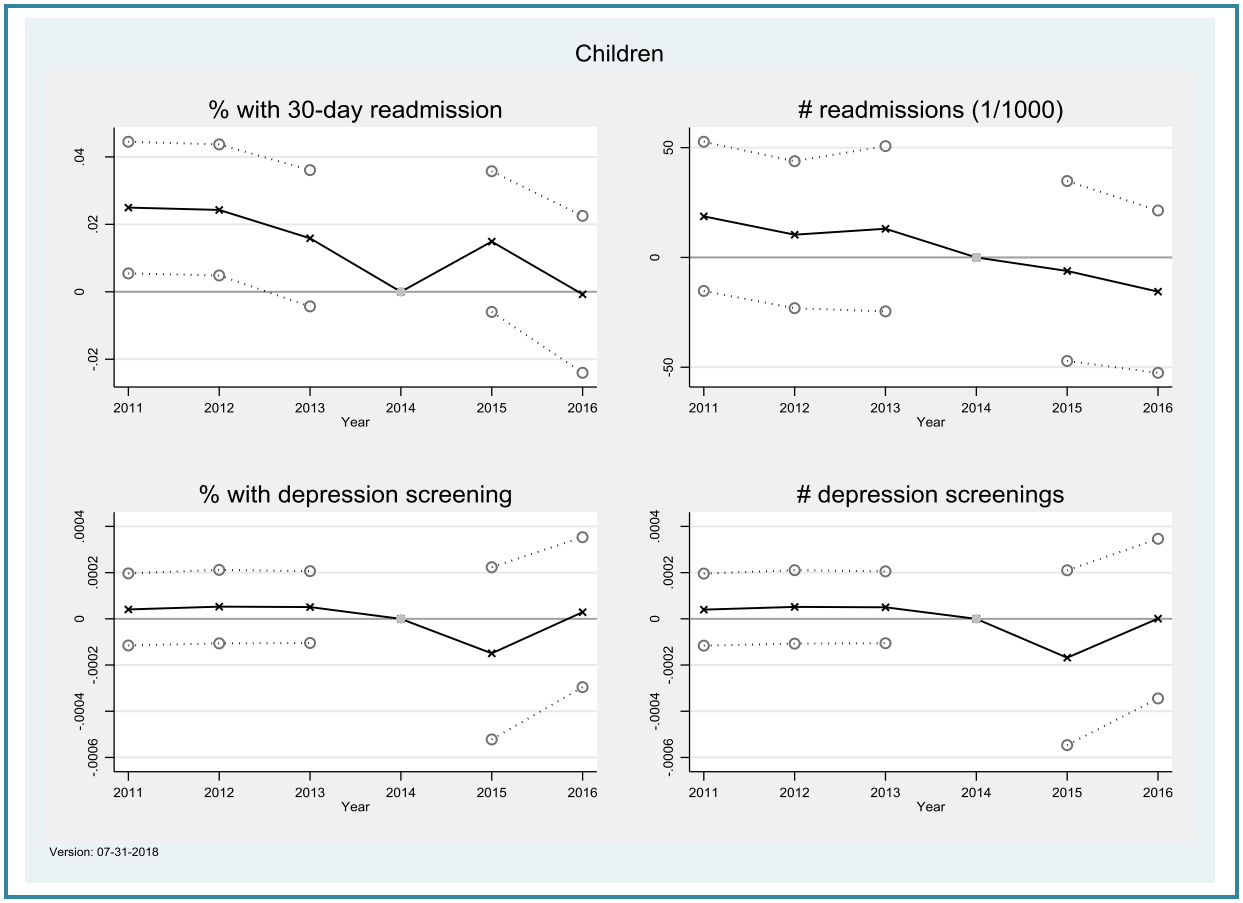
OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board.
Note: Event history plots show the difference in each outcome by year, relative to the final baseline year (2014). Values greater than zero in a given year indicate that the difference between the PEBB value and the OEBB value in that year was greater than the difference in 2014. Dotted lines indicate 90% confidence intervals.

Figure E-2-9. Event history plots for children's physician visit utilization, PEBB vs. OEBB (reference year = 2014)



OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board.
Note: Event history plots show the difference in each outcome by year, relative to the final baseline year (2014). Values greater than zero in a given year indicate that the difference between the PEBB value and the OEBB value in that year was greater than the difference in 2014. Dotted lines indicate 90% confidence intervals.

Figure E-2-10. Event history plots for children's quality of care, PEBB vs OEBB (reference year = 2014)



OEBB = Oregon Educators Benefit Board; PEBB = Public Employees Benefit Board.
Note: Event history plots show the difference in each outcome by year, relative to the final baseline year (2014). Values greater than zero in a given year indicate that the difference between the PEBB value and the OEBB value in that year was greater than the difference in 2014. Dotted lines indicate 90% confidence intervals.

Regression Analysis. The general form of the D-in-D regression analysis for APAC claims data is

$$Y_{it} = BT_t + \beta_1 PEBB_{it} + \beta_2 POST_{it} + \beta_3 PEBB_{it} * POST_{it} + \gamma X_{it} + \epsilon_{it} \qquad \text{(E-2.1)}$$

The coefficient estimate of interest from ***Equation E-2.1*** is β_3, which the D-in-D estimator that measures whether the between-group difference (CCM vs. comparison group) in the performance measure increased ($\beta_3 > 0$), decreased ($\beta_3 < 0$), or did not change ($\beta_3 = 0$) after the CCM was implemented in PEBB. BT_i represents a year dummy for individual baseline years 2011–2013 to control for potentially nonlinear pre-treatment trends. Variables included in the X_{it} vector include age, sex, number of months enrolled in a PEBB/OEBB plan, share of months

attributed to a PCPCH,[38] and a dummy for MSA of residence. Standard errors were adjusted for clustering within observations on the same individual.

Survey analysis

The models we estimate are designed to answer the research question: Does the spread of the CCM to state employees result in improved perceptions of health care? Using public educators (OEBB) as the control group, and a repeated cross-section design, the analysis appropriate for this question is a set of simple difference in differences models of the form

$$Y_i = f(\beta_1 PEBB_i + \beta_2 YEAR2017_i + \beta_3 PEBB_i * YEAR2017_i + \gamma X_i) \qquad \text{(E-2.2)}$$

The coefficient of interest in ***Equation E-2.2*** is again β_3, as it measures the differential change in outcomes for the CCM treatment group 2015 and 2017 relative to comparison group, to whom the CCM had not yet been introduced. Variables included in the X vector include age group (25–29 (reference group), 30–39, 40–49, 50–54, 55–64, 65–79), gender, education (high school or less (reference group), some college, 4-year college degree, more than 4-year college degree), marital status, race-ethnicity (white non-Hispanic (reference group), non-white non-Hispanic, Hispanic) and fair/poor health status. Because all outcomes in the survey are dichotomous we use a logit function for f(.) and use odds ratios to represent effects. Logistic regressions are weighted using sampling weights.

E-2.2 Methods for Qualitative Data Collection and Analysis

The Oregon SIM Initiative Round 1 Evaluation team collected and analyzed a wide range of qualitative data in the fifth year of the federal SIM Initiative evaluation. These data sources included interviews with key informants and focus groups conducted during in-person site visits in previous evaluation years, a review of relevant documents, and regular evaluation calls with the state officials leading the state's SIM Initiative. This report draws from past evaluation reports, where further detail is provided on previously conducted site visit interviews and focus groups.

E-2.2.1 Document review

The evaluation team used Oregon's quarterly and annual reports, operational plans, and other state documents to obtain updated information on its implementation progress during the SIM Initiative test period. To supplement these documents, we collected relevant news articles on the Oregon SIM Initiative activities and related initiatives, and we searched reform-oriented websites that the state maintains.

[38] An individual is attributed to a PCPCH if the plurality of visits for primary care services are billed by a recognized PCPCH. See the Year Four Annual Report (RTI International, 2018) for details.

In addition, we obtained numbers of providers participating in and populations served by Oregon's different innovation models from reports the states submit to the Innovation Center in conjunction with their quarterly reports. We provide Oregon's reported numbers in ***Appendix E***. Sources for these provider and population data are detailed in the Year Four Annual Report (RTI International, 2018).

E-2.2.2 State evaluation calls

We conducted monthly federal evaluation-specific calls beginning in April 2014 and continued through the end of the SIM Initiative test period. The evaluation team for Oregon, the state officials leading the Oregon SIM team, and the state's Innovation Center project officer typically attended each state evaluation call. Occasionally these calls were joined by state officials responsible for individual SIM activities and CMS technical assistance contractors. The purpose of the calls was to review interim evaluation findings with the state (as available), discuss any outstanding federal evaluation data or other needs, review and discuss state implementation and self-evaluation updates, and gather more in-depth information on select topics of interest for the evaluation.

For each meeting used to collect additional information and perspective from state officials leading the SIM Initiative in Oregon, the evaluation team prepared a list of state-specific questions—including the status of delivery and payment system reform initiatives and other SIM activities, related policy levers, and implementation successes, challenges, and lessons learned. We first reviewed relevant state documents for answers to our questions. When we did not find answers in the document or needed further clarification, we sent the questions to the state ahead of the call and ask the state to have knowledgeable state officials available to answer the questions during the call. We also used the calls to review and ask questions about state evaluation findings from particular SIM-funded activities.

E-2.2.3 Analysis

The evaluation team conducted thematic analysis of each source of qualitative data and then synthesized across information gleaned from site visit key informant interviews, site visit focus groups, document review, and state evaluation calls. Site visit interviews and focus groups were conducted in previous years of the evaluation. For more detail on site visit and focus group methods, see past evaluation reports.

E-2.3 References

Agency for Healthcare Research and Quality. (2016). *Defining the PCMH.* Available at https://pcmh.ahrq.gov/page/defining-pcmh

Centers for Medicare & Medicaid Services. (2015). *Quality rating system measure technical specifications, 2016* [Internet]. Available at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/Downloads/2017_QRS-Measure_Technical_Specifications.pdf

Measure #134 (NQF 0418): *Preventive care and screening: Screening for depression and follow-up plan* [Internet]. [cited 2017 Mar 28]. Available at https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/Practice-Management/quality-measures/2017/Preventative-Care-and-Screening-Screening-for-Depression-and-FollowUp-134-Claims.pdf

Molina Healthcare. (2017). *HEDIS 2016 provider guide and toolkit* [Internet]. Available at http://www.molinahealthcare.com/providers/wa/medicaid/comm/PDF/WA-Medicaid-HEDIS-Provider-Guide-Toolkit-COMBINED.pdf

Oregon Health Authority, Office of Health Analytics. *Technical specifications and guidance documents for CCO incentive measures* [Internet]. [cited 2017 Mar 21]. Available at http://www.oregon.gov/oha/analytics/Pages/CCO-Baseline-Data.aspx

Oregon Health Authority. (n.d.). *Coordinate care: The Oregon difference*. Available at https://www.oregon.gov/oha/HPA/Pages/CCOs-Oregon.aspx

RTI International. (2017). *State Innovation Models (SIM) Initiative Evaluation: Model Test Year Three annual report*. Prepared for the Centers for Medicare & Medicaid Services (CMS). Available at https://downloads.cms.gov/files/cmmi/sim-rd1mt-thirdannrpt.pdf

RTI International. (2018). *State Innovation Models (SIM) Initiative evaluation: Model Test Year Four annual report.* Prepared for the Centers for Medicare & Medicaid Services (CMS). Available at https://downloads.cms.gov/files/cmmi/sim-rd1-mt-fourthannrpt.pdf

Singer, S. J., Friedberg, M. W., Kiang, M. V., Dunn, T., & Kuhn, D. M. (2013). Development and Preliminary Validation of the Patient Perceptions of Integrated Care Survey. *Medical Care Research and Review*, *70*(2), 143–164. https://doi.org/10.1177/1077558712465654

WellCare Health Plans. 2016 HEDIS adult resource guide [Internet]. [cited 2017 Mar 30]. Available at https://www.wellcare.com/~/media/PDFs/NA/Provider/na_hedis_adult_resource_guide_eng_01_2016_v2.ashx

Sub-appendix F-2: Methods for Vermont Analyses

F-2.1 Methods for the Vermont Medicaid SSP Impact Analysis

To estimate the impact of the Medicaid Shared Savings Program (SSP) in Vermont, we conducted difference-in-differences (D-in-D) regression analyses using Medicaid claims data provided by the Department of Vermont Health Access (DVHA). In ***Appendix F, Section F.3***, we present D-in-D analyses for outcomes across four evaluation domains: (1) care coordination, (2) service utilization, (3) expenditures, and (4) quality of care. This sub-appendix details the methods we used for this analysis.

The Medicaid SSP in the context of Vermont Medicaid. From 2014 to 2016, Vermont partnered with existing accountable care organizations (ACOs) to implement an alternative payment model called the Shared Savings Program. With support from the SIM Initiative, Vermont's SSPs allowed participating ACOs to share in savings based on achieving cost and quality targets among their attributed beneficiaries. Vermont's payment reforms included both state and commercial payers; however, the quantitative impact analysis in ***Appendix F, Section F.3***, focuses only on the performance of Medicaid SSP beneficiaries.

Medicaid in Vermont is a combination of fee-for-service and primary care case management with approximately 60 percent of beneficiaries reached by Blueprint for Health in 2014. The start of the Medicaid SSP in Vermont coincided with the expansion of Medicaid eligibility in 2014 under the Patient Protection and Affordable Care Act (ACA). Although Vermont already provided expanded Medicaid coverage or subsidies prior to the ACA through the Vermont Health Access Plan (Medicaid) and Catamount Health (subsidies for private insurance), the expansion still resulted in a significant number of first-time Medicaid enrollees participating in the SSP beyond those who transitioned from the two aforementioned expansion programs (Vermont Digger, 2016). Between July and September 2013 and December 2015, Medicaid enrollment in Vermont increased by approximately 18 percent (CMS, 2016).

Profiles of Medicaid SSP and comparison groups. The level of intervention in Vermont's Medicaid SSP is at the ACO. During the SIM Initiative Vermont had three ACOs—OneCare Vermont, Community Health Accountable Care (CHAC), and Vermont Collaborative Physicians (VCP). Although all three ACOs participated in the commercial SSP, ultimately only CHAC and OneCare chose to participate in the Medicaid SSP. VCP chose not to participate in the Medicaid SSP because of its prior experience with the Medicare ACO SSP.

The three ACOs differed somewhat in terms of their provider network. The majority of OneCare's network included hospitals and hospital-owned physician practices along with some independent physicians. The CHAC network consisted primarily of physicians at federally qualified health centers across the state. VCP—the ACO that did not participate in the Medicaid SSP—consisted of providers working at physician-owned practices throughout the state.

Vermont's Medicaid agency conducted the attribution and provided RTI with lists of both the beneficiaries attributed to the Medicaid SSP ACOs (CHAC, OneCare) and beneficiaries attributed to either VCP or to unaffiliated providers (the comparison group). Outside of the nature of their ACOs, we have limited information with which to compare providers between the Medicaid SSP and comparison groups.

Medicaid restricted beneficiary attribution in the SSP to members with full, unrestricted benefits who were not enrolled in Medicare. Beneficiary attribution occurred on an annual basis, and to be eligible members had to have been enrolled for at least 10 months in the year of attribution. Attribution occurred either through a claims-based algorithm or through the affiliation of the beneficiary's assigned primary care provider.

Because attribution took place yearly, both Medicaid SSP and comparison groups contain beneficiaries whose first attribution occurred after the start of the SSP in 2014. However, the proportion of beneficiaries who were attributed after 2014 is similar between the Medicaid SSP and comparison groups. Among Medicaid SSP and comparison beneficiaries (denoted in parentheses), approximately 26 percent (28 percent) were attributed in 2015 and 13 percent (13 percent) were attributed in 2016.

Our analysis compares pre- and post-periods for the Medicaid SSP and comparison groups using a longitudinal design with an unbalanced panel. This means we included beneficiaries who were eligible for Medicaid for the first time in the year of their attribution. The rationale for an unbalanced panel—as opposed to using a balanced panel approach—is to provide estimation of the program's impact encompassing the entire population of attributed beneficiaries—not a subset based on prior eligibility. The disadvantage of an unbalanced design, however, is the inclusion of beneficiaries without baseline observational data prior to their attribution. Even so, 50 percent of the sample does have a full panel of data. To further assess comparability, we tested for and found parallel trends in key outcomes between the groups during the baseline period. And finally, as shown in the balance tables by year, average beneficiary characteristics do not differ substantially year to year within the Medicaid SSP group or within the comparison group, suggesting that even though some beneficiaries may not have baseline data, the characteristics of the sample are not changing over time (with the exception of being a beneficiary eligible through Medicaid expansion).

In addition to the analysis on the overall population, we conducted two subpopulation analyses: (1) children and adults separately and (2) beneficiaries diagnosed with behavioral health conditions. We also conducted two types of sensitivity analyses. The first analysis modeled outcomes separately for each of the ACOs to evaluate them independently. The second analysis looked at beneficiaries in both the Medicaid SSP and comparison group who were attributed at the beginning of the demonstration and who remained attributed in all 3 years of the

demonstration. This was done to investigate whether the demonstration's impact was stronger among beneficiaries with stable and consistent exposure.

Balancing Medicaid SSP and comparison groups. Using the list of beneficiaries attributed in 2014, 2015, and 2016 provided by the state, we selected as our potential comparison group all beneficiaries who were never attributed to a CHAC or OneCare provider. We then estimated propensity score weights to balance the Medicaid SSP and comparison groups on many individual and county-level characteristics. To account for the impact of Medicaid expansion, we included in our model an indicator for 10 or more months enrollment in the previous year. Propensity score weighting rather than matching was used to avoid dropping any beneficiaries from the final analysis. After propensity score weighting, the standardized differences between the weighted comparison group means and Medicaid SSP group means were all well under the standard 10 percent threshold. More information on propensity score weighting is available in ***Sections F-2.3*** and ***F-2.4.***

Study design. Our analysis uses a longitudinal design with an unbalanced sample of intervention and comparison beneficiaries. The samples are unbalanced because beneficiaries are only included in our models in periods when they are both Medicaid and Medicaid SSP attribution eligible. Some beneficiaries are therefore missing in parts of the pre- and post-period. We model intervention effects using the traditional D-in-D framework where estimates represent the pre-post difference in outcome trends among intervention beneficiaries in the absence of treatment.

Statistical approach. Analyses used D-in-D ordinary least squares models for spending outcomes and logistic regression models for binary outcomes. All analyses used clustered standard errors at the provider level to account for repeated observations of beneficiaries within providers over time. More information on the regression model is available in ***Section F-2.5***.

F-2.1.1 Data sources

Medicaid data. To derive eligibility information and claims-based outcomes for our analytic sample, we used Medicaid claims data provided by Vermont's DVHA. In this report, we used data from 2011 to 2016 to examine the 3 years before (2011–2013) and the 3 years after (2014–2016) the start of the Medicaid SSP. The DVHA data included three linkable files containing monthly enrollment and coverage information along with inpatient, outpatient, and pharmaceutical claims. DVHA includes all Medicaid beneficiaries in the state, but for this report we subset the analytic sample to beneficiaries ages 0 through 64 and excluded dually eligible Medicare-Medicaid enrollees.

Attribution file. We received a list from the DVHA of all Medicaid beneficiaries attributed to providers participating in the Medicaid SSP in 2014 or 2016. Included on those lists were beneficiaries identified during the state's attribution process as being attributed to providers

not participating in the SSP. Beneficiaries in this second group form the comparison group for our analysis. Provider attribution was independent across years, and we received separate lists from the state for each year. Beneficiaries included in these attribution lists were linked using a unique member ID number to the DVHA enrollment and claims data to form the analysis sample. Only beneficiaries attributed at least once (to SSP or comparison providers) are included in our analysis.

Area Health Resource File. The AHRF comprises data collected by the Health Resources and Services Administration from more than 50 sources containing more than 6,000 variables related to health care access at the county level. We used information on health professions supply, hospital bed supply, and population characteristics and economic data to select the comparison group and to use as covariates in the analysis.

F-2.1.2 Outcome measures

Care coordination

To evaluate the impact of the Medicaid SSP in Vermont on care coordination, we report the following care coordination measures. The measures were calculated annually for all eligible beneficiaries in the Medicaid SSP and comparison group.

- Percentage of mental illness–related acute inpatient hospital admissions with a mental health follow-up visit within 7 and 30 days. This is the number of acute inpatient hospital admissions with a primary diagnosis for a mental health disorder (ICD-9 and ICD-10 diagnosis codes in the Healthcare Effectiveness Data and Information Set [HEDIS] measure Follow-Up After Hospitalization for Mental Illness [FUH]) followed by a visit to a provider for a mental health outpatient visit, intensive outpatient encounter, or partial hospitalization (identified by the Current Procedural Terminology (CPT), Healthcare Common Procedure Coding System [HCPCS], Place of Service, UB Revenue codes specified in the HEDIS measure) within 7 or 30 days of discharge date, divided by total number of acute inpatient hospital admissions with a primary diagnosis for a mental disorder. Admissions followed by a readmission to an acute or other facility within 7 or 30 days are excluded from the respective denominators. Discharges because of death are also excluded from the respective denominators. The denominators were limited to patients 6 years or older as of the date of discharge.

  For both indicators, 7-Day Follow-Up and 30-Day Follow-Up, any of the following meet the criteria for a follow-up visit using codes specified in the HEDIS measure FUH:

  – A visit with a mental health practitioner
  – A visit to a behavioral health care facility
  – A visit to a non–behavioral health care facility with a mental health practitioner
  – A visit to a non–behavioral health care facility with a diagnosis of mental illness

- Transitional care management services, where the date of service on the claim is 29 days after the date the patient was discharged with a principal diagnosis of mental illness

The following meets the criteria for only the 30-Day Follow-Up indicator:

- Transitional care management services, where the date of service on the claim is 29 days after the date the patient was discharged with a principal diagnosis of mental illness

In addition, we report the following medication management care coordination measure:

- **Percentage of patients ages 18–64 years with a diagnosis of acute bronchitis who were not dispensed an antibiotic prescription.** This is the percentage of patients identified with acute bronchitis who did not have an antibiotic prescription dispensed to them during the year. To identify patients with acute bronchitis, the patient had to be 18–64 years old, have a diagnosis for acute bronchitis (ICD-9 diagnosis code 466.0; ICD-10 diagnosis codes J20.3–J20.9), and have at least one of the following three episodes during the intake period (January 1–December 24 of the measurement year):
  i. At least one emergency department (ED) visit with acute bronchitis as a diagnosis (CPT code = 99281–99285 or Revenue Code = 045x, 0981)
  ii. At least one observation visit with acute bronchitis as a diagnosis (CPT code = 99217–99220)
  iii. At least one outpatient visit with acute bronchitis as a diagnosis (CPT code = 99201–99205, 99211–99215, 99241–99245, 99341–99345, 99347–99350, 99381–99387, 99391–99397, 99401–99404, 99411, 99412, 99420, 99429, 99455, 99456 or HCPCS code = G0402, G0438, G0439, G0463, T1015 or Revenue Code = 051x, 0520–0523, 0526–0529, 0982, 0983)

The episode with the earliest start date was identified as the index episode.

Patients diagnosed with HIV, malignant neoplasms, emphysema, chronic obstructive pulmonary disease (COPD), cystic fibrosis, and other comorbid conditions in the year prior to the index episode were excluded from the denominator (identified by the ICD-9 and ICD-10 diagnosis codes in the following Value Sets in the HEDIS measure Avoidance of Antibiotic Treatment in Adults with Acute Bronchitis: HIV, Malignant Neoplasms, Emphysema, COPD, Cystic Fibrosis, and Comorbid Conditions).

Patients diagnosed with pharyngitis or another competing diagnosis 30 days prior to 7 days after the index episode (inclusive) were excluded from the denominator (identified by the ICD-9 and ICD-10 diagnosis codes in the following Value Sets in the HEDIS measure: Pharyngitis, Competing Diagnosis).

Patients who had a new or refill prescription for an antibiotic medication in ***Table F-2-1*** filled on or 30 days prior to the index episode were excluded from the denominator.

Table F-2-1. Antibiotic medications list

Description	Prescription		
Aminoglycosides	• Amikacin • Gentamicin	• Kanamycin • Streptomycin	• Tobramycin
Aminopenicillins	• Amoxicillin	• Ampicillin	
Antipseudomonal penicillins	• Piperacillin		
Beta-lactamase inhibitors	• Amoxicillin-clavulanate • Ampicillin-sulbactam	• Piperacillin-tazobactam	• Ticarcillin-clavulanate
First generation cephalosporins	• Cefadroxil	• Cefazolin	• Cephalexin
Fourth generation cephalosporins	• Cefepime		
Ketolides	• Telithromycin		
Lincomycin derivatives	• Clindamycin	• Lincomycin	
Macrolides	• Azithromycin • Clarithromycin	• Erythromycin • Erythromycin ethylsuccinate	• Erythromycin lactobionate • Erythromycin stearate
Miscellaneous antibiotics	• Aztreonam • Chloramphenicol • Dalfopristin-quinupristin	• Daptomycin • Erythromycin-sulfisoxazole • Linezolid	• Metronidazole • Vancomycin
Natural penicillins	• Penicillin G benzathine-procaine • Penicillin G potassium	• Penicillin G procaine • Penicillin G sodium	• Penicillin V potassium • Penicillin G benzathine
Penicillinase resistant penicillins	• Dicloxacillin	• Nafcillin	• Oxacillin
Quinolones	• Ciprofloxacin • Gemifloxacin	• Levofloxacin • Moxifloxacin	• Norfloxacin • Ofloxacin
Rifamycin derivatives	• Rifampin		
Second generation cephalosporin	• Cefaclor • Cefotetan	• Cefoxitin • Cefprozil	• Cefuroxime
Sulfonamides	• Sulfadiazine	• Sulfamethoxazole-trimethoprim	
Tetracyclines	• Doxycycline	• Minocycline	• Tetracycline
Third generation cephalosporins	• Cefdinir • Cefditoren • Cefixime	• Cefotaxime • Cefpodoxime • Ceftazidime	• Ceftibuten • Ceftriaxone
Urinary anti-infectives	• Fosfomycin • Nitrofurantoin • Nitrofurantoin macrocrystals	• Nitrofurantoin macrocrystals-monohydrate • Trimethoprim	

For individuals who met the above acute bronchitis criteria, we flagged whether they were dispensed at least one prescription for one of the antibiotic medications in ***Table F-2-1*** on or within 3 days after the index episode during the measurement year and calculated the inverted

percentage [1 − (number of individuals dispensed at least one prescription for an eligible antibiotic medication/eligible population)]. A higher percentage indicates appropriate treatment of adults with acute bronchitis.

Utilization

Utilization measures are reported as rates per 1,000 covered lives (or discharges for readmissions). For each measure, we first calculate the probability of any use. To calculate the probability, the numerator was an indicator of having had at least one event (inpatient admission or ED visit that did not lead to a hospitalization) and the denominator is the number of eligible plan members (or discharges) in the state enrolled during the period. We multiplied the probability of use by 1,000 to obtain approximate rates of utilization per 1,000 beneficiaries. Multiplying the probability by 1,000 does not produce an exact rate of utilization per 1,000 beneficiaries because it assumes that no person has more than one visit or admission per year. However, we concluded that this is a reasonable approximation because only a small percentage of beneficiaries had counts exceeding one for any of the utilization measures. Events are included in a period's total if discharge or service date on the claim was during the period.

- **Probability of having any inpatient use:** This is a dichotomous variable indicating beneficiaries with at least one admission to an acute-care hospital in the observation period. Acute admissions were identified through claims where place of service indicated the admission was to an inpatient hospital (bill type = 11 or 12). Admissions representing transfers between facilities were collapsed into a single admission. An admission was counted as a facility transfer when the time between the discharge date of the index claim and the admission date of the subsequent claim was no more than 1 day.
- **Probability of having any ED visits that did not lead to a hospitalization (outpatient ED) use:** This is a dichotomous variable indicating beneficiaries with at least one visit to the ED that did not result in an inpatient hospital admission. Outpatient ED visits (including observation stays) are identified in the outpatient services file through claims with a revenue center line item equal to 045X, 0981 (ED care) or 0762 (treatment or observation room, thus including observation stays in the overall count). If the procedure code on every line item of the ED claim equaled 70000–89999 and no line items had a revenue center code equal to 0762, that claim was excluded (thus excluding claims for which only radiology or pathology/laboratory services were provided unless they were observation stays). Multiple ED visits on a single day were counted as a single visit.
- **Probability of having a visit to a primary care or specialty care physician:** These are separate dichotomous variables indicating beneficiaries with at least one visit to the primary care physician or at least one visit to a specialty care physician during the period. Physician visits were identified in the outpatient claims file using CPT codes associated with Evaluation and Management visits (***Table F-2-2***). Visits were classified as either primary or specialty care by referencing the provider taxonomy associated with the claim against the classifications in ***Table F-2-3***.

Table F-2-2. Evaluation and management CPT codes

CPT Code	CPT Code	CPT Code	CPT Code	CPT Code	CPT Code	CPT Code	CPT Code
99201	99242	99315	99337	99350	99379	99393	99408
99202	99243	99316	99339	99358	99380	99394	99409
99203	99244	99318	99340	99359	99381	99395	99411
99204	99245	99324	99341	99366	99382	99396	99412
99205	99304	99325	99342	99367	99383	99397	99420
99211	99305	99326	99343	99368	99384	99401	99429
99212	99306	99327	99344	99374	99385	99402	99441
99213	99307	99328	99345	99375	99386	99403	99442
99214	99308	99334	99347	99376	99387	99404	99443
99215	99309	99335	99348	99377	99391	99406	99444
99241	99310	99336	99349	99378	99392	99407	99496

Table F-2-3. Primary/specialty care taxonomies

Taxonomy	Type	Taxonomy	Type	Taxonomy	Type	Taxonomy	Type
193200000X	Primary	207PT0002X	Specialty	207X00000X	Specialty	208100000X	Specialty
208D00000X	Primary	207PP0204X	Specialty	207YS0123X	Specialty	2082S0099X	Specialty
363A00000X	Primary	207PS0010X	Specialty	207YX0602X	Specialty	2082S0105X	Specialty
363AM0700X	Primary	207PE0005X	Specialty	207YX0905X	Specialty	208200000X	Specialty
363L00000X	Primary	207P00000X	Specialty	207YX0901X	Specialty	2083T0002X	Specialty
363LA2100X	Primary	207QA0401X	Specialty	207YP0228X	Specialty	2083X0100X	Specialty
363LA2200X	Primary	207QB0002X	Specialty	207YX0007X	Specialty	2083P0500X	Specialty
363LC1500X	Primary	207QH0002X	Specialty	207YS0012X	Specialty	2083P0901X	Specialty
363LF0000X	Primary	207QS1201X	Specialty	207Y00000X	Specialty	2083S0010X	Specialty
363LG0600X	Primary	207QS0010X	Specialty	208VP0014X	Specialty	2083P0011X	Specialty
363LP2300X	Primary	207RA0401X	Specialty	208VP0000X	Specialty	2084A0401X	Specialty
363LW0102X	Primary	207RA0201X	Specialty	207ZP0101X	Specialty	2084P0802X	Specialty
364S00000X	Primary	207RB0002X	Specialty	207ZP0102X	Specialty	2084B0002X	Specialty
364SA2100X	Primary	207RC0000X	Specialty	207ZB0001X	Specialty	2084P0804X	Specialty
364SA2200X	Primary	207RI0001X	Specialty	207ZP0104X	Specialty	2084N0600X	Specialty
364SC0200X	Primary	207RC0001X	Specialty	207ZC0006X	Specialty	2084D0003X	Specialty
364SC1501X	Primary	207RC0200X	Specialty	207ZP0105X	Specialty	2084F0202X	Specialty
364SC2300X	Primary	207RE0101X	Specialty	207ZC0500X	Specialty	2084P0805X	Specialty
364SF0001X	Primary	207RG0100X	Specialty	207ZD0900X	Specialty	2084H0002X	Specialty
364SG0600X	Primary	207RH0000X	Specialty	207ZF0201X	Specialty	2084P0005X	Specialty
364SH0200X	Primary	207RH0003X	Specialty	207ZH0000X	Specialty	2084N0400X	Specialty

(continued)

Table F-2-3. Primary/specialty care taxonomies (continued)

Taxonomy	Type	Taxonomy	Type	Taxonomy	Type	Taxonomy	Type
364SW0102X	Primary	207RI0008X	Specialty	207ZI0100X	Specialty	2084N0402X	Specialty
208000000X	Primary	207RH0002X	Specialty	207ZM0300X	Specialty	2084N0008X	Specialty
2080A0000X	Primary	207RI0200X	Specialty	207ZP0007X	Specialty	2084P2900X	Specialty
207V00000X	Primary	207RI0011X	Specialty	207ZN0500X	Specialty	2084P0800X	Specialty
207VG0400X	Primary	207RM1200X	Specialty	207ZP0213X	Specialty	2084P0015X	Specialty
207VM0101X	Primary	207RX0202X	Specialty	2080C0008X	Specialty	2084S0012X	Specialty
207VX0000X	Primary	207RN0300X	Specialty	2080I0007X	Specialty	2084S0010X	Specialty
207Q00000X	Primary	207RP1001X	Specialty	2080P0006X	Specialty	2084V0102X	Specialty
207QA0000X	Primary	207RR0500X	Specialty	2080H0002X	Specialty	2085B0100X	Specialty
207QA0505X	Primary	207RS0012X	Specialty	2080T0002X	Specialty	2085D0003X	Specialty
207QG0300X	Primary	207RS0010X	Specialty	2080N0001X	Specialty	2085R0202X	Specialty
207R00000X	Primary	207RT0003X	Specialty	2080P0008X	Specialty	2085U0001X	Specialty
207RA0000X	Primary	207T00000X	Specialty	2080P0201X	Specialty	2085H0002X	Specialty
207RG0300X	Primary	204D00000X	Specialty	2080P0202X	Specialty	2085N0700X	Specialty
207KA0200X	Specialty	204C00000X	Specialty	2080P0203X	Specialty	2085N0904X	Specialty
207KI0005X	Specialty	207UN0903X	Specialty	2080P0204X	Specialty	2085P0229X	Specialty
207K00000X	Specialty	207UN0901X	Specialty	2080P0205X	Specialty	2085R0001X	Specialty
207LA0401X	Specialty	207UN0902X	Specialty	2080P0206X	Specialty	2085R0205X	Specialty
207LC0200X	Specialty	207U00000X	Specialty	2080P0207X	Specialty	2085R0203X	Specialty
207LH0002X	Specialty	207VB0002X	Specialty	2080P0208X	Specialty	2085R0204X	Specialty
207LP2900X	Specialty	207VC0200X	Specialty	2080P0210X	Specialty	2086H0002X	Specialty
207LP3000X	Specialty	207VX0201X	Specialty	2080P0214X	Specialty	2086S0120X	Specialty
207L00000X	Specialty	207VH0002X	Specialty	2080P0216X	Specialty	2086S0122X	Specialty
208C00000X	Specialty	207VE0102X	Specialty	2080T0004X	Specialty	2086S0105X	Specialty
207NI0002X	Specialty	207W00000X	Specialty	2080S0012X	Specialty	2086S0102X	Specialty
207ND0900X	Specialty	204E00000X	Specialty	2080S0010X	Specialty	2086X0206X	Specialty
207ND0101X	Specialty	207XS0114X	Specialty	202K00000X	Specialty	2086S0127X	Specialty
207NP0225X	Specialty	207XX0004X	Specialty	2081H0002X	Specialty	2086S0129X	Specialty
207NS0135X	Specialty	207XS0106X	Specialty	2081N0008X	Specialty	208600000X	Specialty
207N00000X	Specialty	207XS0117X	Specialty	2081P2900X	Specialty	208G00000X	Specialty
204R00000X	Specialty	207XX0801X	Specialty	2081P0010X	Specialty	204F00000X	Specialty
207PE0004X	Specialty	207XP3100X	Specialty	2081P0004X	Specialty	2088P0231X	Specialty
207PH0002X	Specialty	207XX0005X	Specialty	2081S0010X	Specialty	208800000X	Specialty
193200000X	Primary	207PT0002X	Specialty	207X00000X	Specialty	208100000X	Specialty
208D00000X	Primary	207PP0204X	Specialty	207YS0123X	Specialty	2082S0099X	Specialty
363A00000X	Primary	207PS0010X	Specialty	207YX0602X	Specialty	2082S0105X	Specialty
363AM0700X	Primary	207PE0005X	Specialty	207YX0905X	Specialty	208200000X	Specialty
363L00000X	Primary	207P00000X	Specialty	207YX0901X	Specialty	2083T0002X	Specialty
363LA2100X	Primary	207QA0401X	Specialty	207YP0228X	Specialty	2083X0100X	Specialty

(continued)

Table F-2-3. Primary/specialty care taxonomies (continued)

Taxonomy	Type	Taxonomy	Type	Taxonomy	Type	Taxonomy	Type
363LA2200X	Primary	207QB0002X	Specialty	207YX0007X	Specialty	2083P0500X	Specialty
363LC1500X	Primary	207QH0002X	Specialty	207YS0012X	Specialty	2083P0901X	Specialty
363LF0000X	Primary	207QS1201X	Specialty	207Y00000X	Specialty	2083S0010X	Specialty
363LG0600X	Primary	207QS0010X	Specialty	208VP0014X	Specialty	2083P0011X	Specialty
363LP2300X	Primary	207RA0401X	Specialty	208VP0000X	Specialty	2084A0401X	Specialty
363LW0102X	Primary	207RA0201X	Specialty	207ZP0101X	Specialty	2084P0802X	Specialty
364S00000X	Primary	207RB0002X	Specialty	207ZP0102X	Specialty	2084B0002X	Specialty
364SA2100X	Primary	207RC0000X	Specialty	207ZB0001X	Specialty	2084P0804X	Specialty
364SA2200X	Primary	207RI0001X	Specialty	207ZP0104X	Specialty	2084N0600X	Specialty
364SC0200X	Primary	207RC0001X	Specialty	207ZC0006X	Specialty	2084D0003X	Specialty
364SC1501X	Primary	207RC0200X	Specialty	207ZP0105X	Specialty	2084F0202X	Specialty
364SC2300X	Primary	207RE0101X	Specialty	207ZC0500X	Specialty	2084P0805X	Specialty
364SF0001X	Primary	207RG0100X	Specialty	207ZD0900X	Specialty	2084H0002X	Specialty
364SG0600X	Primary	207RH0000X	Specialty	207ZF0201X	Specialty	2084P0005X	Specialty
364SH0200X	Primary	207RH0003X	Specialty	207ZH0000X	Specialty	2084N0400X	Specialty
364SW0102X	Primary	207RI0008X	Specialty	207ZI0100X	Specialty	2084N0402X	Specialty
208000000X	Primary	207RH0002X	Specialty	207ZM0300X	Specialty	2084N0008X	Specialty
2080A0000X	Primary	207RI0200X	Specialty	207ZP0007X	Specialty	2084P2900X	Specialty
207V00000X	Primary	207RI0011X	Specialty	207ZN0500X	Specialty	2084P0800X	Specialty
207VG0400X	Primary	207RM1200X	Specialty	207ZP0213X	Specialty	2084P0015X	Specialty
207VM0101X	Primary	207RX0202X	Specialty	2080C0008X	Specialty	2084S0012X	Specialty
207VX0000X	Primary	207RN0300X	Specialty	2080I0007X	Specialty	2084S0010X	Specialty
207Q00000X	Primary	207RP1001X	Specialty	2080P0006X	Specialty	2084V0102X	Specialty
207QA0000X	Primary	207RR0500X	Specialty	2080H0002X	Specialty	2085B0100X	Specialty
207QA0505X	Primary	207RS0012X	Specialty	2080T0002X	Specialty	2085D0003X	Specialty
207QG0300X	Primary	207RS0010X	Specialty	2080N0001X	Specialty	2085R0202X	Specialty
207R00000X	Primary	207RT0003X	Specialty	2080P0008X	Specialty	2085U0001X	Specialty
207RA0000X	Primary	207T00000X	Specialty	2080P0201X	Specialty	2085H0002X	Specialty
207RG0300X	Primary	204D00000X	Specialty	2080P0202X	Specialty	2085N0700X	Specialty
207KA0200X	Specialty	204C00000X	Specialty	2080P0203X	Specialty	2085N0904X	Specialty
207KI0005X	Specialty	207UN0903X	Specialty	2080P0204X	Specialty	2085P0229X	Specialty
207K00000X	Specialty	207UN0901X	Specialty	2080P0205X	Specialty	2085R0001X	Specialty
207LA0401X	Specialty	207UN0902X	Specialty	2080P0206X	Specialty	2085R0205X	Specialty
207LC0200X	Specialty	207U00000X	Specialty	2080P0207X	Specialty	2085R0203X	Specialty
207LH0002X	Specialty	207VB0002X	Specialty	2080P0208X	Specialty	2085R0204X	Specialty
207LP2900X	Specialty	207VC0200X	Specialty	2080P0210X	Specialty	2086H0002X	Specialty
207LP3000X	Specialty	207VX0201X	Specialty	2080P0214X	Specialty	2086S0120X	Specialty
207L00000X	Specialty	207VH0002X	Specialty	2080P0216X	Specialty	2086S0122X	Specialty
208C00000X	Specialty	207VE0102X	Specialty	2080T0004X	Specialty	2086S0105X	Specialty

(continued)

Table F-2-3. Primary/specialty care taxonomies (continued)

Taxonomy	Type	Taxonomy	Type	Taxonomy	Type	Taxonomy	Type
207NI0002X	Specialty	207W00000X	Specialty	2080S0012X	Specialty	2086S0102X	Specialty
207ND0900X	Specialty	204E00000X	Specialty	2080S0010X	Specialty	2086X0206X	Specialty
207ND0101X	Specialty	207XS0114X	Specialty	202K00000X	Specialty	2086S0127X	Specialty
207NP0225X	Specialty	207XX0004X	Specialty	2081H0002X	Specialty	2086S0129X	Specialty
207NS0135X	Specialty	207XS0106X	Specialty	2081N0008X	Specialty	208600000X	Specialty
207N00000X	Specialty	207XS0117X	Specialty	2081P2900X	Specialty	208G00000X	Specialty
204R00000X	Specialty	207XX0801X	Specialty	2081P0010X	Specialty	204F00000X	Specialty
207PE0004X	Specialty	207XP3100X	Specialty	2081P0004X	Specialty	2088P0231X	Specialty
207PH0002X	Specialty	207XX0005X	Specialty	2081S0010X	Specialty	208800000X	Specialty

- **Probability of having a readmission within 30 days of hospital discharge:** This is a dichotomous variable indicating beneficiaries with at least one hospital readmission within 30 days of a live discharge. This measure was created only for individuals who were 18 years or older. An index hospital discharge is identified as an inpatient stay with a discharge date within the given measurement period (12 months) minus 30 days from the end of the period. We excluded admissions if the beneficiary died during the hospitalization or did not have 30 days of post-discharge enrollment.

Expenditures

All expenditures outcomes were summed annually then transformed to per beneficiary per month (PBPM) amounts by dividing by 12. Payments were not risk adjusted[39] or price standardized across geographic areas. Negative payments were retained for all payment categories except for total payments where they were set to zero before summing. Depending on the type of claim, claims were included in a period's total if discharge or service date on the claim was during the period.

- **Total PBPM**. Overall net payment amounts for all inpatient, outpatient (facility and professional), and prescription drug claims. Overall sums were calculated at the annual level then transformed to PBPM level by dividing by 12. Averages include all individuals enrolled during the period, so that the figures also reflect the presence of individuals with zero medical costs. The payments were not risk adjusted[40] or price standardized across geographic areas. Negative payments on claims were set to zero for total expenditures. Depending on the type of claim, claims were included in a period's total if discharge or service date on the claim was during the period.

[39] Although the expenditures were not formally risk adjusted, the comparison groups were weighted by the propensity score (see ***Section F-2.3***), which includes some risk adjustment measures.
[40] Although the expenditures were not formally risk adjusted, the comparison groups were weighted by the propensity score (see ***Section F-2.3***), which includes some risk adjustment measures.

- **Inpatient facility PBPM**. Overall net payment amounts for all inpatient claims where claim type is not equal to either M (HCFA 1500) or Y (Professional Crossover).
- **Outpatient facility PBPM**. Overall net payment amounts for all outpatient claims where claim type is not equal to either M (HCFA 1500) or Y (Professional Crossover).
- **Professional PBPM**. Overall net payment amounts for all outpatient or inpatient claims where claim type is equal to either M (HCFA 1500) or Y (Professional Crossover).
- **Pharmacy PBPM**. Overall net payment amounts for all pharmacy claims.

Quality of care

To evaluate the impact on quality of care, we report the following quality measures. The measures were calculated annually for all eligible beneficiaries in the Test state and comparison group.

- **The percentage of adolescent and adult patients ages 13–64 years with a new episode of alcohol and other drug (AOD) dependence who initiated treatment through an inpatient AOD admission, outpatient visit, intensive outpatient encounter, or partial hospitalization within 14 days of the diagnosis.** This is the percentage of adolescent and adult patients with a new episode of AOD dependence who initiated treatment within 14 days of the diagnosis. To identify patients, the patient had to be 13–64 years old and have at least one of the episodes listed below during the intake period (January 1–November 15 of the measurement year). Episodes were identified using Value Sets in the HEDIS measure Initiation and Engagement of Alcohol and Other Drug Dependence Treatment (IET).
  - At least one outpatient visit, intensive outpatient encounter, or partial hospitalization with a diagnosis of AOD
  - At least one detoxification visit
  - At least one ED visit with a diagnosis of AOD
  - At least one acute or nonacute inpatient discharge with either a diagnosis of AOD or an AOD procedure code

  The episode with the earliest start date was identified as the index episode.

  Patients who had a claim with any diagnosis of AOD during the 60 days prior to the index episode were excluded from the denominator.

  For individuals who met the above new episode of AOD criteria, we flagged whether they fulfilled initiation of AOD treatment through an inpatient AOD admission, outpatient visit, intensive outpatient encounter, or partial hospitalization within 14 days of the diagnosis and calculated a percentage (number of patients who initiated AOD treatment/number of index episodes).

  If the index episode and the initiation treatment event occur on the same day, they must have been with different providers for the initiation treatment event to count. If the index

episode was an inpatient discharge, the inpatient stay is considered initiation of treatment. If the index episode was an outpatient, intensive outpatient, partial hospitalization, detoxification, or ED visit, the patient must have at least one of the episodes listed below within 14 days of the index episode. Episodes were identified using Value Sets in the HEDIS measure.

  - At least one acute or nonacute inpatient discharge with a diagnosis of AOD
  - At least one outpatient visit, intensive outpatient encounter, or partial hospitalization with a diagnosis of AOD

Patients with an initiation treatment event of an inpatient stay with a discharge date after December 1 of the measurement year were excluded from the denominator.

- **The percentage of adolescent and adult patients ages 13–64 years with a new episode of AOD dependence who initiated treatment and who had engaged in two or more additional services with an AOD diagnosis within 30 days of the initiation visit.** This is the percentage of adolescent and adult patients with a new episode of AOD dependence who initiated treatment within 14 days of the diagnosis and who had two or more additional services with an AOD diagnosis within 30 days of the initiation visit. To identify patients, the patient had to meet the same new episode of AOD criteria included in the AOD measure above.

For individuals who met the above new episode of AOD criteria, we flagged whether they fulfilled the following engagement of AOD criteria and calculated the percentage (number of patients who initiated AOD treatment *and* had two or more AOD engagement visits/number of index episodes):

  - Initiated treatment through an inpatient AOD admission, outpatient visit, intensive outpatient encounter, or partial hospitalization within 14 days of the diagnosis as stipulated in AOD initiation measure above. Patients with an initiation treatment event of an inpatient stay with a discharge date after December 1 of the measurement year were excluded from the denominator.
  - Had two or more inpatient admissions, outpatient visits, intensive outpatient encounters, or partial hospitalizations with any AOD diagnosis, beginning on the day after the initiation treatment event through 29 days after the initiation treatment event (29 days total). Multiple engagement visits may occur on the same day, but they must have been with different providers to count as unique engagement visits. Visits were identified using Value Sets in the HEDIS measure IET.

- **Probability of having an admission for an Ambulatory Care Sensitive Condition based on the Prevention Quality Indicators (PQI) composite of chronic conditions (Agency for Healthcare Research and Quality, 2016):** The denominator includes the Vermont non-dual Medicaid population ages 18 and older. The numerator includes discharges that meet the inclusion and exclusion rules for any of the nine conditions in the PQI Chronic Composite (PQI #92) (Agency for Healthcare Research and Quality, 2016).

  - PQI #01 Diabetes Short-Term Complications Admission Rate
  - PQI #03 Diabetes Long-Term Complications Admission Rate
  - PQI #05 Chronic Obstructive Pulmonary Disease or Asthma in Older Adults Admission Rate
  - PQI #07 Hypertension Admission Rate
  - PQI #08 Heart Failure Admission Rate
  - PQI #13 Angina Without Procedure Admission Rate
  - PQI #14 Uncontrolled Diabetes Admission Rate
  - PQI #15 Asthma in Younger Adults Admission Rate
  - PQI #16 Rate of Lower-Extremity Amputation among Patients with Diabetes
- **Percentage of children ages 1–3 years screened for risk of developmental, behavioral, and social delays using a standardized screening tool in the 12 months preceding their first 3 years of life.** This is the percentage of members ages 1–3 years during the year who had at least one screening for risk of developmental, behavioral, and social delays using a standardized tool during the 12 months preceding their first 3 years of life. A screening counts as a screening for risk of developmental, behavioral, and social delays if the claim includes a CPT code of 96110 (developmental testing, with interpretation and report).
- **Percentage of adolescents 12–21 years of age who had at least one comprehensive well-care visit with a primary care provider (PCP) or OB/GYN practitioner.** This is the percentage of adolescents ages 12–21 years who had at least one comprehensive well-care visit with a PCP or OB/GYN practitioner during the year. A visit counts as a well-care visit if the claim includes a well-care visit CPT, HCPCS, or diagnosis code (CPT code = 99381–99385, 99391–99395, 99461, or HCPCS code = G0438, G0439, or ICD-9 diagnosis code = V20.2, V20.31, V20.32, V70.0, V70.3, V70.5, V70.6, V70.8, V70.9, or ICD-10 diagnosis code = Z00.00, Z00.01, Z00.110, Z00.111, Z00.121, Z00.129, Z00.5, Z00.8, Z02.0–Z02.6, Z02.71, Z02.79, Z02.81, Z02.82, Z02.83, Z02.89, Z02.9).

F-2.1.3 Population studied

In the report, the Medicaid SSP group consists of all Medicaid beneficiaries attributed to Medicaid SSP providers between 2014 to 2016. The beneficiary designations to Medicaid SSP participating providers or others were provided to us from the state agency responsible for attribution. We considered someone ever attributed during the 3-year intervention period to be part of the overall Medicaid SSP group. In addition to the overall group, we conducted analyses on specific subpopulations:

- Adults (age > 18)
- Children (age <= 18)

- Beneficiaries with behavioral health conditions (in year prior to prior attribution)
  - Behavioral health condition defined using diagnosis code found in the HEDIS "Mental Health Diagnosis" or "Chemical Dependency" value sets
- Beneficiaries attributed in all 3 years of the Medicaid SSP
- Beneficiaries attributed to CHAC (OneCare) in their first attribution year

F-2.1.4 Comparison group and propensity score weighting

For the impact analysis, we are using a pre-post comparison group design, in which the comparison group provides an estimate of what would have happened among Medicaid SSP beneficiaries absent the Medicaid SSP. The difference in the changes over time from the pre-period to the intervention period between SSP beneficiaries and their comparison group provides an estimate of the impact of the Medicaid SSP. The comparison group should be similar to the Medicaid SSP group on all relevant dimensions (e.g., demographic, socioeconomic, political, regulatory, and health and health systems) except for the policy change being tested.

In the following section, we detail the procedures we used to select the comparison group for the Medicaid SSP in Vermont.

Selection of comparison group

In 2014 and 2015, as part of their Medicaid SSP program the DVHA assigned eligible Medicaid beneficiaries to primary care providers using a specifically designed attribution algorithm. This algorithm was run for all SSP-eligible Medicaid beneficiaries in the state. In this analysis, the Medicaid SSP group comprises beneficiaries who were attributed to providers associated with either of the two ACOs participating in the SSP—OneCare and CHAC. The comparison group comprises the remaining attributed beneficiaries who were assigned to either the remaining ACO not participating in the Medicaid SSP—VCP—or to unaffiliated independent physicians. Because attribution is performed independently each year, beneficiaries attributed to OneCare or CHAC in either 2014 or 2015 belong to the Medicaid SSP group. Thus, beneficiaries who were attributed but never to OneCare or CHAC represent the comparison group.

Calculation of person-level weights

To balance the population characteristics for the claims-based analyses, we estimated propensity scores for all individuals from the comparison group. A propensity score is the probability that an individual is in the Medicaid SSP group rather than the comparison group.

The objective of propensity score modeling is to create a weighted comparison group with characteristics equivalent to those for the Medicaid SSP population. To the extent that these characteristics are correlated with expenditure, utilization, and quality outcomes, propensity weighting will help balance pre-SSP levels of the outcomes as well.

There are other methods to apply propensity scores to an analysis. Aside from weighting, one frequently used method is matching, whereby a Medicaid SSP beneficiary is matched to a comparison group beneficiary who has a similar propensity score. Although we considered this method, we decided not to pursue matching for several reasons. First, propensity score weighting has been shown to produce less biased estimates, less modeling error (e.g., mean squared error, type 1 error), and more accurate variance estimation and confidence intervals when modeling dichotomous outcomes, and this analysis includes many dichotomous utilization and quality of care outcomes. Second, matching may exclude many comparison group beneficiaries from the analysis if a good match cannot be found. Weighting has the advantage of preserving sample size.

Person-level characteristics

The initial step in the process was to select person-level characteristics to be used in each propensity score model. ***Table F-2-4*** shows the characteristics we used grouped by whether they control for demographics, enrollment, attribution, or beneficiary health status. Because there is limited information available in claims data, we considered also including county-level characteristics to control for geographic characteristics such as physician supply and median income to account for potential differences in access to care or other geographic differences. However, we found that there was little variation in county-level characteristics, which made it difficult to balance on these variables. To optimize the balance and to avoid extreme weights, we therefore excluded county-level covariates from the propensity score model. However, we do control for county-level characteristics in the outcome model.

Estimation and weighting procedures

Using the characteristics listed in ***Table F-2-4***, we estimated propensity models by logistic regression, in which the outcome was one for beneficiaries attributed to a Medicaid SSP provider and zero for the comparison group. Separate models were estimated for 2011, 2012, 2013, 2014, 2015, and 2016 data.

Table F-2-4. Covariates for propensity score logistic regressions

Demographic characteristics
Gender
Age <1
Age 1–18
Age 19–64
Disabled
Federal Poverty Level
Enrollment
Enrolled previous year[a]
Current enrollment length (10, 11, or 12 months)
Enrolled as part of Medicaid expansion
Attribution
Attributed via CPT codes (vs. selection or auto-assignment)
Attributed to Blueprint Medical Home
Health status measures
CDPS (lagged or pegged prior to attribution)
Total Medicaid payments (lagged or pegged prior to attribution)
Probability of inpatient admission (lagged or pegged prior to attribution)
Probability of ED visit (lagged or pegged prior to attribution)
Vermont Chronic Care Initiative[b]

CDPS = Chronic Illness and Disability Payment System; CPT = Current Procedural Terminology; ED = emergency department.

[a] Enrollment in previous year is counted if member was eligible for 10 or more months.

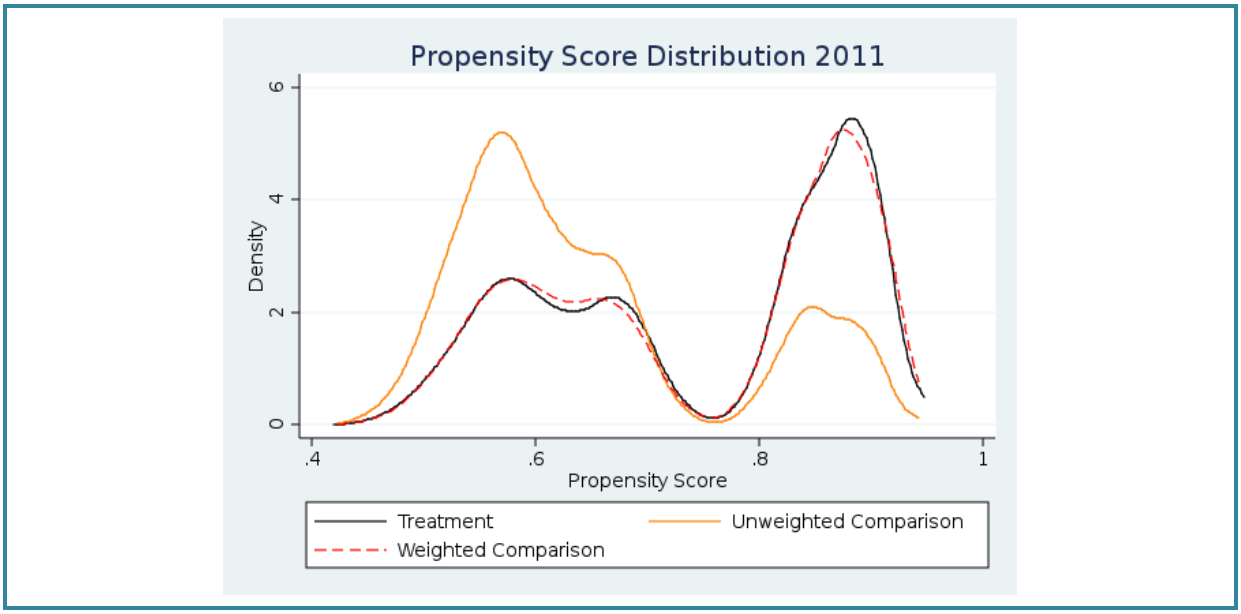
[b] A Vermont Medicaid program that targets members at risk for adverse health outcomes. It provides case management and social support services to improve their health and reduce costs.

We set propensity weights to 1 for all individuals in the Medicaid SSP group. The propensity weight for a comparison individual was a function of his or her predicted propensity score—where weight = p/(1-p), with p the predicted propensity. Our procedure typically includes trimming weights that are either less than 0.05 or greater than 20, although in this analysis no weights needed trimming.

F-2.1.5 Propensity model evaluation

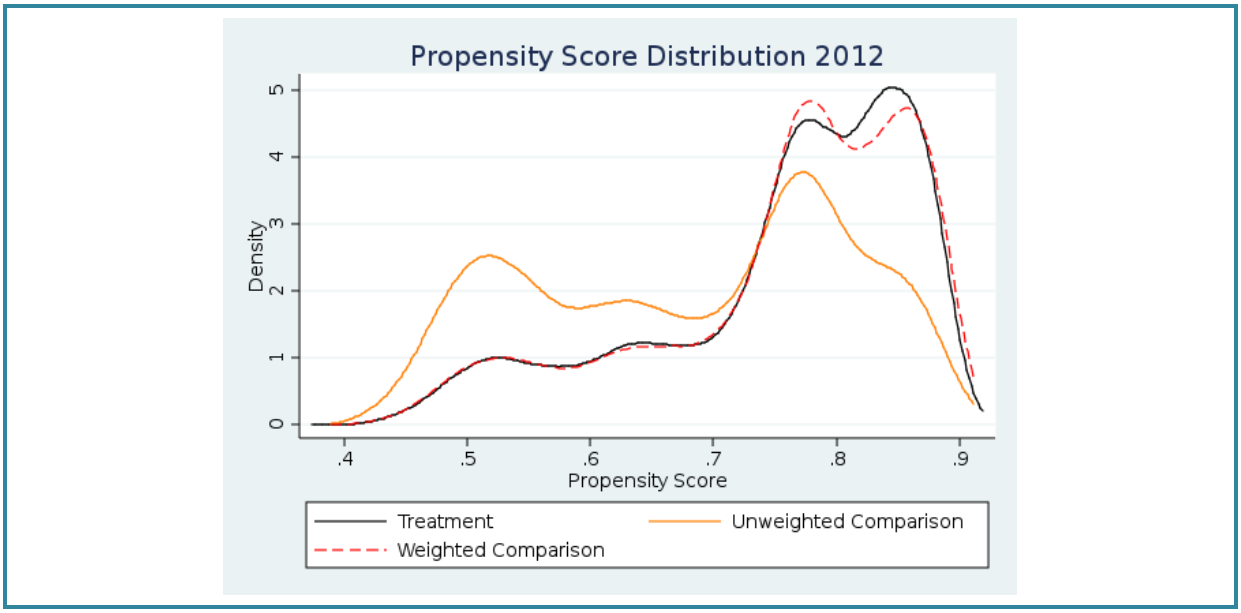
We evaluated several aspects of the propensity score models. First, we examined plots of predicted probabilities to ensure sufficient overlap in the distributions of the Medicaid SSP and comparison groups. This feature, known as common support, is critical because it provides the basis for inferring effects from group comparisons (***Figure F-2-1*** to ***F-2-6***).

Figure F-2-1. Weighted and unweighted propensity score density plots for the Medicaid SSP and comparison groups, 2011[41]



SSP = Shared Savings Program.

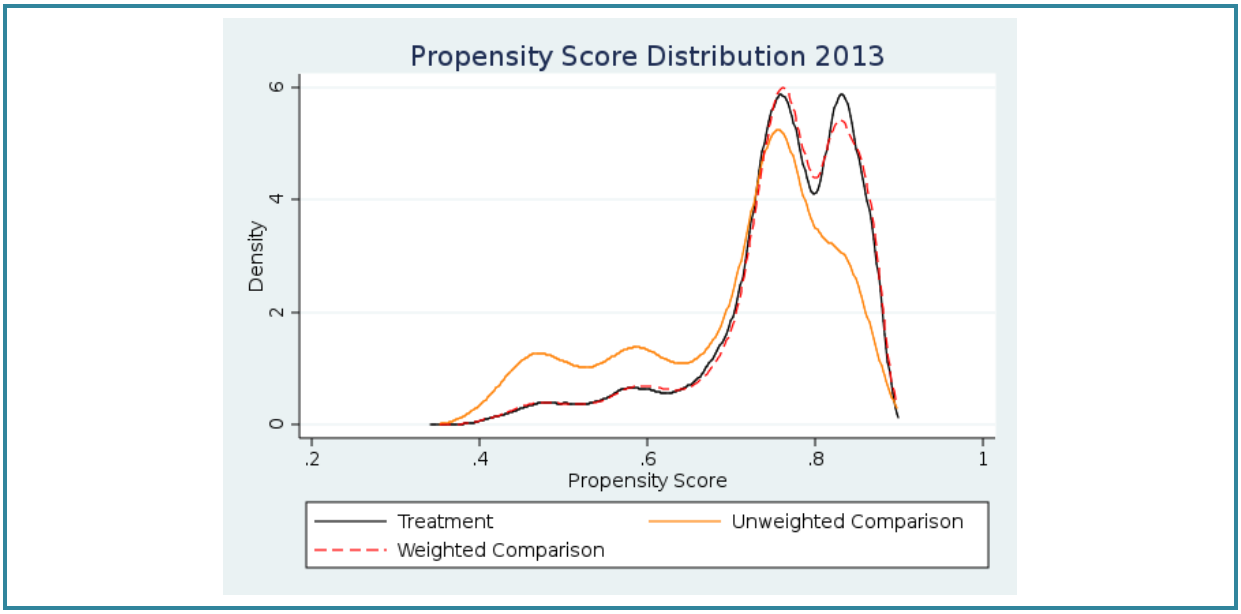
Figure F-2-2. Weighted and unweighted propensity score density plots for the Medicaid SSP and comparison groups, 2012



SSP = Shared Savings Program.

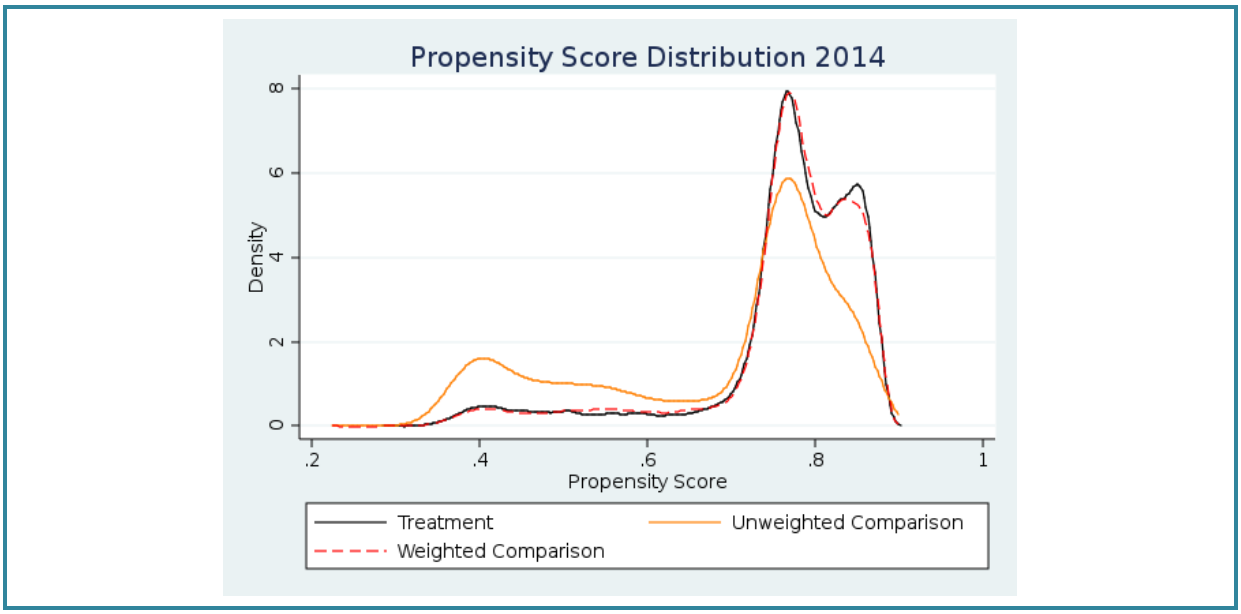
[41] In ***Figures F-2-1*** through ***F-2-5***, the Treatment lines represent those in the SSP group.

Figure F-2-3. Weighted and unweighted propensity score density plots for the Medicaid SSP and comparison groups, 2013



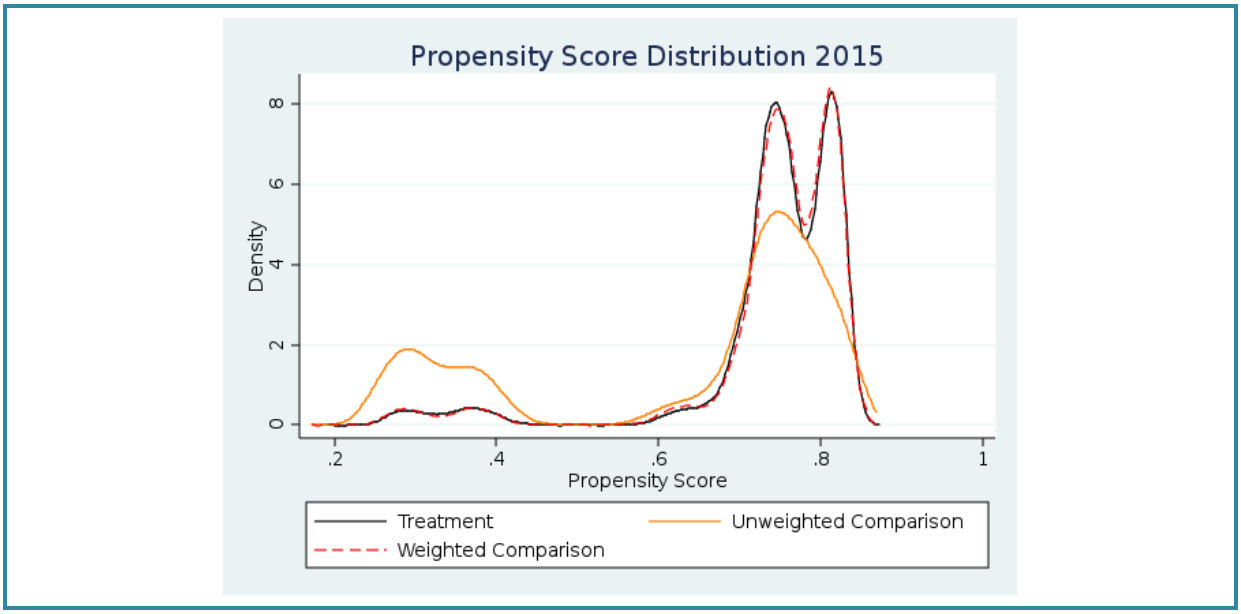
SSP = Shared Savings Program.

Figure F-2-4. Weighted and unweighted propensity score density plots for the Medicaid SSP and comparison groups, 2014



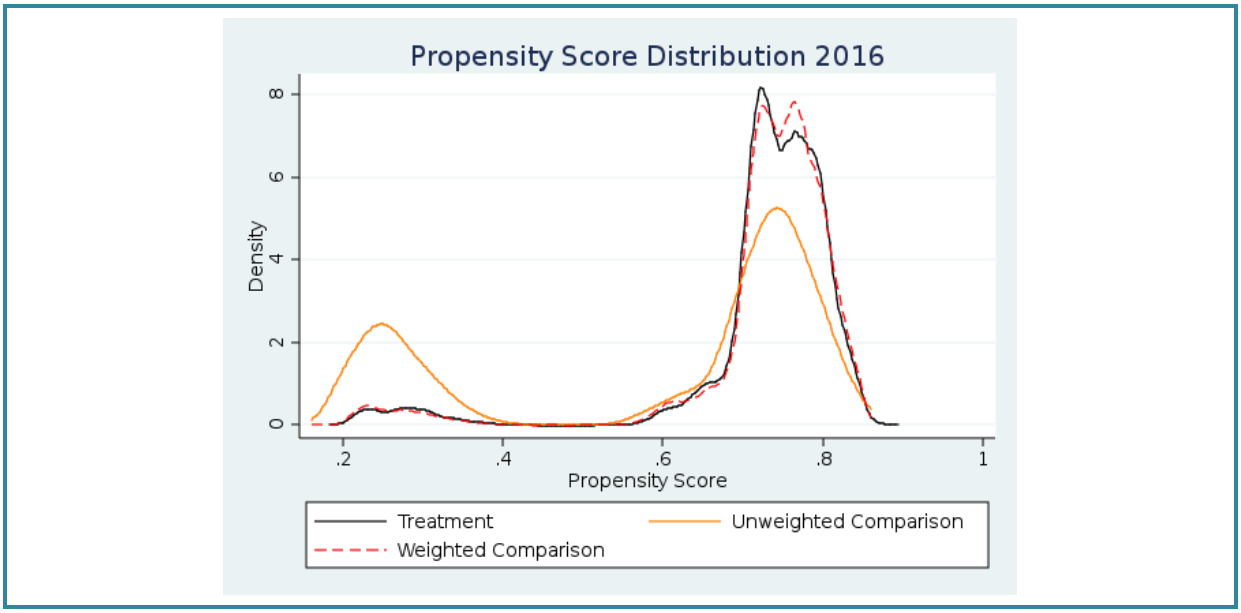
SSP = Shared Savings Program.

Figure F-2-5. Weighted and unweighted propensity score density plots for the Medicaid SSP and comparison groups, 2015



SSP = Shared Savings Program.

Figure F-2-6. Weighted and unweighted propensity score density plots for the Medicaid SSP and comparison groups, 2016



SSP = Shared Savings Program.

In all years, we found the comparison group passed the common support assumption ($P(D = 1|X)>0$) for almost the entire range of the Medicaid SSP group's propensity scores. The only exceptions were in the uppermost percentiles of the Medicaid SSP group's distribution (above the 99th percentile). These plots provide ample evidence that the common support assumption is upheld.

Second, we compared the logistic results of the models to see which variables had the greatest impact on the propensity score weights. We found that the major differences between the groups were in their age, their socioeconomic status, and their proportion of beneficiaries attributed to Blueprint medical homes. Overall, we found that Medicaid SSP beneficiaries were much more likely to be attributed to Blueprint medical homes, were somewhat more likely to be adults, and on average had lower incomes and resided in relatively more rural areas. We found these differences to be fairly consistent and stable over time.

Finally, we compared unweighted and propensity-weighted means for the characteristics in the model. As expected, we found that, after weighting, the comparison group means were within a few percentage points of the values for the Medicaid SSP group.

Tables F-2-5 to ***F-2-10*** show unweighted and (propensity score) weighted means/proportions for 2011–2016. The notable group differences in the unweighted samples—age, attribution, and socioeconomic factors—are substantially mitigated post-weighting as evidenced by the minimized standardized differences.

Table F-2-5. Unweighted and weighted means and standardized differences, Medicaid SSP and comparison groups, 2011

Characteristic	Unweighted			Weighted			
	Medicaid SSP group	Comparison group	Standardized difference[a]	Medicaid SSP group	Comparison group	Standardized difference[a]	p-value
N	36,688	14,252		36,688	36,744		
Female	53.3	52.1	2.5	53.3	53.9	1.2	0.09
Age <1	0.9	1.2	3.0	0.9	1.1	1.9	0.01
Age 1–18	58.7	71.1	26.1	58.7	60.5	3.5	<0.001
Age 19–64	40.4	27.7	26.9	40.4	38.5	3.9	<0.001
Disabled	15.1	10.8	12.7	15.1	15.1	0.2	0.81
Federal Poverty Level	65.6	77.0	16.7	65.6	63.4	3.4	<0.001
Enrolled previous year[a]	0	0	.	0	0	.	.
Enrolled 10 months out of year	3.8	3.3	2.3	3.6	4.0	1.2	0.10
Enrolled 11 months out of year	5.4	5.1	1.4	5.4	5.3	0.3	0.64
Enrolled as part of Medicaid expansion							
Attributed via CPT codes (vs. selection or auto-assignment)	79.1	81.2	4.1	79.6	79.8	0.4	0.59
Attributed to Blueprint Medical Home	53.7	21.0	71.3	53.5	53.7	0.5	0.53
CDPS (lagged or pegged prior to attribution)	1.2	1.2	1.4	1.2	1.3	2.0	0.01
Vermont Chronic Care Initiative	28.2	23.5	10.7	28.2	27.8	0.9	0.22

CDPS = Chronic Illness and Disability Payment System; CPT = Current Procedural Terminology; ED = emergency department; SSP = Shared Savings Program.

[a] Absolute standardized differences are expressed as percentages.

Table F-2-6. Unweighted and weighted means and standardized differences, Medicaid SSP and comparison groups, 2012

Characteristic	Unweighted			Weighted			
	Medicaid SSP group	Comparison group	Standardized difference[a]	Medicaid SSP group	Comparison group	Standardized difference[a]	p-value
N	52,227	18,621		52,227	52,329		
Female	53.6	52.7	1.8	53.6	54.4	1.6	0.01
Age <1	0.8	1.1	2.9	0.8	1.0	2.2	<0.001
Age 1–18	55.5	66.1	21.7	55.5	56.0	1.0	0.10
Age 19–64	43.6	32.8	22.4	43.6	42.9	1.5	0.02
Disabled	14.2	10.9	10.1	14.2	14.0	0.7	0.29
Federal Poverty Level	66.7	76.0	14.1	66.7	65.5	1.9	<0.001
Enrolled previous year[a]	65.3	71.6	13.6	65.3	64.3	2.0	<0.001
Enrolled 10 months out of year	4.3	4.1	1.0	4.3	4.5	1.0	0.11
Enrolled 11 months out of year	6.0	6.0	0.0	6.0	6.2	0.7	0.23
Enrolled as part of Medicaid expansion	0	0		0	0		
Attributed via CPT codes (vs. selection or auto-assignment)	79.6	81.1	3.6	79.6	79.8	0.4	0.50
Attributed to Blueprint Medical Home	72.4	47.4	53.0	72.4	72.5	0.2	0.73
CDPS (lagged or pegged prior to attribution)	1.3	1.2	1.0	1.3	1.3	1.3	0.03
Total Medicaid payments (lagged or pegged prior to attribution)	4,344.7	4,151.4	1.7	4,344.7	4,444.0	0.9	0.16
Probability of inpatient admission (lagged or pegged prior to attribution)	3.9	3.4	2.7	3.9	4.0	0.9	0.16
Probability of ED visit (lagged or pegged prior to attribution)	23.5	22.6	2.1	23.5	23.8	0.6	0.38
Vermont Chronic Care Initiative	13.6	11.5	6.2	13.6	13.9	0.8	0.19

CDPS = Chronic Illness and Disability Payment System; CPT = Current Procedural Terminology; ED = emergency department; SSP = Shared Savings Program.
[a] Absolute standardized differences are expressed as percentages.

Table F-2-7. Unweighted and weighted means and standardized differences, Medicaid SSP and comparison groups, 2013

Characteristic	Unweighted			Weighted			
	Medicaid SSP group	Comparison group	Standardized difference[a]	Medicaid SSP group	Comparison group	Standardized difference[a]	p-value
N	64,643	21,975		64,643	64,796		
Female	53.1	52.4	1.4	53.1	53.3	0.4	0.53
Age <1	0.8	0.9	1.4	0.8	0.8	0.9	0.12
Age 1–18	53.2	63.4	21.0	53.2	54.6	2.8	<0.001
Age 19–64	46.1	35.7	21.3	46.1	44.6	3.0	<0.001
Disabled	14.3	11.1	9.4	14.3	14.0	0.9	0.13
Federal Poverty Level	67.0	77.5	15.0	67.0	66.9	0.2	0.73
Enrolled previous year[a]	74.3	77.6	7.6	74.3	72.7	3.6	<0.001
Enrolled 10 months out of year	4.4	4.3	0.6	4.4	4.5	0.6	0.28
Enrolled 11 months out of year	5.9	5.5	1.7	5.9	6.1	0.6	0.30
Enrolled as part of Medicaid expansion	0	0		0	0		
Attributed via CPT codes (vs. selection or auto-assignment)	79.4	81.1	4.2	79.4	79.69	0.8	0.17
Attributed to Blueprint Medical Home	81.4	63.0	41.8	81.4	81.6	0.6	0.30
CDPS (lagged or pegged prior to attribution)	1.3	1.2	1.4	1.3	1.3	0.9	0.10
Total Medicaid payments (lagged or pegged prior to attribution)	4,960.7	4,551.9	3.3	4,960.7	4,954.4	0.1	0.93
Probability of inpatient admission (lagged or pegged prior to attribution)	4.5	3.6	4.6	4.5	4.5	0.3	0.59
Probability of ED visit (lagged or pegged prior to attribution)	27.3	24.3	6.8	27.3	26.7	1.4	0.01
Vermont Chronic Care Initiative	16.4	14.3	5.9	16.4	16.6	0.6	0.28

CDPS = Chronic Illness and Disability Payment System; CPT = Current Procedural Terminology; ED = emergency department; SSP = Shared Savings Program.

[a] Absolute standardized differences are expressed as percentages.

Table F-2-8. Unweighted and weighted means and standardized differences, Medicaid SSP and comparison groups, 2014

Characteristic	Unweighted			Weighted			
	Medicaid SSP group	Comparison group	Standardized difference[a]	Medicaid SSP group	Comparison group	Standardized difference[a]	p-value
N	86,241	29,538		86,241	85,594		
Female	52.2	52.3	0.3	52.16	52.2	0.1	0.88
Age <1	0.5	0.8	2.8	0.5	0.6	0.8	0.11
Age 1–18	45.6	54.7	18.3	45.6	47.2	3.3	<0.001
Age 19–64	53.9	44.5	18.8	53.9	52.16	3.4	<0.001
Disabled	13.3	11.2	6.4	13.3	12.85	1.3	0.01
Federal Poverty Level	70.8	80.1	12.0	70.8	70.2	0.8	0.11
Enrolled previous year[a]	73.5	72.4	2.6	73.5	72.4	2.5	<0.001
Enrolled 10 months out of year	4.4	5.0	2.8	4.4	4.75	1.5	<0.001
Enrolled 11 months out of year	3.6	3.8	1.3	3.6	3.7	0.8	0.12
Enrolled as part of Medicaid expansion	23.4	17.4	15.1	23.4	22.8	1.5	<0.001
Attributed via CPT codes (vs. selection or auto-assignment)	76.0	78.4	5.7	76.0	76.38	0.9	0.06
Attributed to Blueprint Medical Home	80.7	58.9	48.9	80.7	80.78	0.2	0.71
CDPS (lagged or pegged prior to attribution)	1.2	1.2	1.1	1.2	1.2	0.9	0.07
Total Medicaid payments (lagged or pegged prior to attribution)	5,020.4	4,583.1	3.3	5,020.4	5,088.1	0.5	0.29
Probability of inpatient admission (lagged or pegged prior to attribution)	4.7	3.6	5.4	4.7	4.7	0.1	0.78
Probability of ED visit (lagged or pegged prior to attribution)	25.9	22.2	8.7	25.9	25.7	0.5	0.32
Vermont Chronic Care Initiative	16.5	14.5	5.5	16.5	16.5	0.2	0.68

CDPS = Chronic Illness and Disability Payment System; CPT = Current Procedural Terminology; ED = emergency department; SSP = Shared Savings Program.
[a] Absolute standardized differences are expressed as percentages.

Table F-2-9. Unweighted and weighted means and standardized differences, Medicaid SSP and comparison groups, 2015

Characteristic	Unweighted			Weighted			
	Medicaid SSP group	Comparison group	Standardized difference[a]	Medicaid SSP group	Comparison group	Standardized difference[a]	p-value
N	83,813	33,885		83,813	83,977		
Female	51.5	52.2	1.3	51.5	51.5	0.0	0.97
Age <1	0.5	0.7	2.3	0.5	0.6	1.3	0.01
Age 1–18	44.4	51.2	13.5	44.4	44.7	0.6	0.24
Age 19–64	55.1	48.2	13.8	55.1	54.7	0.8	0.12
Disabled	11.6	10.0	5.4	11.6	11.5	0.5	0.35
Federal Poverty Level	69.3	75.4	7.5	69.3	68.5	0.9	0.05
Enrolled previous year[a]	83.9	82.0	5.2	83.9	82.7	3.2	<0.001
Enrolled 10 months out of year	1.3	1.5	1.1	1.3	1.4	0.9	0.06
Enrolled 11 months out of year	2.3	2.3	0.1	2.3	2.6	1.7	<0.001
Enrolled as part of Medicaid expansion	32.0	27.9	9.0	32.0	32.3	0.5	0.30
Attributed via CPT codes (vs. selection or auto-assignment)	75.4	78.1	6.4	75.4	75.8	0.8	0.12
Attributed to Blueprint Medical Home	82.2	58.9	52.8	82.2	82.4	0.5	0.30
CDPS (lagged or pegged prior to attribution)	1.2	1.1	1.0	1.2	1.2	0.2	0.76
Total Medicaid payments (lagged or pegged prior to attribution)	4,611.1	4,365.7	2.0	4,611.1	4,648.1	0.3	0.54
Probability of inpatient admission (lagged or pegged prior to attribution)	4.3	3.5	4.4	4.3	4.3	0.1	0.86
Probability of ED visit (lagged or pegged prior to attribution)	24.6	21.0	8.5	24.6	24.3	0.8	0.10
Vermont Chronic Care Initiative	14.9	13.6	3.7	14.9	15.0	0.1	0.82

CDPS = Chronic Illness and Disability Payment System; CPT = Current Procedural Terminology; ED = emergency department; SSP = Shared Savings Program.
[a] Absolute standardized differences are expressed as percentages.

Table F-2-10. Unweighted and weighted means and standardized differences, Medicaid SSP and comparison groups, 2016

Characteristic	Unweighted			Weighted			
	Medicaid SSP group	Comparison group	Standardized difference[a]	Medicaid SSP group	Comparison group	Standardized difference[a]	p-value
N	63,454	29,052		63,454	63,720		
Female	51.5	52.0	1.13	51.5	51.47	0.02	0.97
Age <1	0.3	0.4	1.7	0.3	0.5	2.3	<0.001
Age 1–18	50.5	56.9	12.8	50.5	50.3	0.5	0.39
Age 19–64	49.2	42.7	13.0	49.2	49.3	0.2	0.73
Disabled	7.6	6.2	5.6	7.6	7.4	0.7	0.22
Federal Poverty Level	69.3	74.7	6.3	69.3	68.2	1.4	0.01
Enrolled previous year[a]	84.5	86.2	4.9	84.5	83.4	2.9	<0.001
Enrolled 10 months out of year	3.5	3.4	0.6	3.5	3.6	0.9	0.09
Enrolled 11 months out of year	3.9	3.7	0.6	3.9	3.9	0.2	0.68
Enrolled as part of Medicaid expansion	34.1	30.3	8.3	34.1	34.9	1.7	<0.001
Attributed via CPT codes (vs. selection or auto-assignment)	78.7	81.2	6.3	78.7	79.0	0.9	0.11
Attributed to Blueprint Medical Home	84.0	58.5	58.7	84.0	84.2	0.6	0.29
CDPS (lagged or pegged prior to attribution)	1.1	1.1	1.9	1.1	1.1	0.0	0.96
Total Medicaid payments (lagged or pegged prior to attribution)	4,539.4	4,162.7	3.0	4,539.4	4,529.8	0.1	0.89
Probability of inpatient admission (lagged or pegged prior to attribution)	4.2	3.1	5.4	4.2	4.1	0.4	0.50
Probability of ED visit (lagged or pegged prior to attribution)	23.4	19.5	9.4	23.4	22.6	1.7	<0.001

CDPS = Chronic Illness and Disability Payment System; CPT = Current Procedural Terminology; ED = emergency department; SSP = Shared Savings Program.

The Vermont Chronic Care Initiative variable was omitted from the 2016 propensity model because the Department of Vermont Health Access included it in the Medicaid data files for 2011–2015 only.

[a] Absolute standardized differences are expressed as percentages.

F.2.1.6 Propensity model evaluation for subpopulation

In addition to the overall model, we evaluated common support graphs and standardized differences of the propensity score models for the subpopulation analyses. Because age was not a factor in our adult and children subpopulations, we found that the most common attribute to exhibit imbalance between the Medicaid SSP and comparison group was attribution to a Blueprint medical home. In all three cases—adults, children, and beneficiaries with behavioral health conditions—the Medicaid SSP group was overwhelmingly more Blueprint attributed than the comparison group. In the Medicaid SSP group, Blueprint attribution was commonly in the 70 to 80 percent range while in the comparison group it fell in the 50 to 60 percent rage. These differences, however, were erased after applying the propensity score weights.

F-2.1.7 Statistical analysis

Regression model

The underlying assumption in D-in-D models estimation is that trends in the Medicaid SSP group would be similar to that of the comparison group in the absence of the initiative (i.e., that the two were on "parallel paths" prior to the start of the Medicaid SSP).

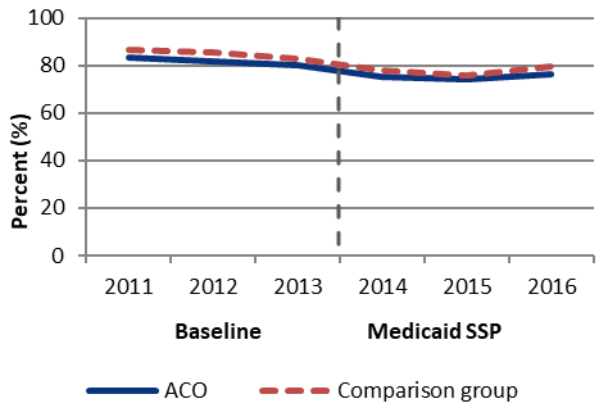
To assess the parallel assumption's validity more empirically, we modeled core expenditure and utilization outcomes during the baseline period with a linear time trend interacted with a dichotomous variable indicating the beneficiary was attributed to a Medicaid SSP provider (i.e., the "test" group). The following section describes the baseline analysis we conducted to inform the D-in-D model.

To examine descriptively whether the trends in the test and comparison group are parallel, we present graphs of annual unadjusted averages for Medicaid SSP-attributed beneficiaries and the comparison group for the baseline period (2011–2013) and the 3 years of the implementation (2014–2016).

Figures F-2-7 to ***F-2-11*** provide the unadjusted averages of the care coordination measures by year.

- The baseline trends were parallel for primary care **provider visits, specialty care provider visits,** and **7-day follow-ups following a mental health hospitalization**, but not parallel for **adults with acute bronchitis not prescribed antibiotic treatment** and **30-day follow-ups following a mental health hospitalization**.

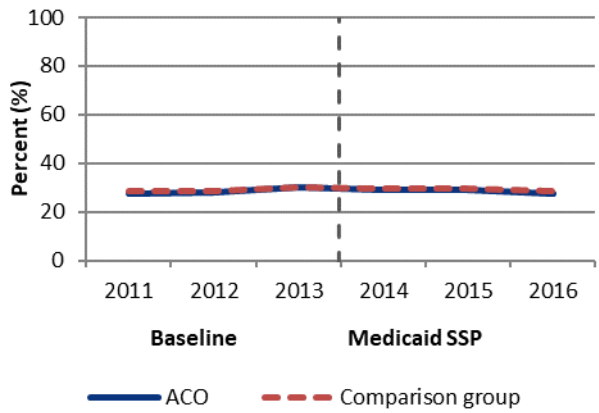
Figure F-2-7. Percentage of beneficiaries with a visit to a primary care provider, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The percentage of beneficiaries with a **primary care provider visit** among SSP-attributed Medicaid beneficiaries and comparison group beneficiaries was similar throughout the baseline and Medicaid SSP implementation period (***Figure F-2-7***). In both groups, the likelihood of a visit was parallel and trended slightly downward during the baseline period, while the likelihood of a visit was fairly stable for both groups during the Medicaid SSP implementation period. The comparison group consistently had a slightly higher percentage of beneficiaries with a primary care provider visit than the SSP-attributed group throughout the baseline and implementation periods.

ACO = accountable care organization; SSP = Shared Savings Program.

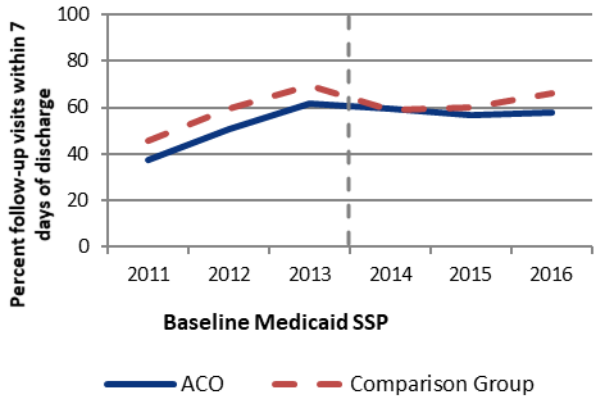
Figure F-2-8. Percentage of beneficiaries with a visit to a specialty care provider, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The likelihood of a **specialty care provider visit** among SSP-attributed Medicaid beneficiaries and comparison group beneficiaries was similar throughout the baseline and Medicaid SSP implementation period (***Figure F-2-8***). In both groups, the likelihood of a visit was fairly stable during the baseline and Medicaid SSP implementation period.

ACO = accountable care organization; SSP = Shared Savings Program.

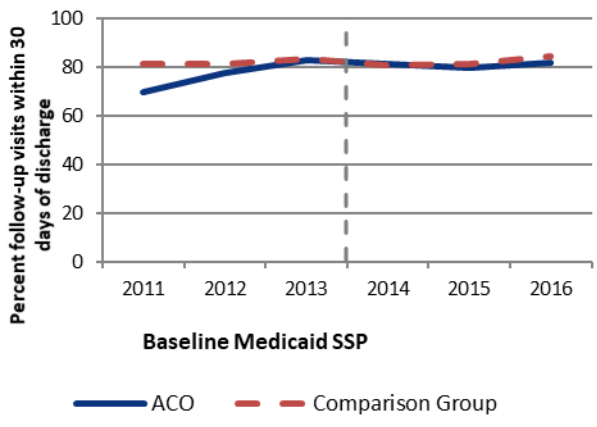
Figure F-2-9. Percentage of Medicaid beneficiaries with a follow-up visit within 7 days of discharge from hospitalization for mental illness, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The percentage of **7-day follow-ups following a mental health hospitalization** trended slightly upward during the baseline period and remained steady during the Medicaid SSP implementation period among SSP-attributed Medicaid beneficiaries and comparison group beneficiaries (***Figure F-2-9***). The comparison group consistently had slightly higher percentages than the Medicaid SSP-attributed group throughout the baseline and implementation periods.

ACO = accountable care organization; SSP = Shared Savings Program.

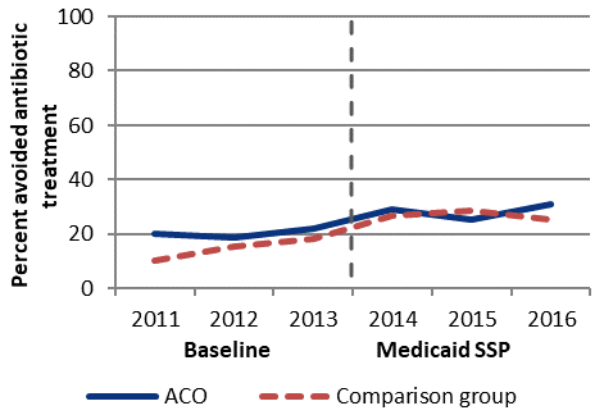
Figure F-2-10. Percentage of Medicaid beneficiaries with a follow-up visit within 30 days of discharge from hospitalization for mental illness, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The percentage of **30-day follow-ups following a mental health hospitalization** among SSP-attributed Medicaid beneficiaries trended upward during the baseline period and remained stable during the implementation period. The percentage among comparison group beneficiaries remained steady throughout the baseline and implementation period (***Figure F-2-10***).

ACO = accountable care organization; SSP = Shared Savings Program.

Figure F-2-11. Percentage of Medicaid adult beneficiaries with acute bronchitis who avoided antibiotic treatment, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



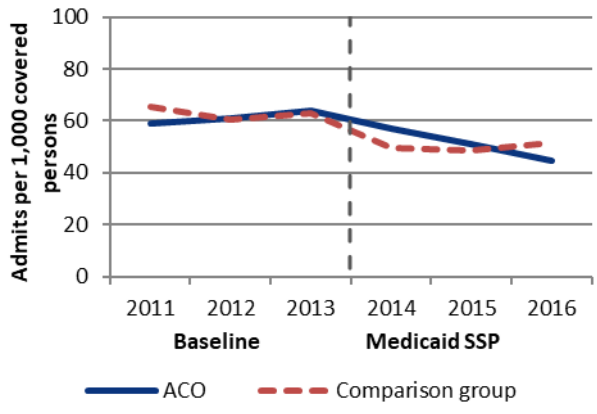
The percentage of **adults with acute bronchitis not prescribed antibiotic treatment** among SSP-attributed Medicaid beneficiaries and comparison group beneficiaries was similar throughout the baseline and Medicaid SSP implementation periods (***Figure F-2-11***). In both groups, the baseline trend was parallel and trending slightly upward, while the percentage remained steady during the implementation period.

ACO = accountable care organization; SSP = Shared Savings Program.

Figures F-2-12 to ***F-2-14*** provide unadjusted annual averages of inpatient admissions, outpatient ED visits, and 30-day readmissions per 1,000 Medicaid beneficiaries by year, respectively.

- The baseline trends were parallel for the **outpatient ED visit rate**, but not parallel for the rate of **acute inpatient admissions** and **30-day readmissions**.

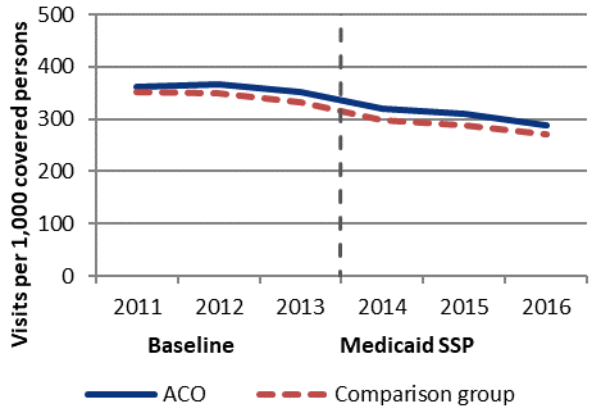
Figure F-2-12. All-cause acute inpatient admissions per 1,000 Medicaid beneficiaries, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The rate of **acute inpatient admissions** among SSP-attributed Medicaid beneficiaries and comparison group beneficiaries was similar throughout the baseline period with the admission rate remaining steady. During the Medicaid SSP implementation period, the admission rate trended downward among SSP-attributed Medicaid beneficiaries and dropped and then trended slightly upward among comparison group beneficiaries (***Figure F-2-12***).

ACO = accountable care organization; SSP = Shared Savings Program.

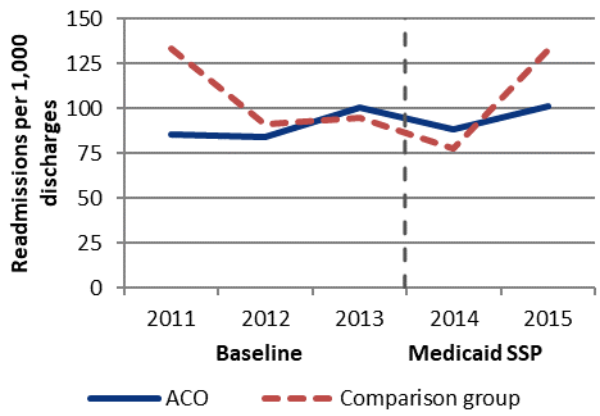
Figure F-2-13. Emergency department visits that did not lead to a hospitalization per 1,000 Medicaid beneficiaries, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The rate of **outpatient ED visits** among SSP-attributed Medicaid beneficiaries and comparison beneficiaries was similar throughout the baseline and Medicaid SSP implementation periods (***Figure F-2-13***). The ED visit rate showed seasonal fluctuations and trended downward for both groups throughout the baseline and implementation periods.

ACO = accountable care organization; ED = emergency department; SSP = Shared Savings Program.

Figure F-2-14. Discharges with a readmission within 30 days per 1,000 Medicaid beneficiaries, 2011 through 2015, Vermont Medicaid SSP-attributed beneficiaries and comparison group

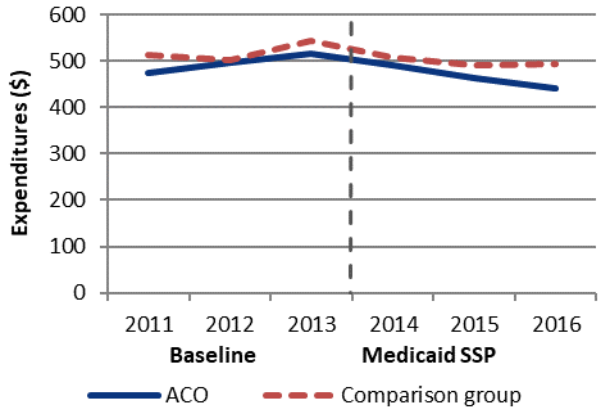


The rate of discharges with a **readmission within 30 days** among SSP-attributed Medicaid beneficiaries remained steady through the baseline and trended slightly upward during the implementation period (***Figure F-2-14***). The rate among comparison group beneficiaries fluctuated throughout the baseline and implementation periods.

ACO = accountable care organization; SSP = Shared Savings Program.

Figures F-2-15 to F-2-19 provide unadjusted annual averages of total, inpatient facility, other facility, professional, and prescription drug PBPM expenditures.

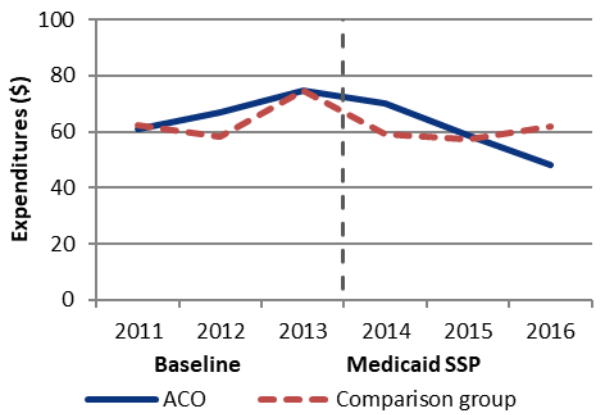
Figure F-2-15. Average total PBPM expenditures, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



For Medicaid beneficiaries, the average **total PBPM expenditures** trend was parallel and remained fairly constant over the baseline and Medicaid SSP implementation periods for the SSP-attributed group and the comparison group (***Figure F-2-15***). The comparison group consistently had slightly higher total PBPM expenditures than the SSP-attributed group throughout the baseline and implementation periods.

ACO = accountable care organization; PBPM = per beneficiary per month; SSP = Shared Savings Program.

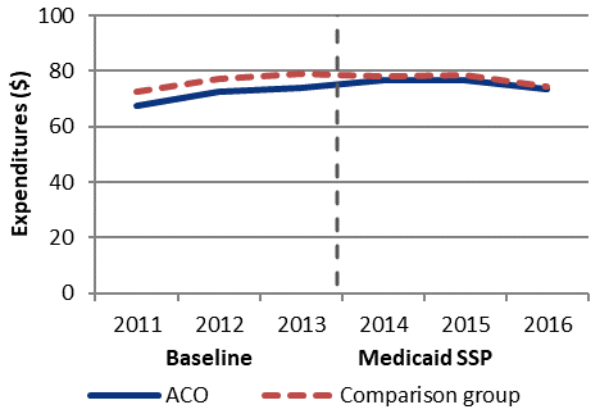
Figure F-2-16. Average inpatient facility PBPM expenditures, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



For SSP-attributed Medicaid beneficiaries, the average **inpatient facility PBPM expenditures** trended upward during the baseline period and trended downward during the Medicaid SSP implementation period (***Figure F-2-16***). Expenditures among the comparison group fluctuated throughout the baseline period and remained fairly steady during the implementation period.

ACO = accountable care organization; PBPM = per beneficiary per month; SSP = Shared Savings Program.

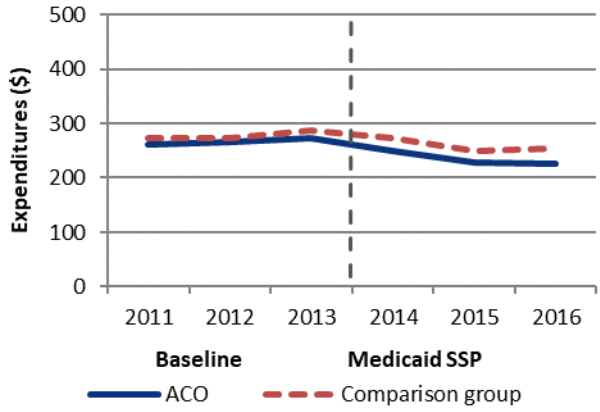
Figure F-2-17. Average other facility PBPM expenditures, 2011 through 2016, Vermont Medicaid SSP-attributed expenditures and comparison group



For Medicaid beneficiaries, the average **other facility PBPM expenditures** trend was parallel and remained fairly constant over the baseline and Medicaid SSP implementation periods for the SSP-attributed group and the comparison group (***Figure F-2-17***). The comparison group consistently had slightly higher other facility PBPM expenditures than the SSP-attributed group throughout the baseline and implementation periods.

ACO = accountable care organization; PBPM = per beneficiary per month; SSP = Shared Savings Program.

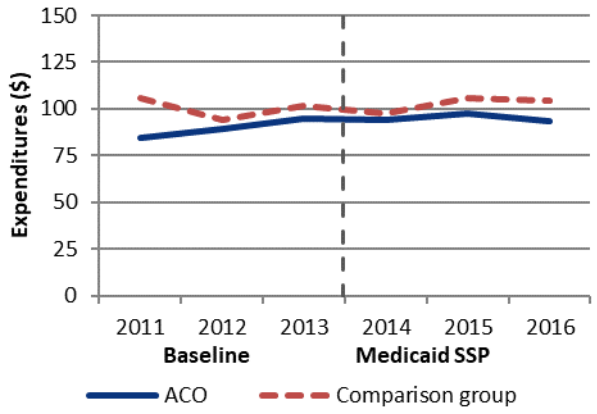
Figure F-2-18. Average professional PBPM expenditures, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



For Medicaid beneficiaries, the average **professional PBPM expenditures** trend was parallel and remained fairly constant over the baseline and Medicaid SSP implementation periods for the SSP-attributed group and the comparison group (***Figure F-2-18***). The comparison group consistently had slightly higher professional PBPM expenditures than the SSP-attributed group throughout the baseline and implementation periods.

ACO = accountable care organization; PBPM = per beneficiary per month; SSP = Shared Savings Program.

Figure F-2-19. Average prescription PBPM expenditures, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



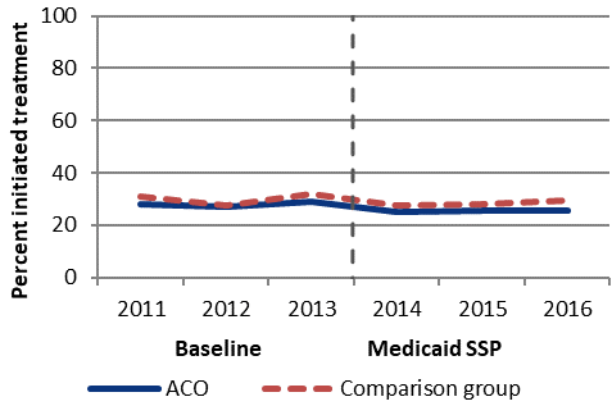
For Medicaid beneficiaries, the average **prescription PBPM expenditures** trend remained fairly constant over the baseline and Medicaid SSP implementation periods for the SSP-attributed group and the comparison group (***Figure F-2-19***). The comparison group consistently had slightly higher professional PBPM expenditures than the SSP-attributed group throughout the baseline and implementation periods.

ACO = accountable care organization; PBPM = per beneficiary per month; SSP = Shared Savings Program.

Figures F-2-20 to ***F-2-24*** provide the unadjusted averages of the quality of care measures by year.

- The baseline trends were parallel for **initiation of treatment after episode of alcohol and other drug dependence**, **engagement of treatment after episode of alcohol and other drug dependence**, and **adolescent well-care visits,** but not parallel for **hospitalizations for ambulatory care sensitive conditions based on chronic PQI** and **developmental screenings**.

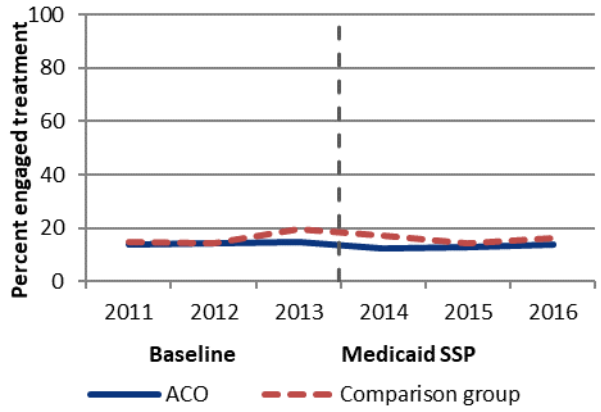
Figure F-2-20. Percentage of Medicaid beneficiaries who initiated treatment after an episode of alcohol and other drug dependence, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The percentage of **initiation of treatment after episode of alcohol and other drug dependence** among SSP-attributed Medicaid beneficiaries and comparison group beneficiaries was similar throughout the baseline and Medicaid SSP implementation periods (***Figure F-2-20***). The percentage for both outcomes remained steady during the baseline and Medicaid SSP implementation period.

ACO = accountable care organization; SSP = Shared Savings Program.

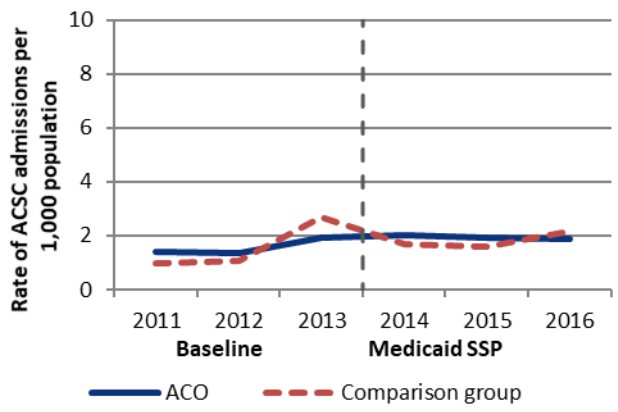
Figure F-2-21. Percentage of Medicaid beneficiaries who engaged treatment after an episode of alcohol and other drug dependence, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The percentage of **engagement of treatment after episode of alcohol and other drug dependence** among SSP-attributed Medicaid beneficiaries and comparison group beneficiaries was similar throughout the baseline and Medicaid SSP implementation periods (***Figure F-2-21***). The percentage for both outcomes remained steady during the baseline and Medicaid SSP implementation period.

ACO = accountable care organization; SSP = Shared Savings Program.

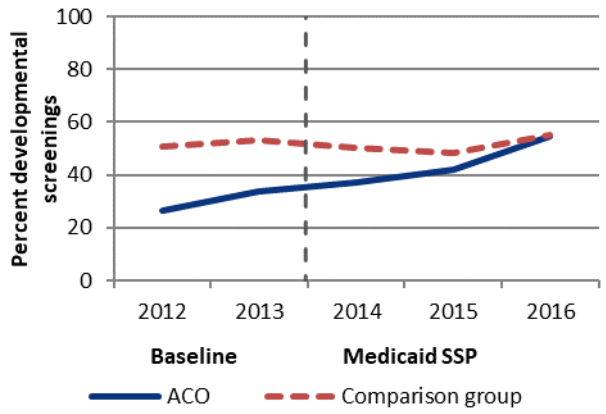
Figure F-2-22. Rate of admissions for ambulatory care sensitive conditions per 1,000 population, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The rate of **hospitalizations for ambulatory care sensitive conditions based on chronic PQI** among SSP-attributed Medicaid beneficiaries and comparison group beneficiaries trended upward during the baseline period and remained steady during the implementation period (***Figure F-2-22***).

ACO = accountable care organization; PQI = Prevention Quality Indicators; SSP = Shared Savings Program.

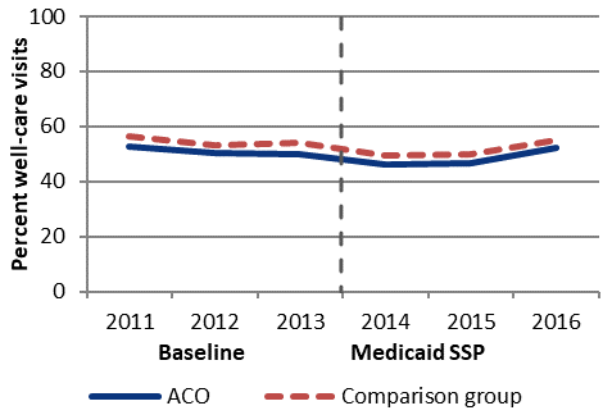
Figure F-2-23. Percentage of Medicaid beneficiaries who had a developmental screening, 2012 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The percentage of **developmental screenings** among SSP-attributed Medicaid beneficiaries trended upward during the baseline and Medicaid SSP implementation period. Among comparison group beneficiaries, the percentage remained stable throughout the baseline and Medicaid SSP implementation periods (***Figure F-2-23***).

ACO = accountable care organization; SSP = Shared Savings Program.

Figure F-2-24. Percentage of Medicaid beneficiaries who had an adolescent well-care visit, 2011 through 2016, Vermont Medicaid SSP-attributed beneficiaries and comparison group



The percentage of **adolescent well-care visits** among SSP-attributed Medicaid beneficiaries and comparison group beneficiaries was similar throughout the baseline and Medicaid SSP implementation periods (***Figure F-2-24***). The percentage for both outcomes trended slightly downward during the baseline period and trended slightly upward during the Medicaid SSP implementation period.

ACO = accountable care organization; SSP = Shared Savings Program.

An annual fixed-effects model considered for the evaluation is shown in ***Equation F-2.1***:

$$\gamma = \alpha_0 + \alpha_1 I + \sum \beta_n Q_{n,b} + \sum \phi_t Q_{t,p} \bullet I + \delta X + \mu \qquad \text{(F-2.1)}$$

where

y = a performance measure (e.g., total PBPM cost per year) for the i-th beneficiary in the j-th group (test or comparison), in period t (i,j,t subscripts suppressed).

I	=	a 0,1 indicator (0 = comparison group, 1 = Medicaid SSP).
X	=	a vector of patient and demographic characteristics.
$Q_{n,b}$, $Q_{t,d}$	=	0,1 indicator of the *n*-th or *t*-th calendar year in the base (*b*) or post (*p*) period (*n* starts counting at first baseline period, whereas *t* starts with first Medicaid SSP year).
μ	=	error term.

The model in ***Equation F-2.1*** assumes that, except for an intercept difference α_1, the outcomes for beneficiaries in the Medicaid SSP group and beneficiaries in the comparison groups followed a similar growth trend during the baseline period. We investigated whether the baseline period before the start of the Medicaid SSP satisfied the baseline trend assumptions of the D-in-D model in ***Equation F-2.1***—that is, whether the outcome trends for beneficiaries in Medicaid ACO and in the comparison group were similar during this period.

To test the similarity of baseline trends, we used a model with a linear trend during the baseline period. We tested whether this trend differed for Medicaid SSP participants relative to comparison group participants. Specifically, the model for the outcomes may be written as follows:

$$y = \alpha_0 + \alpha_1 I + \theta \bullet t + \lambda I \bullet t + \delta X + \mu. \qquad \text{(F-2.2)}$$

In ***Equation F-2.2***, *y, I,* X, and μ are defined as in ***Equation F-2.1***. The variable *t* is linear time ranging from 1 to 3. The linear time trend in the comparison group is θ•*t*, whereas for Medicaid SSP beneficiaries ($I = 1$) it is $(\theta + \lambda) * t$. Hence, λ measures the difference in linear trends and the *t*-statistic for this coefficient can be used to test the null hypothesis of equal trends ($\lambda = 0$). In other words, rejecting the null hypothesis would suggest that the assumption of equal trends underlying our outcome models is not met.

The parameters of ***Equation F-2.2*** were estimated using weighted least-squares regression models for three key outcomes. The weights are a function of the eligibility fraction and propensity scores. For each outcome, we report estimates and standard errors of the difference between the baseline trend in the test and the comparison groups (λ).

Table F-2-11 show estimates of the baseline trend differences for the following outcomes:

- Total expenditures
- Probability of an acute inpatient stay
- Probability of an outpatient ED visit

Table F-2-11. Differences in average expenditure and utilization outcomes during the baseline period, Medicaid SSP beneficiaries and comparison group beneficiaries

Parameter estimate	Total PBPM ($)	Any inpatient	Any outpatient ED visit
Test–CG trend difference	−0.11	.0002*	.0005
	(1.52)	(.0001)	(.0004)

CG = comparison group; ED = emergency department; PBPM = per beneficiary per month; SSP = Shared Savings Program.

Baseline is the period January 2011–December 2013. The trend (slope) is the year-to-year change in the outcome variable. Standard errors are given in parentheses. *p < 0.10

Relative to the comparison group, there was no statistically significant difference in the baseline trend for total PBPM expenditures, but the inpatient admission rate increased faster relative to the comparison group (statistically significant at the p<0.10 level). The magnitude of the difference was very small, however. Based on the overall results, we concluded that in general beneficiaries in the Medicaid SSP were on a similar trajectory with comparison beneficiaries prior to the Medicaid SSP, and thus the parallel trend assumption of the D-in-D model was satisfied.

D-in-D regression model. The D-in-D model is shown in ***Equation F-2.3***. The model is an annual fixed effects model as shown in ***Equation F-2.1.*** As in ***Equation F-2.1***, Y_{ijt} is the outcome for individual i (test or comparison group) in state j in year t; I_{ij} (=0,1) is an indicator equal to 1 if the individual is in the Medicaid SSP and 0 if the individual is in its comparison group; Qn is a series of yearly dummies for the baseline period (years 1 to 3); and Q_t is a series of yearly dummies for the post years (years 2014 to 2016). The interaction of the Medicaid SSP indicator and Q_t ($I_{ij} * Q_t$) measures the difference in the pre-post change between the Medicaid SSP and its comparison group.

$$Y_{ijt} = \alpha_0 + \beta_1 I_{ij} + \sum \beta_2 Q_n + \sum \alpha_2 Q_t + \sum \gamma I_{ijt} * Q_t + \lambda X_{ijt} + \varepsilon_{ijt} \qquad \text{(F-2.3)}$$

Table F-2-12 illustrates the interpretation of the D-in-D estimate from this model. The coefficient β_1 in ***Equation F-2.3*** is the difference in the measure between test beneficiaries and comparison beneficiaries at the start of the baseline period, holding constant other variables in the equation. The β_2 and α_2 coefficients are for the annual fixed effects and capture differences over time for each baseline and post year, respectively. The coefficient of the interaction term between Q_t and I measures any differences for the Medicaid SSP relative to the comparison group in the post years relative to the baseline year. Thus, in the post-period, the comparison group mean is captured by $\alpha_0 + \alpha_2$, whereas the Medicaid SSP mean is captured by $(\alpha_0 + \beta_1) + (\alpha_2 + \gamma)$. In other words, the between-group difference changes from β_1 during the baseline years

Table F-2-12. Difference-in-differences estimate

Group	Pre-period	Post-period	Pre-post difference
Test	$\alpha_0 + \beta_1 + \beta_2$	$(\alpha_0 + \beta_1) + (\alpha_2 + \gamma)$	$\alpha_2 + \gamma$
Comparison	$\alpha_0 + \beta_2$	$\alpha_0 + \alpha_2$	α_2
Between group	β_1	$\beta_1 + \gamma$	γ

to $\beta_1 + \gamma$ during the post-period. The D-in-D parameter, γ, shows whether the between-group difference increased ($\gamma > 0$) or decreased ($\gamma < 0$) after the Medicaid SSP was implemented. Using the annual fixed effects model, we calculated overall estimates by taking linear combinations of the yearly estimates.

Models for unplanned readmissions and mental health follow-ups were estimated at the annual admission level. All other outcomes were estimated with the beneficiary year as the unit of analysis.

The outcome model for total Medicaid PBPM expenditures was estimated using ordinary least squares. To show the adjusted means in the pre- and post-periods for the Medicaid SSP and comparison groups, we used a linear model that allows for the calculation of means that will sum to the D-in-D estimate. Although this model has strong assumptions of normality of the outcome, the ordinary least squares model still produces unbiased estimates even when the normality assumptions is violated as long as errors are uncorrelated and have a constant variance (Gauss-Markov Theorem). However, we can and do control for the correlation and variance in errors with clustered standard errors. Additionally, the model yields estimates that are readily interpretable in dollars and do not require additional transformation.

For all other outcomes, we converted utilization counts into binary outcomes (1 = any use) and used weighted logistic regression models. Count models are not appropriate because of the low occurrence of most types of utilization for individual beneficiaries in any year; however, we multiplied the marginal effect from the logistic regression models by 1,000 to obtain approximate rates of utilization per 1,000 beneficiaries. Multiplying the marginal effect by 1,000 does not produce an exact rate of utilization per 1,000 beneficiaries because it assumes no person has more than one visit or admission per year. However, we concluded that this is a reasonable approximation because only a small percentage of beneficiaries had counts exceeding 1 for any of the utilization measures. For expenditure outcomes, we used weighted generalized linear models with a normal distribution and identity link.

The models for total expenditures, inpatient admissions, ED visits, and readmissions were run separately for children and adults. In addition, we ran these outcomes and behavioral health–

related expenditures, ED visits, and inpatient admissions separately for people with behavioral health conditions.

Control variables. In all models we controlled for the following variables:

- Age (<1, 1 to 18, 19 to 64, 65 plus)
- Disability
- Gender
- Beneficiary's lagged or pre-period classification on the Chronic Illness and Disability Payment System
- Beneficiary's lagged or pre-period total payments
- Beneficiary's lagged or pre-period probability of inpatient admission or ED visit
- Beneficiary's participation in the Chronic Care Initiative
- Number of months beneficiary was Medicaid eligible during year (minimum of 10)
- If beneficiary was Medicaid eligible 10 or more months in previous year
- Attribution method of beneficiary (claims-based or choice/auto-assigned)
- If beneficiary was attributed to a Vermont Blueprint for Health medical home
- If beneficiary was eligible through Medicaid expansion
- Federal Poverty Level

Weighting and clustering. All of the regression models were estimated using weighted regressions and weighted by the propensity score times the eligibility fraction. In addition, standard errors were clustered at the provider level to account for clustering of beneficiaries within providers.

Adjusted means. The regression-adjusted D-in-D estimate and the D-in-D calculated from regression-adjusted means will differ for one of two reasons. First, in nonlinear specifications the D-in-D calculated from the regression-adjusted means is known to be a biased estimator for the treatment effect. To address this bias, we use the nonlinear D-in-D approach described in Puhani (2012). In some cases the bias may be extreme, leading to substantial differences between the regression-adjusted D-in-D estimates versus the D-in-D calculated from regression-adjusted means.

Second, in linear specifications the D-in-D calculated from the overall regression-adjusted means may be substantially different than the overall regression-adjusted D-in-D estimate because we use different weights to obtain the overall figures. Specifically, the regression-adjusted D-in-D estimates are weighted using the number of treatment beneficiaries

observed in each year relative to the total number of treatment beneficiaries ever observed during the test period. This is mathematically equivalent to weighting the test-period adjusted means for both groups with the same weights that are applied to the treatment group. However, the test-period adjusted means that are presented for the comparison group are weighted using the number of comparison beneficiaries observed in each year relative to the total number of comparison beneficiaries ever observed during the test period. The implication of this is that in cases where there are large differences in the rates of rolling entry or exit across the two groups, we may observe large differences in the D-in-D calculated from the overall regression-adjusted means versus the overall regression-adjusted D-in-D estimate.

F-2.2 Methods for Qualitative Data Collection and Analysis

The Vermont SIM Initiative Round 1 Evaluation team collected and analyzed a wide range of qualitative data in the fifth year of the federal SIM Initiative evaluation. These data sources included interviews with key informants and focus groups conducted during in-person site visits in previous evaluation years, a review of relevant documents, and regular evaluation calls with the state officials leading the state's SIM Initiative. This report draws from past evaluation reports, where further detail is provided on previously conducted site visit interviews and focus groups.

F-2.2.1 Document review

We used Vermont's quarterly and annual reports, operational plans, and other state documents to obtain updated information on their implementation progress during the SIM Initiative test period. To supplement these documents, we collected relevant news articles on the Vermont SIM Initiative activities and related initiatives, and we searched reform-oriented websites that the state maintains.

In addition, we obtained numbers of providers participating in and populations served by the different innovation models from quarterly reports Vermont submits to CMS and personal communication with the state. We provide Vermont's reported numbers in ***Appendix F***. Sources for these provider and population data are detailed in the Year Four Annual Report (RTI International, 2018).

F-2.2.2 State evaluation calls

We conducted monthly federal evaluation-specific calls beginning in April 2014 and continued through the end of the SIM Initiative test period. The RTI/NASHP evaluation team for Vermont, the state's SIM Initiative team, and the state's Innovation Center project officer typically attended each state evaluation call. The purpose of the calls was to review interim evaluation findings with the state (as available), discuss any outstanding federal evaluation data or other needs, review and discuss state implementation and self-evaluation updates, and gather more in-depth information on select topics of interest for the evaluation.

For each meeting used to collect additional information and perspective from state officials leading the SIM Initiative in Vermont, the evaluation team prepared a list of state-specific questions—including the status of related policy levers and implementation successes, challenges, and lessons learned. We first reviewed relevant state documents for answers to our questions. When we did not find answers in the document or needed further clarification, we sent the questions to the state ahead of the call and asked the state to have knowledgeable state officials available to answer the questions during the call.

F-2.2.3 Analysis

The RTI/NASHP evaluation team conducted thematic analysis of each source of qualitative data and then synthesized across information gleaned from site visits, focus groups, document review, and state evaluation calls. Site visit interviews and focus groups were conducted in previous years of the evaluation. For more detail on site visit and focus group methods, see past evaluation reports.

F-2.3 References

Agency for Healthcare Research and Quality. (2016). *Prevention Quality Indicators overview*. Available at https://www.qualityindicators.ahrq.gov/modules/pqi_resources.aspx

CMS. (2016). *Medicaid & CHIP: December 2015 monthly applications, eligibility determinations and enrollment report*. Available at https://www.medicaid.gov/medicaid/program-information/downloads/december-2015-enrollment-report.pdf

Puhani, P. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear ‘‘difference-in-differences’’ models. *Economics Letters, 115*, 85–87.

RTI International. (2018, March). *State Innovation Models (SIM) Initiative evaluation: Model Test Year four annual report.* Prepared for the Centers for Medicare & Medicaid Services (CMS). Available at https://downloads.cms.gov/files/cmmi/sim-rd1-mt-fourthannrpt.pdf

Vermont Digger (March 6, 2016). *Medicaid by the numbers: Reforms, demographic changes drive up enrollment*. Available at https://vtdigger.org/2016/03/06/medicaid-by-the-numbers-reforms-demographic-changes-drive-up-enrollment/#.Wgig1FKZPow

[this page intentionally left blank]

Appendix G: Methods for Statewide Analyses

G.1 Comparison Groups for Analysis of Measures in Claims Data

For statewide analysis of measures in claims data, we used a pre-post comparison group design, in which the comparison group provides a benchmark for comparison for what was happening in similar populations during the same time period as the SIM Initiative. The difference in the changes in outcomes over time from the pre-test to the test period between a Test state and its comparison group provides an estimate of the changes in the Test state that coincided with the SIM Initiative. The comparison group should be similar to the Test state on all relevant dimensions (e.g., demographic, socioeconomic, political, regulatory, and health and health systems) except for the policy change being tested.

We used a two-stage procedure to create a comparison group for each Test state. First, we identified three states that best resemble the Test state on key characteristics. Second, for each of the three payer databases (MarketScan, Medicare, and Medicaid), we weight individuals within the comparison states, so the population characteristics of the comparison states together are similar to those in the SIM Initiative target state. The weights—which are based on propensity scores computed from logistic regression of the probability that the individual resides in the Test state—are re-estimated annually.

The methods for selecting the comparison states for Round 1 Test states are reported in detail in our First, Second and Third Annual Reports and summarized below. It was our intent to use the same three comparison states for each payer, however, we lacked sufficient post period Medicaid data in the MAX/Alpha-MAX data system for comparison states, which reduces the number of comparison states to two each for the Test states. For the Fifth Annual Report, we changed the comparison group for the Medicaid data due to lack of availability in Medicaid data for previously selected comparison group states. We estimate propensity scores and weights for all years in which we have Medicaid data for the Test state and at least two comparison states.

G.1.1 Selection of comparison states

Relying on a single comparison state may be prone to bias, because contrasts may reflect idiosyncratic features of the comparison or Test state. To reduce the risk of this type of bias, we identified three comparison states for each Test state, using the following procedures:

- Identified the pool of potential comparison states
- Computed Euclidean distance scores (defined below) based on a broad array of state-level characteristics to summarize the difference between each Test state and each potential comparison state

- Used a boosted regression (defined below) to identify any additional characteristics unique to a Test state
- Rank-ordered comparison states by their distance scores
- Identified the states with the three smallest difference scores
- Reviewed the identified states for appropriateness
- Replaced inappropriate states with the next state in the rank-ordering until three comparison states had been identified

G.1.1.1 State-level characteristics

To select states comparable to the six Test states, we compiled a data base of 25 baseline (pre-SIM Initiative) state-level characteristics in the following dimensions:

- key outcomes of interest, including expenditures, utilization, care coordination, quality of care, provider, and population health
- demographic characteristics of the state's population, including age distribution, income levels, and employment
- access to care measures, such as the percentage of children and adults with no insurance, adults with a usual source of care, and children with medical and preventive care visits
- characteristics of the state's public and private health care systems, including Medicaid eligibility levels, managed care penetration levels, and provider supply
- health policy reforms, including implementation of the Patient Protection and Affordable Care Act Medicaid expansions, and the number of other Center for Medicare and Medicaid Innovation (Innovation Center) payment and delivery system initiatives

As reported in the Third Annual Report, the states in the final comparison group on average exhibit much smaller differences across these covariates than the entire pool of potential comparison states.

G.1.1.2 State selection procedures

Using this database of state characteristics, we assessed the similarity of each Test state to the pool of 16 Design, three Pre-Test, and 25 non–SIM Initiative comparison states. As noted, similarity was measured by a statistical measure of "distance" between two states, known as the Euclidean distance, which is based on the relative magnitude of the differences in state-level means. Distances are summed over characteristics to create a total distance score. The smaller the distance score, the more similar are the two states. We also computed another common distance measure, the Mahalanobis score, but found those scores to be unstable given the large number of characteristics under consideration.

We based the distance scores on the set of 25 characteristics listed in ***Section G.1.1.1*** for each Test state. However, since a Test state might have other extreme or unusual characteristics that should also be considered when selecting comparison states, we used boosted regression to examine more than 100 additional characteristics in our database. Boosted regression is a data mining technique that iteratively identifies influential predictors of an outcome, using an algorithm that can be efficiently applied to a variety of datasets. For three Test states, all influential variables identified by boosted regression were already part of the base set of 25 state characteristics. For two Test states, the addition of influential variables did not affect distance score rankings. For the remaining Test state, the variables identified by boosted regression resulted in some alterations of the rank-ordering of the top five potential comparison states.

The final step in the state selection process was to produce a list of comparisons for each Test state rank-ordered by distance scores, with the smallest scores at the top of the list. These lists were then reviewed by the evaluation team for problems. We initially removed comparison states from the list for one of two reasons: (1) unavailability of recent Medicaid claims or encounter data (Wisconsin, Pennsylvania, and New York) and (2) geographic distance or uniqueness (Hawaii). For the final report, we also removed the following states for the statewide Medicaid analysis only due to lack of Medicaid data availability with sufficient run out to produce stable estimates in the MAX data: Alabama, New Hampshire, Rhode Island, Colorado. In addition, because of incomplete encounter data following adoption of managed care among Medicaid enrollees in Kentucky, we dropped it as a comparison state for the Arkansas Medicaid analyses. We replaced each eliminated state with the next state in the rank order. For the Medicaid analysis, we replaced the eliminated states with the next state in the rank order that also had available data for the analysis. Due to lack of available data, we were also only able to include two comparison states for the Medicaid analysis.

Table G-1 shows the selected states and their distance scores for the MarketScan and Medicare analyses. A total of 10 different states were selected as comparisons for the Round 1 Test states. The three comparison states for Arkansas were not part of the SIM Initiative in Round 1, but Kentucky and Oklahoma were Round 2 Design states. The remaining seven comparison states were all SIM Round 1 Design or Pre-Test states;, six of these states (Colorado, Connecticut, Iowa, Michigan, Rhode Island, Washington), became Round 2 Test states and New Hampshire became a Round 2 Design state.

Table G-1. Comparison states selected for each SIM Test state for the MarketScan and Medicare analyses

Test state	Comparison states	Distance function value
Arkansas		
	Kentucky	11.42
	Alabama	15.82
	Oklahoma	18.45
Maine		
	New Hampshire	20.74
	Rhode Island	35.70
	Connecticut	39.76
Massachusetts		
	Connecticut	25.24
	New Hampshire	31.30
	Rhode Island	34.42
Minnesota		
	Colorado	29.20
	Iowa	33.83
	Washington	34.04
Oregon		
	Colorado	14.14
	Washington	18.66
	Michigan	19.41
Vermont		
	New Hampshire	20.44
	Iowa	30.04
	Connecticut	44.15

Table G-2 shows the selected states and their distance scores for the Medicaid analyses. We include Medicaid claims data in the analyses only if they had 2 or more quarters of run-out. ***Table G-2*** also shows the latest quarter meeting this criterion for each Test state and its comparison states.

Table G-2. Comparison states selected for each SIM Test state for the Medicaid analysis

Test state Comparison states	Distance function value	End quarter
Arkansas		Q1 2015
Michigan	24.59	Q3 2015
Oklahoma	18.45	Q3 2014
Maine		Q3 2015
Michigan	28.15	Q3 2015
Connecticut	39.76	Q1 2015
Massachusetts		Q3 2015
Michigan	45.56	Q3 2015
Connecticut	25.24	Q1 2015
Minnesota		Q3 2015
Iowa	33.83	Q3 2015
Washington	34.04	Q4 2014
Oregon		Q2 2015
Washington	18.66	Q4 2014
Michigan	19.41	Q3 2015
Vermont		Q3 2015
Iowa	30.04	Q3 2015
Connecticut	44.15	Q1 2015

G.1.2 Calculation of person-level weights

While the state selection process provides a set of up to three comparison states that are similar in major respects to each Test state, differences may remain between the populations of the Test and comparison states. To balance the population characteristics for the claims-based analyses, we estimated propensity scores for all individuals from the comparison states in each payer database annually. A propensity score is the probability that an individual is from the Test state rather than a comparison state.

The objective of propensity score modeling is to create a weighted comparison group with payer and beneficiary characteristics equivalent to those for the Test state population. To the extent that these characteristics are correlated with care coordination, quality of care, utilization and expenditures, propensity weighting will help balance pre-Initiative levels of the outcomes as well.

G.1.2.1 Person-level characteristics

The initial step in the process was to select person-level characteristics to be used in each propensity score model. We extracted these characteristics from the respective payer databases; therefore, each is unique to the particular database. ***Table G-3*** shows the characteristics used in each database grouped by whether they control for demographic, health plan, or health status characteristics.

Table G-3. Covariates for propensity score logistic regressions by payer type

Covariates	Medicaid	MarketScan	Medicare
Demographic characteristics			
Gender	✓	✓	✓
Age (age and age squared)	✓	✓	✓
Disabled (yes/no)	(a)	—	✓
White race (yes/no)	✓	—	✓
Resides in metropolitan area (yes/no)	—	—	—
Health plan characteristics			
Medicaid eligibility category (infant, child, nondisabled adult, blind/disabled)	—	—	—
Continuous enrollment indicator (yes/no)	✓	—	—
Also enrolled in Medicaid (yes/no)	—	—	✓
Employee relationship (employee/spouse/child-other)	—	✓	—
Pharmaceutical claims (yes/no)	—	✓	—
Mental health claims coverage (yes/no)	—	✓	—
Consumer-driven or high-deductible health plan (yes/no)	—	✓	—
Individual vs. employer plan	—	✓	—
Health status measures			
Hierarchical Condition Categories risk score	—	(b)	X
Chronic Illness and Disability Payment score (count of major comorbidities)	✓	—	—
Past year total expenditures	✓	✓	✓
Any inpatient admissions in the past year (yes/no)	✓	✓	✓
Any ED visits in the past year (yes/no)	✓	✓	✓

ED = emergency department.

(a) Blind/disabled is one of the eligibility categories we use for the Medicaid propensity score models.

(b) Hierarchical Condition Categories (HCCs) are calculated using three separate models: infants (0–1), children (2–20); adults (21+)

G.1.2.2 Estimation and weighting procedures

Using the characteristics listed in ***Table G-3***, we estimated propensity models by logistic regression, in which the outcome was 1 = Test state resident and 0 = comparison state resident. Separate models were estimated for each payer by year. For Medicaid, we ran the models for baseline years 2010, 2011, 2012 and 2013 as well as test year 2014. For Medicare and MarketScan, we ran the models for baseline years 2011, 2012 and 2013 as well as test years 2014, 2015, and 2016. Separate models were estimated by several key subpopulations within each payer. Within Medicaid and MarketScan, we estimated models separately for children (ages 0–18) and adults (19–64) as well as persons with a mental or behavioral health condition documented during the baseline period. Within Medicare, we estimated models separately for beneficiaries dually eligible for both Medicare and Medicaid.

We set analysis weights to 1 for all individuals in a Test state. The weight for a comparison state individual was initially a function of his/her predicted propensity score—where weight = p/(1-p), with p the predicted propensity. We then upweighted scores to a minimum value of .05 and capped larger weights at a maximum value of 20 to prevent any single individual from having undue influence on the results.

G.1.3 Propensity model evaluation

We evaluated several aspects of the propensity score models. First, we examined plots of predicted probabilities to ensure sufficient overlap in the distributions of the Test and the combined comparison states. This feature, known as common support, is critical because it provides the basis for inferring effects from group comparisons. In all years for all payers, we found the comparison group passed the common support assumption ($P(D = 1|X)>0$) for the Test and comparison groups (data not shown).

Second, we compared the logistic results for the same states in the three pre-demonstration years, to determine whether the same characteristics were influential over time. With a few minor exceptions, we found that the models were similar each year. This is not surprising, because the same individuals frequently appear in the databases for multiple years. In the MarketScan data, the variables with the greatest impact in the propensity score models were presence of mental health coverage and health plan status (individual vs. employer plan). Thus, the major differences between the Test state and comparison state populations were found for types of insurance coverage. In the Medicare data, the only two factors with comparatively large effects for more than one state were racial group and residence in a metropolitan area.

Finally, we compared unweighted and propensity-weighted means for the characteristics in the model. As expected, we found that, after weighting, the comparison group means were within a few percentage points of the values for their respective Test state each year within each payer source indicating that the propensity score weights substantially mitigate observed covariate balance (data not shown).

G.2 Claims-Based Outcomes: Data and Measures

G.2.1 Data sources

For the Year 5 Annual Report, we produced estimates of selected measures of (1) care coordination, (2) quality of care, (3) utilization, and (4) expenditures for three populations—Medicaid beneficiaries, the commercially insured in MarketScan, and Medicare beneficiaries. We describe the data sources and methods used below.

G.2.1.1 Medicaid data

The RTI evaluation team used Medicaid data from the CMS Medicaid Analytic eXtract (MAX) and Alpha-MAX research files made available through the CCW enclave for Arkansas, Minnesota, Oregon and Vermont analyses. Each state's Medicaid Statistical Information System (MSIS) data are the source of the MAX and Alpha-MAX files. The MAX processing adds enhancements such as claims adjustments, creation of a national type of service field, and state-specific quality issues corrections; Alpha-MAX provides fewer enhancements. The MAX and Alpha-MAX files include a person summary (PS) file, with all enrollment information and summary claims information and four claims files: inpatient hospital (IP), long-term care (LT), prescription drugs (RX), and other (OT) claims. The quarterly Alpha-MAX files are generated for a state once all five MSIS file types for a single quarter are approved. The quarterly files are overwritten and updated each time a new quarter of run-out data is added. Quarterly versions of Alpha-MAX are being produced for each state through 7 quarters of run-out data; therefore, the quarterly files are based on 0 to 7 quarters of run-out time. Annual calendar-year MAX files are prepared from data with 8 quarters of run-out time. For simplicity, we refer to the MAX and Alpha-MAX data as simply MAX data for the remainder of this appendix.

Availability of MAX data files varies by state and neither Maine nor Massachusetts has MAX data available in the CCW enclave. We obtained Maine Medicaid (MaineCare) data from the state's data vendor, Molina Medicaid Solutions. These data contain demographic and enrollment information, including a monthly indicator of enrollment. The data also include medical and pharmaceutical claims information for all facility and professional services, both inpatient and outpatient. We also obtained Medicaid claims data from the Massachusetts Medicaid program. In addition to monthly enrollment and demographic information, the Massachusetts data contain inpatient, outpatient, pharmacy, and capitation records for the traditional fee-for-service (FFS) plan beneficiaries as well as the managed-care plan enrollees.

G.2.1.2 MarketScan data

We used data from MarketScan Research Databases (©2018 from Truven Health Analytics Inc., an IBM Company), to calculate outcomes for the commercially insured population in SIM Round 1 Test and comparison states. In addition, we used data from the Maine and Vermont all-payer claims databases (APCDs). MarketScan may not be as

representative of the states' commercially insured population as the APCDs, but it provides similarly constructed comparison state data not otherwise available. The MarketScan data included in this report are from fourth quarter 2010 through fourth quarter 2016

MarketScan is the largest available database of commercial insurance claims and contains payment and utilization data for all claim types. The MarketScan commercial insurance claims are constructed from data contributed by around 350 payers, although the exact number of contributors varies by year. Individuals represented in the database are covered under plan types with a wide variety of delivery and payment types—including FFS, fully and partially capitated plans, and various plan models (such as preferred provider organizations). The MarketScan data include covered individuals from all 50 states and the District of Columbia. These data do not contain the same benefit design for everyone included in the sample. In particular, drug claims and mental health/substance abuse claims are not submitted or covered for everyone in the sample. Further, the database over-represents the self-insured market. Nevertheless, MarketScan is the largest and most complete source of timely commercial claims data in the United States, and importantly, it includes comparable claims in a uniform format for both Test and comparison states. To assess the generalizability of the MarketScan findings, we used the Minnesota All Payer Claims Database (MN APCD) provided by the Minnesota Department of Health (MDH), the Vermont Health Care Uniform Reporting and Evaluation System (VHCURES) data provided by the Green Mountain Care Board, and the Oregon All Payer All Claims data provided by the Oregon Health Authority to obtain a more complete picture of the commercially insured population in Minnesota, Vermont, and Oregon, respectively. The MarketScan and all-payer claims data include clinical, financial, and demographic fields used to calculate cross-state care coordination, quality of care, utilization and expenditure outcomes. We created the following analytic files using the MarketScan and APCD data files:

- **Annual enrollment file**. The Annual Enrollment Summary Table for MarketScan and member enrollment files in the APCDs contain enrollment information for every person enrolled during the year, including a monthly indicator of enrollment. We used the enrollment files to calculate fraction of time each person was enrolled and total number of people enrolled per year in each state.
- **Claims data**. MarketScan and the APCDs include files that contain complete header information for all facility claims, all facility and professional encounters and paid claims for inpatient and outpatient services, and outpatient pharmaceutical claims data for a portion of the covered individuals. We used these files to calculate the care coordination, quality of care, utilization and expenditure outcomes.

G.2.1.3 Medicare data

We used Medicare claims and enrollment data for fourth quarter 2010 through fourth quarter 2016 from the Chronic Conditions Data Warehouse. These data include: (1) denominator information that indicates number of beneficiaries alive and enrolled in Medicare during the period; (2) enrollment information that indicates number of days beneficiaries were enrolled in

Medicare during the period; and (3) claims experience for each beneficiary—including inpatient, hospital outpatient, physician, skilled nursing facility, home health agency, hospice, and durable medical equipment claims.

G.2.2 Population

For the statewide trend analyses, the target populations are all individuals included in the Medicaid, MarketScan, and Medicare databases for all states except Arkansas, Maine, Massachusetts and Oregon. In Arkansas, we do not include Medicare beneficiaries because they were not a targeted population under the SIM Initiative in Arkansas. Likewise, in Maine, we exclude the commercially insured population because they were not targeted by any of the Maine initiatives and there was little chance of spillover of SIM activities to the commercially insured. In Massachusetts, we do not include any data for Medicare or MarketScan (commercially insured population), because activities funded under the SIM Initiative in that state reached providers who served primarily Medicaid beneficiaries and supported only payment reform under the state's Medicaid agency; therefore, it is not informative to analyze differences in Massachusetts and its comparison states between the pre- and post-SIM Initiative periods. In Oregon, we restricted the Medicare population in to Medicare-Medicaid beneficiaries, because over half of Medicare-Medicaid beneficiaries in Oregon are enrolled in a coordinated care organization. The complete inclusion and exclusion criteria are described in detail in ***Section G.2.2.1***. In addition, because of the great variation in health care needs among select population subgroups, we conducted separate analyses of key subpopulations.

G.2.2.1 Population inclusions and exclusions

For each Test state and comparison group, we include all Medicaid beneficiaries eligible for full benefits; we exclude Medicaid beneficiaries eligible for only a restricted set of benefits, such as family planning program beneficiaries and those eligible for emergency Medicaid at the time of labor and delivery. Because Medicaid claims present only a partial picture of health care use among Medicare-Medicaid beneficiaries, we report utilization measures for Medicaid-only beneficiaries. However, we do present descriptive statistics and regression models for total Medicaid payments made on behalf of Medicare-Medicaid beneficiaries.

For the care coordination, quality of care, and utilization outcomes, the target commercial population was all individuals in the MarketScan database identified as enrolled in an included commercial plan at any point during the given analysis year. Because capitated plans may not have complete expenditure data in the MarketScan database, we restricted the sample for expenditure outcomes to commercially insured individuals identified as enrolled at any point during the year in an FFS plan and having no capitated payments in the database. Similarly, to calculate expenditures we restricted the Vermont and Maine APCD sample to the FFS population.

Because Medicare Advantage (i.e., managed care) enrollees may not have complete utilization and expenditure data, we excluded beneficiaries with any months of enrollment in Medicare managed care. We restricted the Medicare sample to beneficiaries who were alive at the beginning of the year, had at least 1 month of both Part A and Part B enrollment, had no months of Part A only or Part B only, and had no months of Medicare managed care enrollment. In addition, we restricted the Oregon Medicare population to Medicare-Medicaid beneficiaries.

G.2.2.2 Population subgroups

Health care use varies by age and health conditions category for Medicaid and commercial beneficiaries. Therefore, we report descriptive results for the overall beneficiary population and separately for children, adults and among those with documented mental or behavioral health conditions during the baseline period. For each year, we used age as of last enrollment month to define an individual's age group. Because Medicaid claims represent only a partial picture of health care use among dually eligible Medicare-Medicaid beneficiaries, we do not report care coordination, quality of care or utilization outcomes for beneficiaries in this group. We do, however, report total Medicaid payments separately for Medicare-Medicaid and Medicaid-only beneficiaries.

We report descriptive results for the overall Medicare population and by whether the beneficiaries were Medicare-Medicaid beneficiaries (who have different health care needs and utilization patterns than Medicare-only beneficiaries). Beneficiaries were designated as dually eligible Medicare-Medicaid enrollees for the year if they were enrolled in Medicaid for at least one month during the year.

G.2.2.3 Population weights

Eligibility fraction

Because some individuals are not enrolled in insurance throughout an entire period, we calculate eligibility fractions for each individual. The eligibility fraction is defined as total number of months the person was enrolled in a given period divided by total number of months in the period. For example, an individual enrolled in insurance nine months of a year has an eligibility fraction of 0.75 for that year. The eligibility fraction is used to inflate expenditure and utilization data if an individual was not enrolled for an entire period. The eligibility fractions are also used as weights in calculating weighted average outcomes. This prevents individuals with limited enrollment but extreme outcomes from strongly influencing the results.

Propensity score

For the comparison groups, outcomes are weighted by the eligibility fraction times the propensity score weight. We used propensity score weights to create a pooled, weighted comparison group from the comparison states for each target Test state and payer. A description of the methods used to develop the propensity score weights can be found in ***Appendix B***.

Balancing weight

To reduce the risk of bias from often unobserved individual state idiosyncrasies, we used two or three states to form a pooled comparison group for each Test state. We then created population balancing weights for the Medicaid, MarketScan, and Medicare populations to insure equal contribution from each of the comparison states in the pooled comparison group, regardless of population size in the comparison state. We created the balancing weight for each comparison state using the formula:

BW_i = [(sum of all eligible persons from all three comparison states)/3] / (sum of eligible persons in comparison state i)

For Medicaid analyses where we had to reduce the comparison group to only two states because of unavailable data for the third, we revised the formula to:

BW_i = [(sum of all eligible persons from the two comparison states)/2] / (sum of eligible persons in comparison state i)

G.2.3 Measures

We present estimates from claims data for four domains of performance: (1) care coordination, (2) quality of care, (3) utilization and (4) expenditures. We present graphical presentations of annual estimates for the outcome measures as well as the difference-in-differences (D-in-D) regression analyses.

G.2.3.1 Care coordination measures

One of the objectives of the SIM Initiative is to address the shortcomings of historically fragmented delivery systems in the Test states through better care coordination. Whether the state implements a medical or health home, accountable care organization, episode-of-care payment, or another model, the innovation models are expected to include features that improve care coordination for patients.

To evaluate the changes in care coordination, we report the following care coordination measures for all payers:

- **Percentage of beneficiaries with a visit to a primary care provider.** Visits to primary care providers were counted if the provider type was any of the primary care provider types listed in ***Table G-4,*** and one of the following primary care evaluation and management Current Procedural Terminology (CPT) codes was included on the claim for the visit:

  99201–99205, 99211–99215, 99241–99245, 99304–99310, 99315–99316, 99318, 99324–99328, 99334–99350, 99358–99359, 99366–99368, 99374–99397, 99401–99412, 99420, 99429, 99441–99444, 99495, 99496

- **Percentage of beneficiaries with a visit to a specialty provider.** Visits to specialty providers were counted if the provider type was any of the specialty provider types listed in ***Table G-4***, and one of the primary care evaluation and management CPT codes shown above was included on the claim for the visit.

Table G-4. Primary and specialty provider types

Primary care providers	Specialty providers	
General practice	Allergy/immunology	General surgery
Family practice	Otolaryngology	Anesthesiology
Internal medicine	Cardiology	Neurosurgery
Pediatrics (for MarketScan)	Dermatology	Oral surgery (dentists only)
Geriatric medicine	Gastroenterology	Orthopedic surgery
Multispecialty clinic or group practice	Neurology	Plastic and reconstructive surgery
Preventive medicine	Ophthalmology	Colorectal surgery
Nurse practitioner	Pathology	Thoracic surgery
Physician assistant	Physical medicine and rehabilitation	Hand surgery
Obstetrics/gynecology (for MarketScan; specialty provider for Medicare)	Psychiatry	Vascular surgery
	Pulmonary disease	Cardiac surgery
	Diagnostic radiology	Maxillofacial surgery
	Urology	Surgical oncology
	Nephrology	Sports medicine
	Infectious disease	Geriatric psychiatry
	Endocrinology	Palliative medicine
	Rheumatology	Sleep medicine
	Peripheral vascular disease	Pain management
	Critical care (intensivists)	Osteopathic
	Hematology/oncology	Nuclear medicine
	Neuropsychiatry	Radiology
	Medical oncology	Addiction medicine
	Emergency medicine	

- **Percentage of beneficiaries with any visit to a physician (Medicaid only).** Due to data constraints, we were not able to separate out visits by type of provider for Medicaid beneficiaries, so we report the percentage of beneficiaries with an E&M physician visit using CPT codes listed above.
- **Percent of acute inpatient hospital admissions with a follow-up visit within 14 days.** This is the number of acute inpatient hospital admissions that are followed by a visit to a provider within 14 days of discharge date, divided by the total number of acute inpatient hospital admissions. We used the following CPT codes to identify a follow-up visit:

  99201–99205, 99211–99215, 99241–99245, 99304–99310, 99315–99316, 99318, 99324–99328, 99334–99350

- **Percentage of mental illness related acute inpatient hospital admissions with a mental health follow-up visit within 7 and 30 days.** This is the number of acute inpatient hospital admissions with a primary diagnosis for a behavioral health condition (ICD-9 diagnosis codes 291, 292, 303, 304, 305, 293–302, 306–316) that are followed by a visit to a provider for a mental health visit (identified by visits with any of the below CPT or revenue codes) within 7 or 30 days of discharge date, divided by the total number of acute inpatient hospital admissions with a primary diagnosis for a behavioral health condition. Admissions that are followed by a readmission to an acute or other facility within 7 or 30 days are excluded from the respective denominators.

  Procedure code= 90801, 90802, 90804–90819, 90821–90824, 90826–90829, 90845, 90847, 90849, 90853, 90857, 90862, 90870, 90875, 90876, 98960–98962, 99078, 99201–99205, 99211–99215, 99217–99223, 99231–99233, 99238, 99239, 99241–99245, 99251–99255, 99341–99345, 99347–99350, 99383–99387, 99393–99397, 99401–99404, 99411, 99412, 99510. G0155, G0176, G0177, H0002, H0004, H0031, H0034-H0037, H0039, H0040, H2000, H2001, H2010-H2020, M0064, S0201, S9480, S9484, S9485

  OR

  Revenue code =0513, 0900–0905, 0907, 0911–0917, 0919

G.2.3.2 Quality of care measures

For all three payers, we include three baseline measures of quality of care: (1) ambulatory sensitive condition hospitalization rates; (2) influenza immunization rates; and (3) breast cancer screening rates. For MarketScan and Medicaid, we additionally report two well-child visit measures: (1) the percentage of children age 3–6 years who have 1 or more well-child visits and (2) percentage of 15 month olds with 0 or 6 or more well-child visits in the first 15 months of life and two medication management measures: (1) asthma medication management and (2) depression medication management. For Medicare, we additionally present the percentage of patients age 18 years and older seen for a visit who were screened for tobacco use and who received cessation counseling if identified as a user. We also calculated the tobacco screening rates for Medicaid and MarketScan, but we do not present the rates because the sample size of claims was too low to provide meaningful rates. Each measure is described in detail below.

- **Prevention Quality Indicators (ambulatory sensitive condition hospitalization rates).** For each payer, we evaluated the rates of avoidable hospitalizations using the composite prevention quality indicators (PQIs) that the Agency for Healthcare Research and Quality has stewarded as ambulatory care sensitive conditions. The idea behind PQIs is that certain hospitalizations may be avoided with adequate and quality access to primary care services. Given the low rates of the individual measures, we

report on the three composite PQIs.[42] The first, the ***Overall Composite (PQI #90)***, includes 12 of the 14 individual PQIs:

- PQI #01 Diabetes Short-Term Complications Admission Rate
- PQI #11 Bacterial Pneumonia Admission Rate
- PQI #03 Diabetes Long-Term Complications Admission Rate
- PQI #12 Urinary Tract Infection Admission Rate
- PQI #05 Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate
- PQI #13 Angina without Procedure Admission Rate
- PQI #07 Hypertension Admission Rate
- PQI #14 Uncontrolled Diabetes Admission Rate
- PQI #08 Heart Failure Admission Rate
- PQI #15 Asthma in Younger Adults Admission Rate
- PQI #10 Dehydration Admission Rate
- PQI #16 Rate of Lower-Extremity Amputation Among Patients With Diabetes

The second is the ***Acute Composite (PQI #91)*** and includes three individual PQIs:

- PQI #10 Dehydration Admission Rate
- PQI #12 Urinary Tract Infection Admission Rate
- PQI #11 Bacterial Pneumonia Admission Rate

Finally, the ***Chronic Composite (PQI #92)*** measure includes nine individual PQIs:

- PQI #01 Diabetes Short-Term Complications Admission Rate
- PQI #13 Angina without Procedure Admission Rate
- PQI #03 Diabetes Long-Term Complications Admission Rate
- PQI #14 Uncontrolled Diabetes Admission Rate
- PQI #05 Chronic Obstructive Pulmonary Disease or Asthma in Older Adults Admission Rate
- PQI #15 Asthma in Younger Adults Admission Rate
- PQI #07 Hypertension Admission Rate
- PQI #16 Rate of Lower-Extremity Amputation Among Patients With Diabetes
- PQI #08 Congestive Heart Failure Admission Rate

[42] PQI rates will be calculated per 100,000 patients. Only observable rates will be reported, as risk-adjusted rates posted by the Agency for Healthcare Research and Quality for the PQIs are established based on the general population in a geographic area, and will be incorrect when limited to the MarketScan population.

- **Percentage of patients age 1 year and older seen for a visit between October 1 and March 31 who received an influenza immunization during the visit.** This is the percentage of individuals who had a physician visit (as identified by CPT codes given below) during the flu season (service date fell between October 1–March 31) who received an influenza immunization. Individuals were identified as having an influenza immunization if they had one of the following procedure codes: G8482, G8483, G0919, G8484, 90653, 90654, 90656, 90658, 90660, 90661, 90662, 90664, 90666, 90667, 90668, 90672, 90673, 90686, 90688, G0008, Q2034, Q2035, Q2036, Q2037, Q2038, Q2039.

  Procedure codes to identify evaluation and management (E&M) visits:

90945	90959	90969	99215	99324	99342
90947	90960	90970	99304	99325	99343
90951	90961	99201	99305	99326	99344
90952	90962	99202	99306	99327	99345
90953	90963	99203	99307	99328	99347
90954	90964	99204	99308	99334	99348
90955	90965	99205	99309	99335	99349
90956	90966	99212	99310	99336	99350
90957	90967	99213	99315	99337	G0438
90958	90968	99214	99316	99341	G0439

- **Percentage of women 41-69 years old who had a mammogram to screen for breast cancer during the measurement year.** This is the percentage of women age 41–69 years old at the start of the measurement year who were screened for breast cancer (procedure code = 8736, 8737, 77055–77057, G0202, G0204 or G0206 or revenue code= 0401 or 0403). Women were excluded from the denominator if they were not enrolled for at least 11 of the 12 months of the year or ever had a bilateral mastectomy or two unilateral mastectomies (procedure code = 8541, 8543, 8545, 8547 or 19303–19307).

- **Percentage of children age 3–6 years who have 1 or more well-child visits during the measurement year.** The percentage of members who were 3–6 years old during the year who had at least 1 well-child visit during the year. A visit counts as a well-child visit if the claim includes a diagnosis code of V20.2, V70.0, V70.3, V70.5, V70.6, V70.8, or V70.9 or a procedure code of 99382, 99383, 99392, or 99393.

- **Well child visits within 15 months of age**. The percentage of members who turned 15 months old during the measurement year and who had the following number of well-child visits during their first 15 months of life:

  – No well-child visits

  – One well-child visit Six or more well-child visits

The denominator includes all infants in MarketScan and Medicaid who turn 15 months in the given year and who are continuously enrolled from 1 month to 15 months of age. The numerator is the count of children with 0 to 6 or more well-child visits. A visit counts as a well-

child visit if the claim includes a diagnosis code of V20.2, V20.3, V20.31, V20.32, V70.0, V70.3, V70.5, V70.6, V70.8, or V70.9, or a procedure code of 99381, 99382, 99391, 99392, or 99461.

- **Percentage of patients age 5-64 years with persistent asthma who were appropriately prescribed medication during the year.** This is the percentage of patients identified with persistent asthma who had an asthma medication dispensed to them during the year. To identify patients with persistent asthma, the patient had to be 5–64 years old and have a diagnosis for asthma (ICD-9 diagnosis codes 493.0, 493.1, 493.8, 493.9) that met at least 1 of the following 4 criteria:

  i. ***At least one ED visit with asthma as the principal diagnosis.*** (CPT code = 99281–99285 or revenue code=045x, 0981)

  ii. ***At least one acute inpatient discharge with asthma as the principal diagnosis.*** (CPT code=99221–99223, 99231–99233, 99238, 99239, 99251–99255, 99291 or revenue code=010x, 0110–0114, 0119, 0120–0124, 0129, 0130–0134, 0139, 0140–0144, 0149, 0150–0154, 0159, 016x, 020x, 021x, 072x, 0987)

  iii. ***At least four outpatient visits on different dates of service, with asthma as one of the listed diagnoses and at least two asthma medication dispensing events.*** To identify outpatient visits, CPT code=99201–99205, 99211–99215, 99217–99220, 99241–99245, 99341–99345, 99347–99350, 99382–99386, 99392–99396, 99401–99404, 99411, 99412, 99420, 99429 and revenue code =051x, 0520–0523, 0526–0529, 057x–059x, 0982, 0983. Asthma medication events were identified using the list of asthma medications in the table below.

  iv. ***At least four asthma medication dispensing events.*** Asthma medication events were identified using the list of asthma medications in the table below. If all four dispensing events were "leukotriene modifiers," then the individual also needed a diagnosis of asthma for any kind of service.

  Patients diagnosed with emphysema, COPD, cystic fibrosis, and acute respiratory failure in the prior year were excluded from the denominator (ICD-9 diagnosis codes 492, 518.1, 518.2, 491.2, 493.2, 496, 506.4, 277.0, and 518.81).

  For individuals who met the above criteria, we flagged whether or not they were dispensed at least one prescription for one of the asthma controller medications in the table below during the measurement year and calculated the percentage.

Description	Prescriptions		
Antiasthmatic combinations	Dyphylline-guaifenesin	Guaifenesin-theophylline	Potassium iodide-theophylline
Antibody inhibitor	Omalizumab		
Inhaled steroid combinations	Budesonide-formoterol	Fluticasone-salmeterol	Mometasone-formoterol
Inhaled corticosteroids	Beclomethasone Budesonide Ciclesonide	Flunisolide Fluticasone CFC free Mometasone	Triamcinolone
Leukotriene modifiers	Montelukast	Zafirlukast	Zileuton
Long-acting, inhaled beta-2 agonists	Aformoterol Indacaterol	Formoterol Salmeterol	
Mast cell stabilizers	Cromolyn	Nedocromil	
Methylxanthines	Aminophylline Dyphylline	Oxtriphylline Theophylline	
Short-acting, inhaled beta-2 agonists	Albuterol Levalbuterol	Metaproterenol Pirbuterol	

- **Percentage of patients age 18 years and older diagnosed with a new episode of major depression and treated with antidepressant medication who remained on medication treatment at least 84 and 180 days.** The percentage of patients 18 years of age and older who were diagnosed with a new episode of major depression and treated with antidepressant medication, and who remained on an antidepressant medication treatment. Two rates are reported:
  - *Effective Acute Phase Treatment.* The percentage of newly diagnosed and treated patients who remained on an antidepressant medication for at least 84 days (12 weeks).
  - *Effective Continuation Phase Treatment.* The percentage of newly diagnosed and treated patients who remained on an antidepressant medication for at least 180 days (6 months).

To identify patients with a new episode of major depression, the patient had to be at least 18 years old and have a diagnosis for major depression (ICD-9 diagnosis codes 296.20–296.25, 296.30–296.35, 298.0, 311) that met at least 1 of the following criteria:

- At least one principal diagnosis of major depression in any outpatient, ED, intensive outpatient or partial hospitalization setting (as indicated by the procedure or revenue codes given below)
- At least two visits in an outpatient, ED, intensive outpatient or partial hospitalization setting (as indicated by the procedure or revenue codes given below) on different dates of service with any diagnosis of major depression
- At least one inpatient (acute or nonacute) claim/encounter with any diagnosis of major depression

To identify the date of the first diagnosis, we used the date of the first claim/encounter that met one of the above criteria. To identify the date the medication was dispensed, we used the date that an antidepressant medication (as shown in the below table) was dispensed during the period 30 days prior to 14 days after the date of the first diagnosis.

We then checked whether the antidepressant medication was dispensed for at least 84 days (12 weeks) and 180 days (6 months) of continuous treatment with no more than 30 or 51 gap days in treatment, respectively.

Patients who received an antidepressant medication any time 3 months prior to the date the antidepressant medication was dispensed and those who were not continuously enrolled for 45 days prior to and 245 days after the depression diagnosis were excluded from the denominator.

Codes to identify visits

Description	CPT	HCPCS	UB Revenue
ED	99281–99285		045x, 0981
Outpatient, intensive outpatient and partial hospitalization	90804–90815, 98960–98962, 99078, 99201–99205, 99211–99215, 99217–99220, 99241–99245, 99341–99345, 99347–99350, 99384–99387, 99394–99397, 99401–99404, 99411, 99412, 99510	G0155, G0176, G0177, G0409–G0411, H0002, H0004, H0031, H0034-H0037, H0039, H0040, H2000, H2001, H2010-H2020, M0064, S0201, S9480, S9484, S9485	0510, 0513, 0515–0517, 0519–0523, 0526–0529, 0900, 0901, 0902–0905, 0907, 0911–0917, 0919, 0982, 0983
	CPT		**POS**
	90801, 90802, 90816–90819, 90821–90824, 90826–90829, 90845, 90847, 90849, 90853, 90857, 90862, 90870, 90875, 90876, 99221–99223, 99231–99233, 99238, 99239, 99251–99255	***WITH***	03, 05, 07, 09, 11, 12, 13 14, 15, 20, 22, 24, 33, 49, 50, 52, 53, 71, 72

Antidepressant medications

Description	Prescription		
Miscellaneous antidepressants	Bupropion	Vilazodone	
Monoamine oxidase inhibitors	Isocarboxazid	Selegiline	
	Phenelzine	Tranylcypromine	
Phenylpiperazine antidepressants	Nefazodone	Trazodone	
Psychotherapeutic combinations	Amitriptyline-chlordiazepoxide		Fluoxetine-olanzapine
	Amitriptyline-perphenazine		
SSNRI antidepressants	Desvenlafaxine	Duloxetine	Venlafaxine
SSRI antidepressants	Citalopram	Fluoxetine	Paroxetine
	Escitalopram	Fluvoxamine	Sertraline
Tetracyclic antidepressants	Maprotiline	Mirtazapine	
Tricyclic antidepressants	Amitriptyline	Desipramine	Nortriptyline
	Amoxapine	Doxepin	Protriptyline
	Clomipramine	Imipramine	Trimipramine

- **Percentage of patients age 18 years and older seen for a visit who were screened for tobacco use and who received cessation counseling if identified as user in measurement year.** This is the percentage of individuals who had a physician visit (as identified by CPT codes given below) who received screening and counseling for tobacco use (CPT code = 4004F or 1036F).

  Procedure codes to identify physician visits:

90791	90839	92014	97004	99205	99406
90792	90845	96150	99201	99212	99407
90832	92002	96151	99202	99213	G0438
90834	92004	96152	99203	99214	G0439
90837	92012	97003	99204	99215	

G.2.3.3 Utilization measures

Utilization measures are reported as rates per 1,000 covered lives (or discharges for readmissions). For each measure, the numerator is the weighted sum of number of events (inpatient admissions, 30-day readmissions, and ER visits that did not lead to a hospitalization). Events are included in a period's total if discharge or service date on the claim was during the period. The denominator is the number of eligible plan members in the state enrolled during the analytic year.

- **Rate of all-cause inpatient hospitalizations (per 1,000 covered lives):** This is an indicator of whether the beneficiary had at least one admission to an acute-care hospital reported in the inpatient file for the year, divided by the number of beneficiaries in the same year. For Medicaid, we identified acute care hospital admission by including all admissions in the MAX inpatient (IP) file with a type of service that indicated admission was to an inpatient hospital (type of service = 01) and all inpatient admissions in the MaineCare data with a bill type of 11 or 12. For MarketScan, we identified acute care hospital admission by including all admissions with a place of service that indicated admission was to an inpatient hospital (place of service = 21). For Medicare, we identified all hospital admissions in which the last four digits of the provider values were 0001–0879 (acute inpatient) or 1300–1399 (critical access hospitals). For all data sources, some records in the inpatient claims files may appear to be multiple admissions but are in fact transfers between facilities; these records are counted as a single admission. To combine transfers into one acute admission, we identified claims that had no more than 1 elapsed day between discharge date of the index claim and admission date of the subsequent claim. We combined the claims into one record by taking earliest admission date and latest discharge date and summing all payment amounts. This same roll-up procedure was applied to claims with overlapping or identical admission and discharge dates (i.e., claims associated with the same visit).
- **Rate of ED visits that did not result in an inpatient hospital admission (per 1,000 covered persons):** This is an indicator of whether the beneficiary had at least one visit to the ED that did not result in an inpatient hospital admission, divided by the number of beneficiaries in the same period. ED visits, including observation stays,

were identified in the outpatient services file as visits with a revenue center line item equal to 045X or 0981 (ER care) or 0762 (treatment or observation room, thus counting observation stays in the overall count). If the procedure code on every line item of the ED claim equals 70000 through 79999 or 80000 through 89999, and no line items have a revenue center code equal to 0762, we excluded these claims (thus excluding claims where only radiological or pathology/laboratory services were provided, unless it was an observation stay). Because not all states submit complete revenue code information in their Medicaid data, we additionally identified visits that included the following procedure codes as outpatient ED visits in the Medicaid data: 99281, 99282, 99283, 99284, or 99285.

- **Rate of 30-day readmissions (per 1,000 discharges):** This is an indicator of whether the beneficiary had at least one acute hospitalization that occurred within 30 days following a live discharge for beneficiaries ages 18 or older for the year, divided by the number of inpatient discharges in the same year. Index hospital discharges were identified as inpatient stays with a discharge date within the given measurement period (12 months) minus 30 days from the end of the period. If an index stay had another admission within 30 days, the numerator is set to 1.

G.2.3.4 Expenditure measures

Weighted average payments are calculated on a per member per month (PMPM) basis. For each individual, PMPM payments were estimated as one-twelfth of their annual payments. Expenditures are defined as payments made by the payer (Medicaid, commercial, or Medicare); enrollee cost-sharing was not included (and is nonexistent or minimal in Medicaid). All individuals enrolled in the period for the state were included in calculating the averages, so the figures also include individuals with zero medical costs. The payments were not risk-adjusted or price-standardized across geographic areas. Claims were included in an analytic year based on the date of discharge or service date on the claim was during the year.

Medicaid expenditures

Current Medicaid program designs often include a complex mix of traditional FFS plans and managed care plans with innovative delivery systems (fully or partially capitated plans, primary care case management [PCCM] plans, vulnerable population plans, service carve-out plans, etc.). Due to potential inaccuracies, the Medicaid paid amount for managed care encounter records is set to zero in MAX data. We therefore do not present payment by type of service for Medicaid. Managed care payments—including capitated payments to HMO plans, pre-paid health plans, and PCCM plans—were included as premium payment records with a capitated type of service code. We present the following categories of payments for Medicaid:

- **Total payments.** Total payments represent overall net payment amounts from all FFS claims and all capitated payments made to HMOs, pre-paid health plans, and PCCM plans. Total payments include all FFS payments made for inpatient, other therapy, long-term care, and pharmacy claims. We present total PMPM payments for each state for Medicaid-only enrollees and Medicare-Medicaid enrollees separately. In

addition, we present the average FFS, PCCM, capitated, and total payments by year for each state.

MarketScan and Medicare expenditures

We report the following categories of payments for MarketScan and Medicare:

- **Total payments.** Total payments represent overall net payment amounts from all inpatient and outpatient (facility and professional) claims and encounters, excluding member cost sharing. Although pharmacy component expenditures are included for MarketScan, total payments do not include pharmacy claims, because MarketScan does not include drug claims for every member.
- **Inpatient hospitals facility.** This represents the sum of net facility payments to a hospital for covered services provided during all inpatient admissions. Inpatient admissions were assigned to a period based on discharge date. Inpatient admissions include stays in psychiatric hospitals and rehabilitation facilities, but exclude skilled nursing facility stays.
- **Non–inpatient facility.** This represents the sum of net facility payments for non-inpatient services, including those made for outpatient, home health, hospice, and skilled nursing facility services.
- **Professional.** This represents the overall net payment amounts from all inpatient and outpatient professional claims and encounters, excluding member cost sharing.
- **Pharmaceutical payments.** This is the sum of net payments for outpatient pharmaceutical claims. The denominator for the average pharmaceutical payments is restricted to individuals with drug claims in MarketScan data.

G.2.4 Statistical methods

G.2.4.1 Difference-in-Differences regression analysis

To test for differences in care coordination, quality or care, expenditures and utilization outcomes during the SIM Initiative and the baseline period between the Round 1 Test states and their comparison groups, we use alternative D-in-D regression analyses. We conduct all analyses at the individual beneficiary level with annual observations, so the unit of analysis is person-year. For the utilization outcomes, we convert annual utilization counts into binary outcomes and use logistic regression models. Count models are not appropriate because of the low occurrence of multiple hospitalizations and ED visits for individual beneficiaries in any year; however, we multiplied the marginal effect from the logistic regression models by 1,000 to obtain approximate rates of utilization per 1,000 beneficiaries. Multiplying the marginal effect by 1,000 does not produce an exact rate of utilization per 1,000 beneficiaries because it assumes that no person has more than one visit or admission per year. However, we concluded that this is a reasonable approximation because the majority of the populations had zero or one ED visit or admission per year. Because the D-in-D coefficient of the logit model is an interaction term (Test*Post), it is not readily interpretable from a logistic regression model. We used literature on

D-in-D non-linear models to transform the coefficient to a marginal effect (Dowd, Greene, & Norton, 2014; Puhani, 2008; Puhani, 2012). For expenditure outcomes, we use weighted ordinary least squares (OLS) regression models. For all outcomes, we calculated regression adjusted means for the baseline and Test periods for each Test state and its comparison group. For binary outcomes estimated using nonlinear models, the regression-adjusted D-in-Ds are calculated as the average treatment effect *on the treated* (ATET), whereas the D-in-D derived from the adjusted means represents the average treatment effect. As a result, the regression-adjusted D-in-D and the D-in-D calculated from the adjusted means will differ.

Unadjusted averages

The underlying assumption in D-in-D models estimating the changes that occur in Test states after the SIM Initiative is implemented is that trends in the outcomes among individuals in Test states and their respective comparison groups would be similar absent the SIM Initiative (i.e., that the two were on "parallel paths" prior to the start of the SIM Initiative).

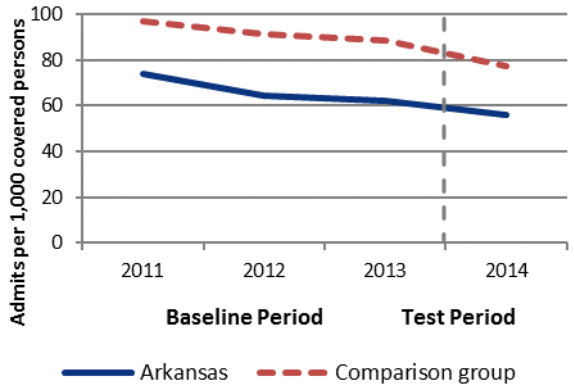
To examine descriptively whether the trends in the test and comparison groups are parallel, we present graphs of annual, unadjusted averages (weighted by the propensity score and eligibility fraction) for the four key outcomes (inpatient admissions per 1,000 covered persons, ED visits per 1,000 covered persons, readmissions per 1,000 discharges, and total PMPM expenditures) by payer for each relevant state (***Figures G-1*** through ***G-56*** below).

We concluded that we cannot assume parallel trends across all outcomes and payers. Therefore, we generate estimates that net out the potential baseline differences between the Test state and the comparison group. Specifically, we include a linear time trend interacted with the dichotomous variable for residing in the Test state in the outcomes model. This model specification allows for differences in estimates in the Round 1 Test states and their comparison groups during the baseline period, and it allows for a straightforward interpretation of the D-in-D coefficient. In this way, the alternative D-in-D model can be thought of as a comparative interrupted time series or structural break equation that captures whether trends in Test states changed relative to the comparison group after the introduction of the SIM Initiative.

Medicaid

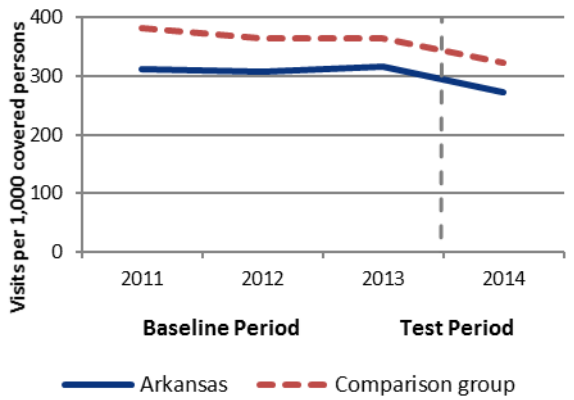
Arkansas

Figure G-1. Average number of inpatient admissions per 1,000 beneficiaries, FY 2011–FY 2014, Arkansas Medicaid beneficiaries and the comparison group



The average number of **inpatient admissions per 1,000 beneficiaries** declined from 2011 to 2014 for Medicaid beneficiaries in both Arkansas and the comparison group.

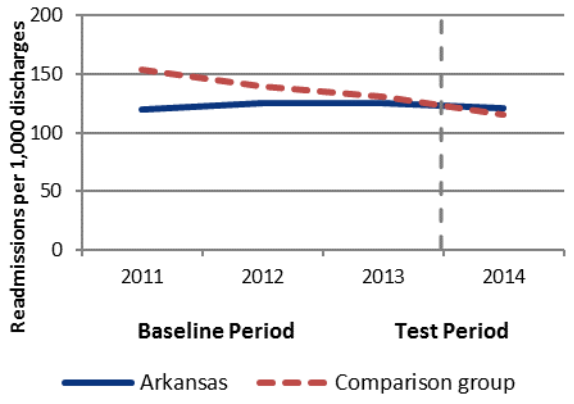
Figure G-2. Average number of emergency department visits that did not lead to a hospitalization per 1,000 beneficiaries, FY 2011–FY 2014, Arkansas Medicaid beneficiaries and the comparison group



The average number of **outpatient ED visits per 1,000 beneficiaries** remained relatively stable from 2011 to 2013, then declined from 2013 to 2014.

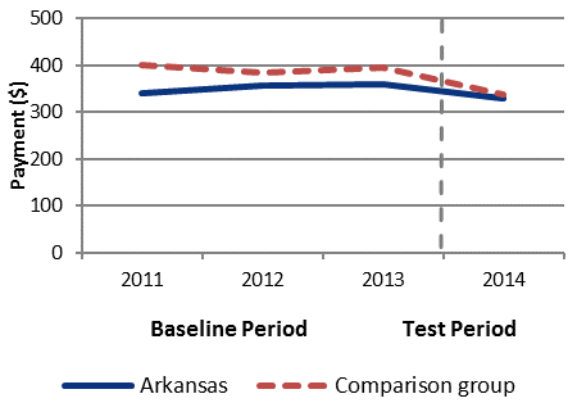
ED = emergency department.

Figure G-3. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2014, Arkansas Medicaid beneficiaries and the comparison group



The average number of **30-day readmissions** per 1,000 discharges increased for Arkansas Medicaid beneficiaries over the baseline period while declining for the comparison group. As a result, the rate for the two groups converged in 2013 and remained similar in 2014.

Figure G-4. Total PMPM expenditures, FY 2011–FY 2014, Arkansas Medicaid beneficiaries and the comparison group

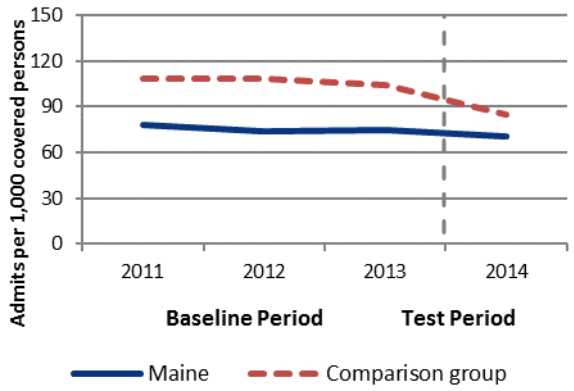


Total PMPM expenditures were lower for Arkansas relative to the comparison group throughout the baseline period and Test period. For both groups, the total PMPM expenditures remained relative stable until 2013, then declined from 2013 to 2014.

PMPM = per member per month.

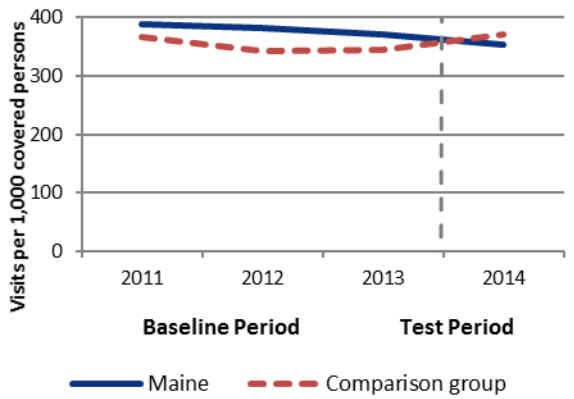
Maine

Figure G-5. Average number of inpatient admissions per 1,000 beneficiaries, FY 2011–FY 2014, Maine Medicaid beneficiaries and the comparison group



The average number of **inpatient admissions per 1,000 beneficiaries** was lower in Maine relative to the comparison group throughout the baseline and Test periods. The admission rate remained stable in Maine throughout the study period, but the rate declined for the comparison group from 2013 to 2014.

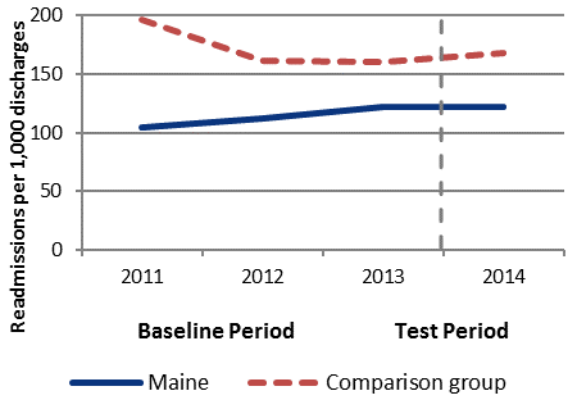
Figure G-6. Average number of emergency department visits that did not lead to a hospitalization per 1,000 beneficiaries, FY 2011–FY 2014, Maine Medicaid beneficiaries and the comparison group



The average number of **outpatient ED visits per 1,000 beneficiaries** was higher in Maine relative to the comparison group until 2014, at which point the ED visit rate increased for the comparison group while declining slightly for Maine Medicaid beneficiaries.

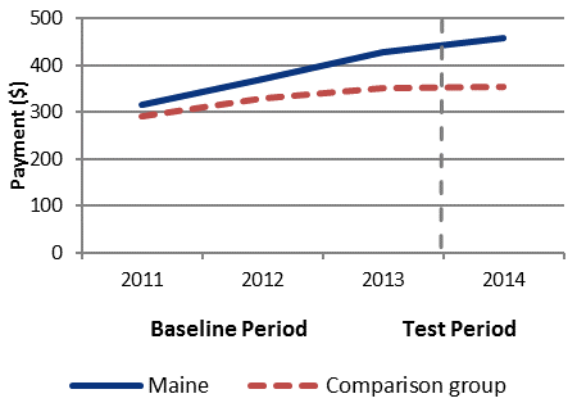
ED = emergency department.

Figure G-7. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2014, Maine Medicaid beneficiaries and the comparison group



The average number of **30-day readmissions** per 1,000 discharges was lower for Maine relative to the comparison group for the entire study period. For Maine Medicaid beneficiaries, the readmission rate increased slightly from 2011 to 2013 then remained stable until 2014. For the comparison group, the readmission rate declined from 2011 to 2012, then remained stable until 2014.

Figure G-8. Total PMPM expenditures, FY 2011–FY 2014, Maine Medicaid beneficiaries and the comparison group

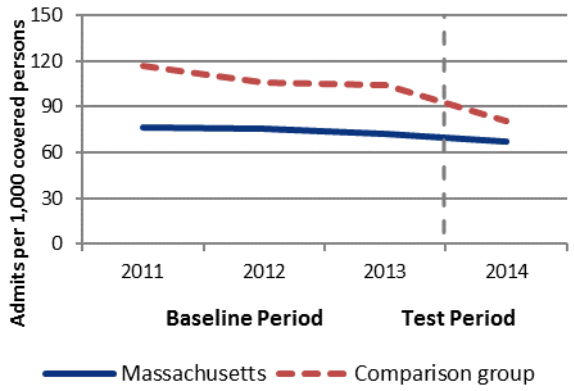


Total PMPM expenditures were higher for Maine relative to the comparison group throughout the baseline period and Test period. For both groups, the total PMPM expenditures increased throughout the baseline period, though the increase was greater for Maine Medicaid beneficiaries. Total PMPM expenditures continued to increase in the Test period for Maine, but remained stable for the comparison group.

PMPM = per member per month.

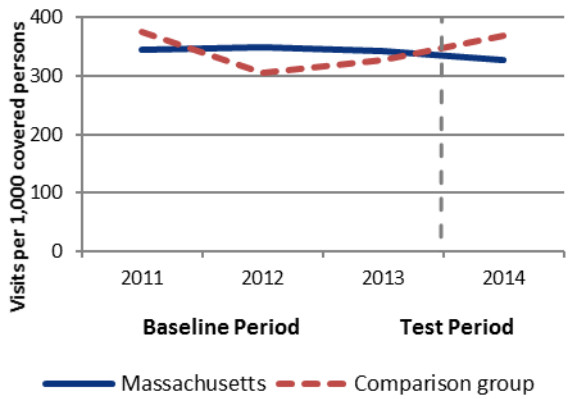
Massachusetts

Figure G-9. Average number of inpatient admissions per 1,000 beneficiaries, 2011–2014, Massachusetts Medicaid beneficiaries and the comparison group



The average number of **inpatient admissions per 1,000 beneficiaries** was lower in Massachusetts relative to the comparison group throughout the baseline and Test periods. The admission rate remained stable in Massachusetts throughout the study period, but the rate declined for the comparison group from 2013 to 2014.

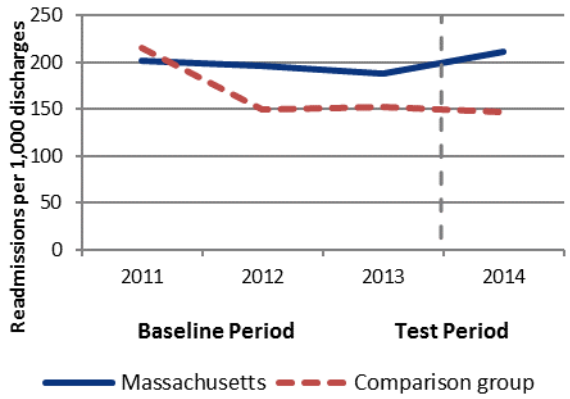
Figure G-10. Average number of emergency department visits that did not lead to a hospitalization per 1,000 beneficiaries, 2011–2014, Massachusetts Medicaid beneficiaries and the comparison group



The average number of **outpatient ED visits per 1,000 beneficiaries** remained relatively stable for Medicaid beneficiaries in Massachusetts, but the ED visit rate declined from 2011 to 2012 then increased from 2012 to 2014 for the comparison group.

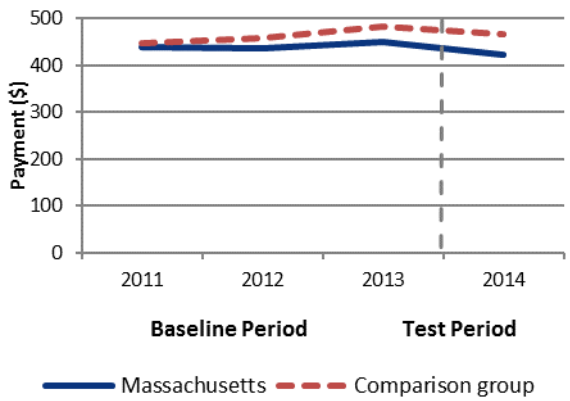
ED = emergency department.

Figure G-11. Average number of 30-day readmissions per 1,000 discharges, 2011–2014, Massachusetts Medicaid beneficiaries and the comparison group



The average number of **30-day readmissions** per 1,000 discharges remained stable from 2011 to 2013 then increased in 2014 for Massachusetts Medicaid beneficiaries. For the comparison group, the readmission rate declined from 2011 to 2012, then remained stable until 2014.

Figure G-12. Total PMPM expenditures, 2011–2014, Massachusetts Medicaid beneficiaries and the comparison group

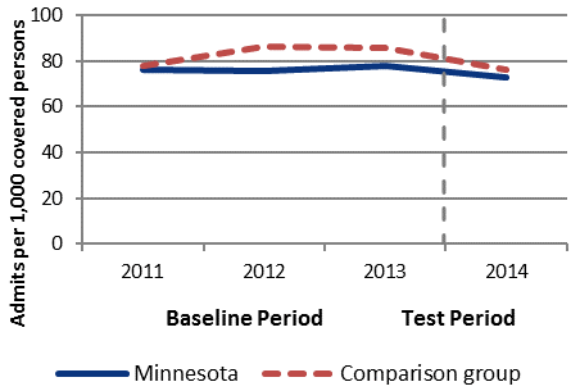


Total PMPM expenditures were slightly lower for Massachusetts relative to the comparison group throughout the baseline period and Test period. For both groups, the total PMPM expenditures increased slightly throughout the baseline period, then declined slightly in 2014.

PMPM = per member per month.

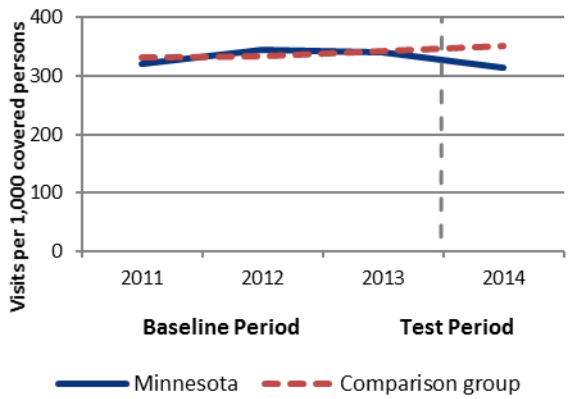
Minnesota

Figure G-13. Average number of inpatient admissions per 1,000 beneficiaries, FY 2011–FY 2014, Minnesota Medicaid beneficiaries and the comparison group



The average number of **inpatient admissions per 1,000 beneficiaries** was lower in Minnesota relative to the comparison group throughout the baseline and Test periods. The admission rate remained stable in Minnesota throughout the baseline period, then declined slightly in 2014. For the comparison group, the rate increased from 2011 to 2012, then declined from 2013 to 2014.

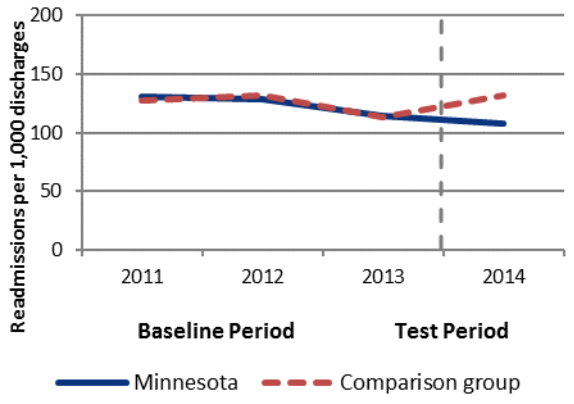
Figure G-14. Average number of emergency department visits that did not lead to a hospitalization per 1,000 beneficiaries, FY 2011–FY 2014, Minnesota Medicaid beneficiaries and the comparison group



The average number of **outpatient ED visits per 1,000 beneficiaries** was similar in Minnesota relative to the comparison group until 2014, at which point the ED visit rate increased for the comparison group while declining slightly for Minnesota Medicaid beneficiaries.

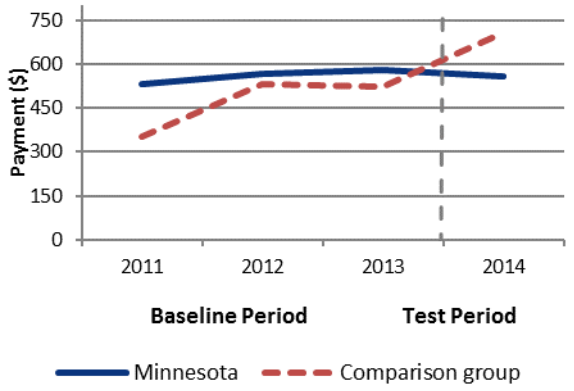
ED = emergency department.

Figure G-15. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2014, Minnesota Medicaid beneficiaries and the comparison group



The average number of **30-day readmissions** per 1,000 discharges was similar for Minnesota relative to the comparison group until 2014, at which point the readmission rate increased for the comparison group while declining slightly for Minnesota Medicaid beneficiaries.

Figure G-16. Total PMPM expenditures, FY 2011–FY 2014, Minnesota Medicaid beneficiaries and the comparison group

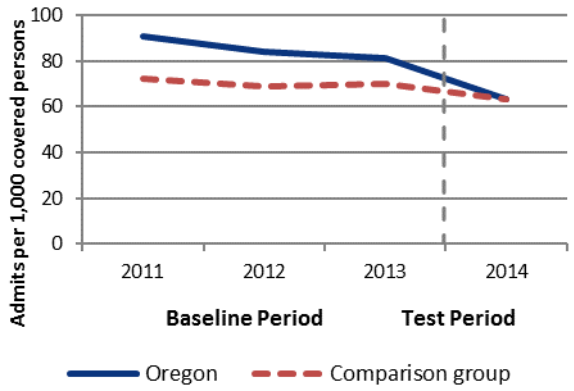


Total PMPM expenditures were higher for Minnesota relative to the comparison group throughout the baseline period, then expenditures were higher for the comparison group during the test period. In Minnesota, total PMPM expenditures remained relatively stable throughout the entire study period. For the comparison group, expenditures increased from 2011 to 2012 then again from 2013 to 2014.

PMPM = per member per month.

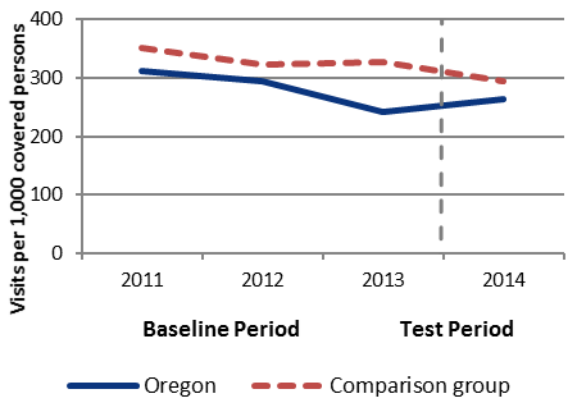
Oregon

Figure G-17. Average number of inpatient admissions per 1,000 beneficiaries, FY 2011–FY 2014, Oregon Medicaid beneficiaries and the comparison group



The average number of **inpatient admissions per 1,000 beneficiaries** was higher in Oregon relative to the comparison group throughout the baseline period, then the two groups had the same rate in the test period. The admission rate declined slightly in Oregon throughout the baseline period, then declined sharply in 2014. For the comparison group, the rate remained relatively stable with a slight decline in the test period.

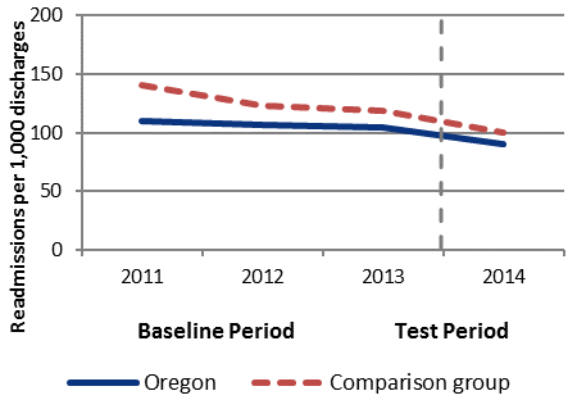
Figure G-18. Average number of emergency department visits that did not lead to a hospitalization per 1,000 beneficiaries, FY 2011–FY 2014, Oregon Medicaid beneficiaries and the comparison group



The average number of **outpatient ED visits per 1,000 beneficiaries** was lower in Oregon relative to the comparison group throughout the study period. For Oregon beneficiaries, the ED visit rate declined slightly from 2011 to 2012, then declined more sharply from 2012 to 2013 before increasing in 2014. For the comparison group, the ED visit rate declined slightly from 2011 to 2012 then again from 2013 to 2014.

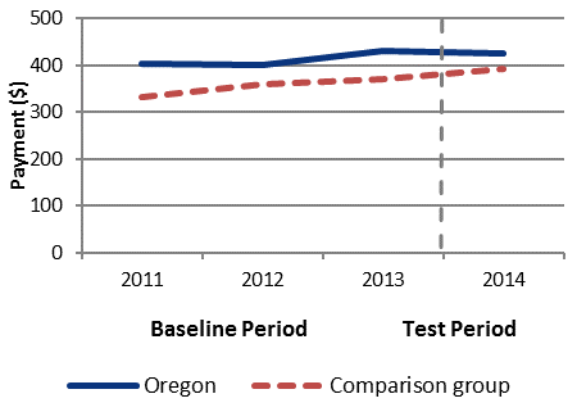
ED = emergency department.

Figure G-19. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2014, Oregon Medicaid beneficiaries and the comparison group



The average number of **30-day readmissions** per 1,000 discharges was lower for Oregon relative to the comparison group throughout the study period. The readmission rate declined in both groups over the study period.

Figure G-20. Total PMPM expenditures, FY 2011–FY 2014, Oregon Medicaid beneficiaries and the comparison group

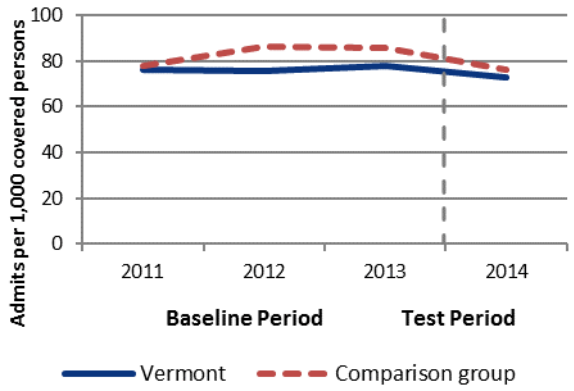


Total PMPM expenditures were higher for Oregon relative to the comparison group throughout the study period. Total PMPM expenditures increased for both groups throughout the study period.

PMPM = per member per month.

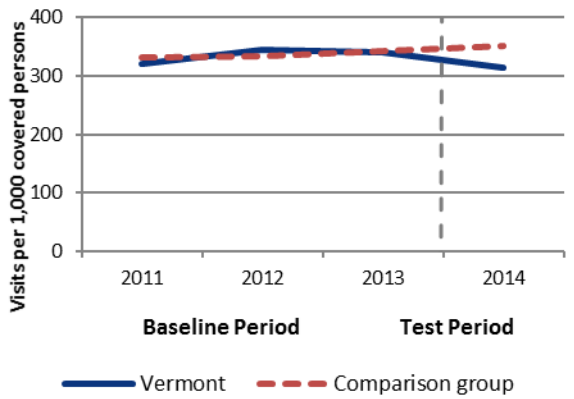
Vermont

Figure G-21. Average number of inpatient admissions per 1,000 beneficiaries, FY 2011–FY 2014, Vermont Medicaid beneficiaries and the comparison group



The average number of **inpatient admissions per 1,000 beneficiaries** was lower in Vermont relative to the comparison group throughout the study period. For Vermont, the admission rate remained stable throughout the baseline period then declined slightly in 2014. For the comparison group, the admission rate increased slightly from 2011 to 2012 then declined from 2013 to 2014.

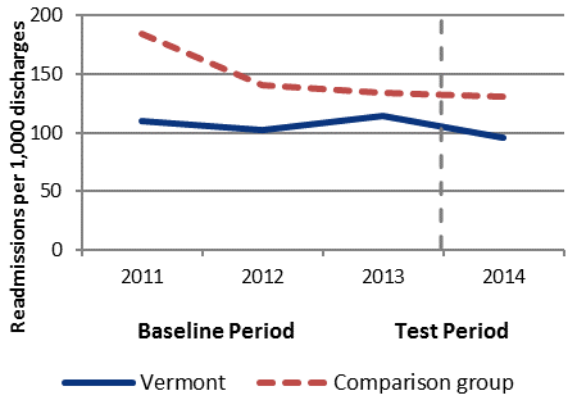
Figure G-22. Average number of emergency department visits that did not lead to a hospitalization per 1,000 beneficiaries, FY 2011–FY 2014, Vermont Medicaid beneficiaries and the comparison group



The average number of **outpatient ED visits per 1,000 beneficiaries** was similar in Vermont relative to the comparison group until 2014, at which point the ED visit rate increased for the comparison group while declining slightly for Vermont Medicaid beneficiaries.

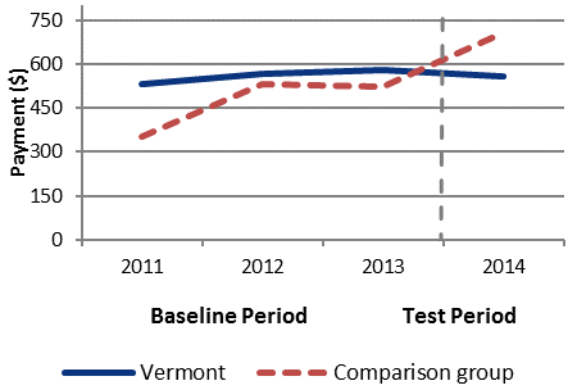
ED = emergency department.

Figure G-23. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2014, Vermont Medicaid beneficiaries and the comparison group



The average number of **30-day readmissions** per 1,000 discharges was lower for Vermont relative to the comparison group throughout the study period. The readmission rate increased in 2013 then declined in 2014 for Vermont. For the comparison group, the readmission rate declined from 2011 to 2012 then remained stable through 2014.

Figure G-24. Total PMPM expenditures, FY 2011–FY 2014, Vermont Medicaid beneficiaries and the comparison group



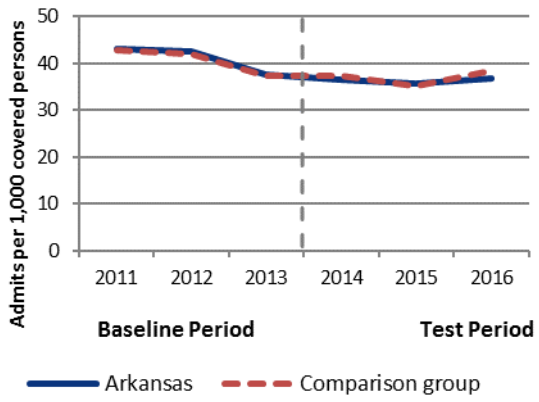
Total PMPM expenditures remained relatively stable in Vermont throughout the study period; expenditures increased slightly through 2013 then declined slightly in 2014. For the comparison group, total PMPM expenditures increased from 2011 to 2012 then again from 2013 to 2014.

PMPM = per member per month.

MarketScan

Arkansas

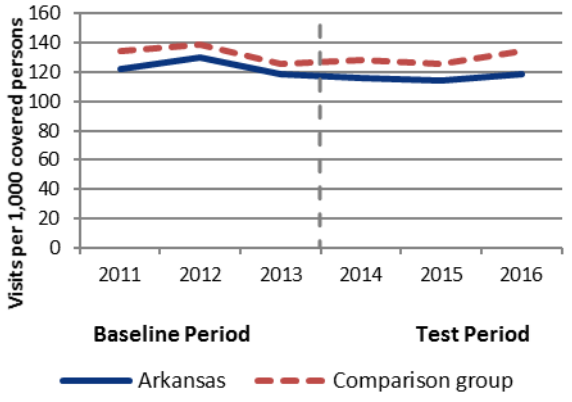
Figure G-25. Average number of inpatient admissions per 1,000 covered persons, FY 2011–FY 2016, Arkansas commercial plan members and the comparison group



The average number of **inpatient admissions per 1,000 covered persons** declined slightly from 2011 to 2016 for commercial plan members in both Arkansas and the comparison group.

MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

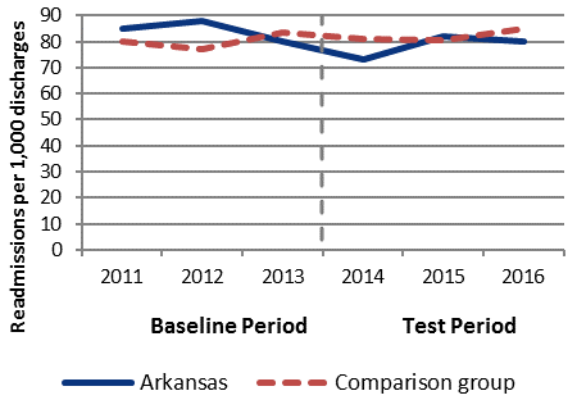
Figure G-26. Average number of emergency department visits that did not lead to a hospitalization per 1,000 covered persons, FY 2011–FY 2016, Arkansas commercial plan members and the comparison group



The average number of **outpatient ED visits per 1,000 covered persons** remained relatively stable from 2011 to 2016, with consistently lower rates among the Arkansas group relative to the comparison group.

ED = emergency department.
MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

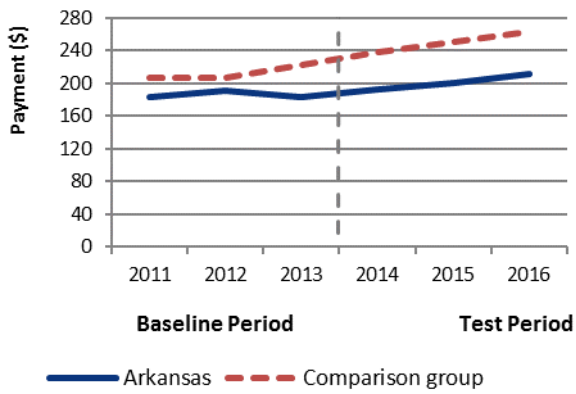
Figure G-27. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2016, Arkansas commercial plan members and the comparison group



The average number of **30-day readmissions** per 1,000 discharges decreased and then increased for Arkansas commercial plan members over the baseline and test periods, while remaining relatively stable for the comparison group. As a result, the rate for the two groups were similar at the end of the test period in 2015 and 2016.

MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

Figure G-28. Total PMPM expenditures, FY 2011–FY 2016, Arkansas commercial plan members and the comparison group

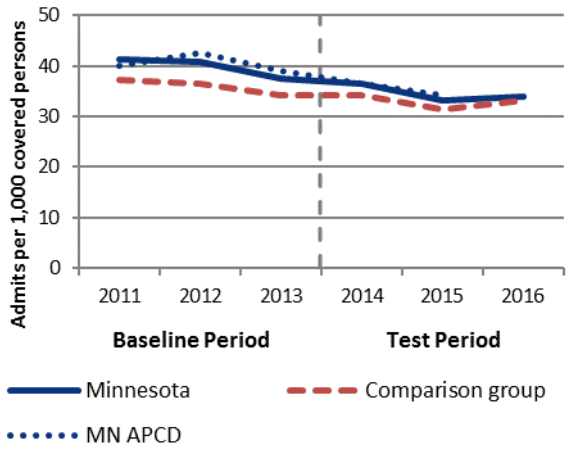


Total PMPM expenditures were lower for Arkansas relative to the comparison group throughout the baseline and test periods. For both groups, the total PMPM expenditures increased from 2011 through 2016, though the rate of increase was faster among the comparison group.

PMPM = per member per month.
MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

Minnesota

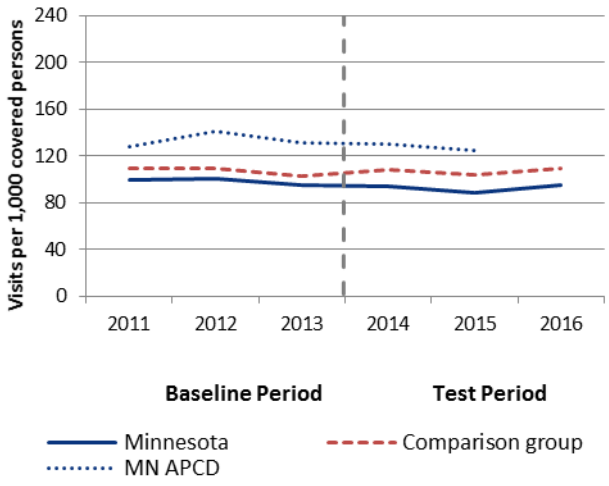
Figure G-29. Average number of inpatient admissions per 1,000 covered persons, FY 2011–FY 2016, Minnesota commercial plan members and the comparison group



The average number of **inpatient admissions per 1,000 covered persons** declined from 2011 to 2016 for commercial plan members in both Minnesota and the comparison group. The trend line calculated from the MN APCD showed a similar decline to the Minnesota commercial plan members in the MarketScan data.

MN APCD = Minnesota All Payer Claims Database.
MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

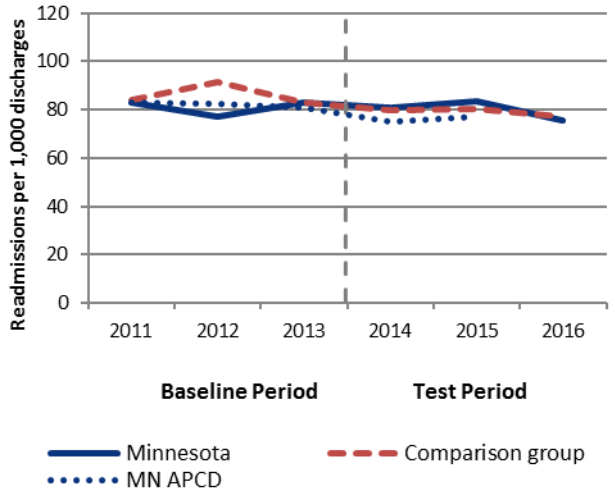
Figure G-30. Average number of emergency department visits that did not lead to a hospitalization per 1,000 covered persons, FY 2011–FY 2016, Minnesota commercial plan members and the comparison group



The average number of **outpatient ED visits per 1,000 covered persons** remained relatively stable from 2011 to 2016, with consistently lower rates among the Minnesota group relative to the comparison group. The trend line calculated from the MN APCD was consistently higher than the rate for commercial plan members in the MarketScan data.

MN APCD = Minnesota All Payer Claims Database.
MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

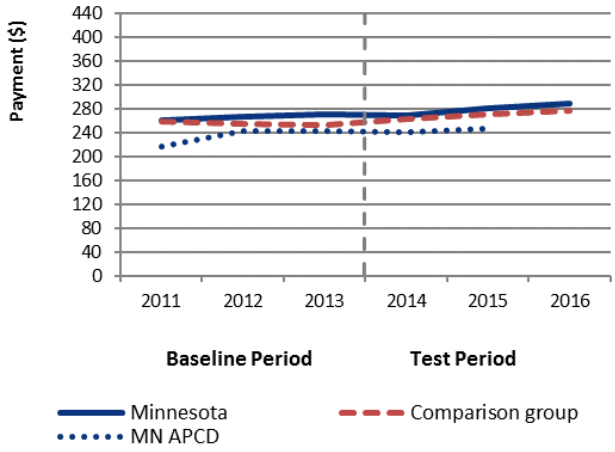
Figure G-31. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2016, Minnesota commercial plan members and the comparison group



The average number of **30-day readmissions** per 1,000 discharges increased and then decreased during the baseline period for the comparison group, while remaining relatively stable for Minnesota commercial plan. The rates for the two groups converged in 2013 and remained similar throughout the test period. The rate calculated from the MN APCD remained relatively stable, but declined slightly from 2011 to 2015.

MN APCD = Minnesota All Payer Claims Database.
MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

Figure G-32. Total PMPM expenditures, FY 2011–FY 2016, Minnesota commercial plan members and the comparison group

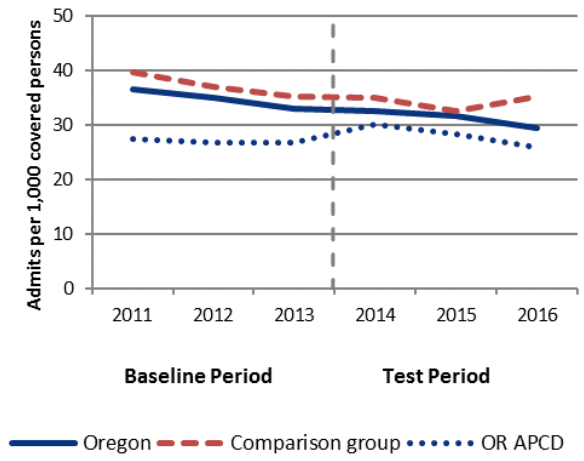


Total PMPM expenditures increased slightly from 2011 to 2016 for commercial plan members in both the Minnesota and comparison group. The MN APCD-derived expenditures also increased slightly from 2011 to 2015, although the expenditures calculated from the MN APCD data were consistently lower than the expenditures for the commercial plan members in the MarketScan data.

MN APCD = Minnesota All Payer Claims Database.
MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

Oregon

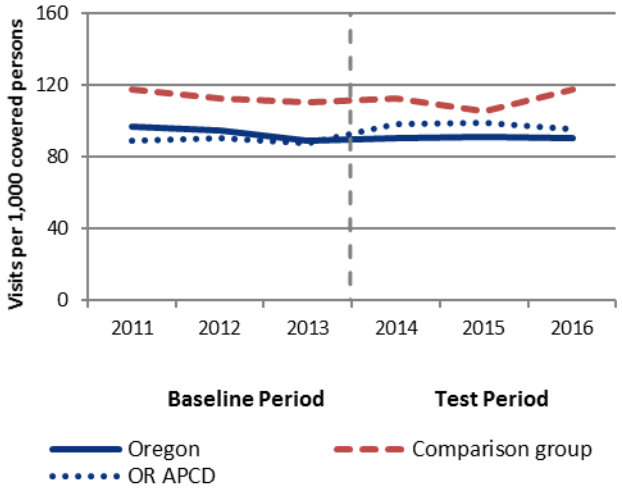
Figure G-33. Average number of inpatient admissions per 1,000 covered persons, FY 2011–FY 2016, Oregon commercial plan members and the comparison group



The average number of **inpatient admissions per 1,000 covered persons** declined from 2011 to 2015 for commercial plan members in both Oregon and the comparison group. In 2016, the Oregon rate continued to decline while the comparison group saw a small increase. The trend line calculated from the OR APCD remained flat through most of the baseline period, then increased from 2013 to 2014 before declining from 2014 to 2016.

OR APCD = Oregon All Payer All Claims database.
MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

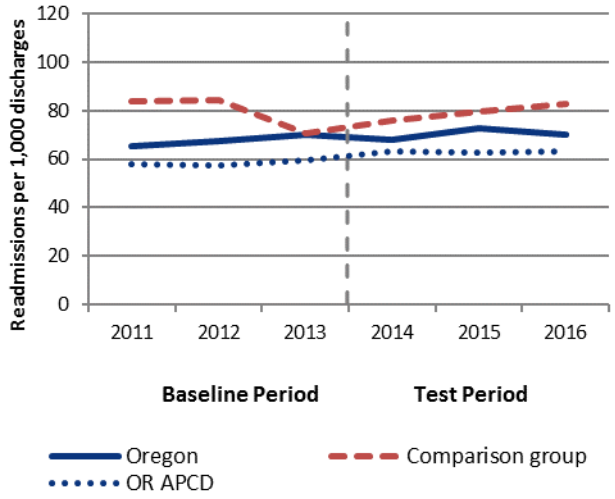
Figure G-34. Average number of emergency department visits that did not lead to a hospitalization per 1,000 covered persons, FY 2011–FY 2016, Oregon commercial plan members and the comparison group



The average number of **outpatient ED visits per 1,000 covered persons** remained relatively stable from 2011 to 2016 for commercial plan members in Oregon. The comparison group rate declined gradually through 2015 but increased again in 2016, resulting in no net change overall from 2011 to 2016. The ED visit rate calculated from the OR APCD remained stable through the baseline period then increased from 2013 to 2014 then remained stable through 2016.

OR APCD = Oregon All Payer All Claims database.
MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

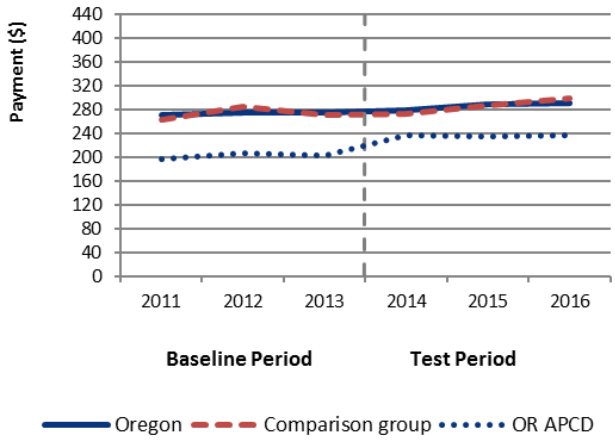
Figure G-35. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2016, Oregon commercial plan members and the comparison group



The average number of **30-day readmissions** per 1,000 discharges increased for Oregon commercial plan members over the baseline period while declining for the comparison group. As a result, the rate for the two groups converged in 2013. In the test period, the rate remained stable for Oregon commercial plan members and increased for the comparison group. The rate calculated from the OR APCD increased slightly from 2011 to 2016.

OR APCD = Oregon All Payer All Claims database.
MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

Figure G-36. Total PMPM expenditures, FY 2011–FY 2016, Oregon commercial plan members and the comparison group

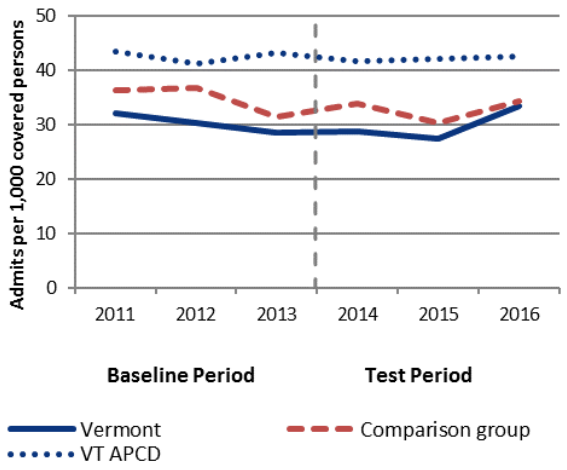


Total PMPM expenditures remained relatively stable from 2011 to 2016 for commercial plan members in Oregon and the comparison group, with a slight increase overall in both groups. Expenditures calculated from the OR APCD remained stable from 2011 to 2012 then increased from 2013 to 2014, then remained stable through 2016.

OR APCD = Oregon All Payer All Claims database.
MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company.

Vermont

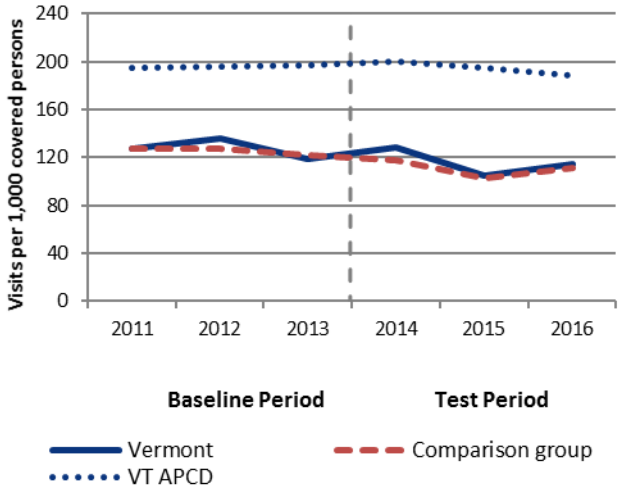
Figure G-37. Average number of inpatient admissions per 1,000 covered persons, FY 2011–FY 2016, Vermont commercial plan members and the comparison group



The average number of **inpatient admissions per 1,000 covered persons** declined slightly from 2011 to 2015 for commercial plan members in both Vermont and the comparison group, before increasing for both groups in 2016. The rate calculated from the VT APCD remained relatively stable across the baseline and test periods and was consistently higher than the commercial plan rate.

VT APCD = Vermont All Payer Claims Database (VHCURES). MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company. The Green Mountain Care Board (GMCB) and Vermont Health Care Uniform Reporting and Evaluation System (VHCURES) are the source for Vermont APCD data. Analyses are solely those of the SIM Round 1 Evaluation team and not GMCB.

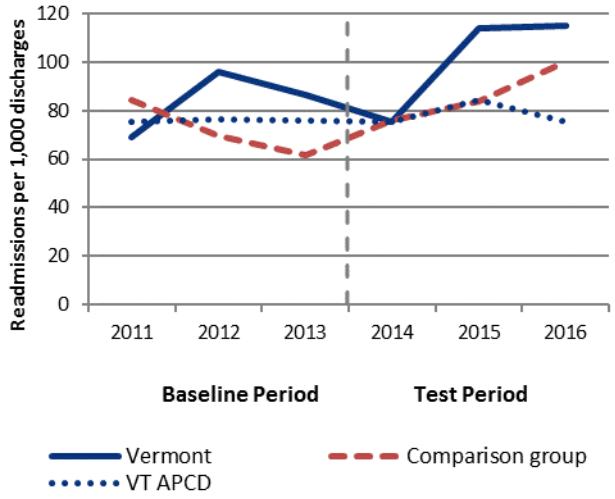
Figure G-38. Average number of emergency department visits that did not lead to a hospitalization per 1,000 covered persons, FY 2011–FY 2016, Vermont commercial plan members and the comparison group



The average number of **outpatient ED visits per 1,000 covered persons** remained relatively stable from 2011 to 2016 for commercial plan members in both Vermont and the comparison group, with a slight overall decline. The rate also remained stable during this period, although there was a slight decline from 2015 to 2016, and the rate calculated from the VT APCD was consistently higher than the commercial plan rate.

VT APCD = Vermont All Payer Claims Database (VHCURES). MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company. The Green Mountain Care Board (GMCB) and Vermont Health Care Uniform Reporting and Evaluation System (VHCURES) are the source for Vermont APCD data. Analyses are solely those of the SIM Round 1 Evaluation team and not GMCB.

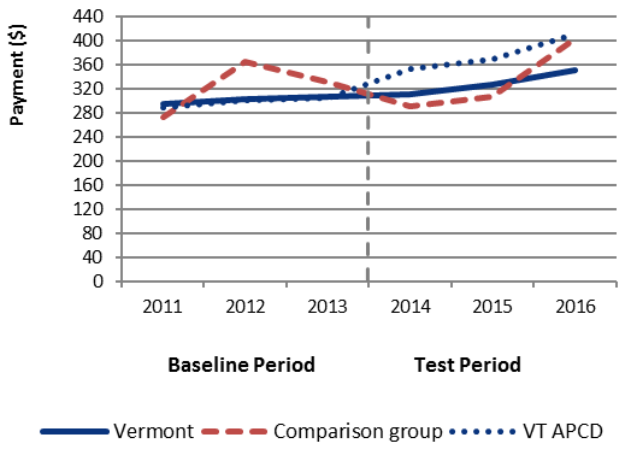
Figure G-39. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2016, Vermont commercial plan members and the comparison group



The average number of **30-day readmissions** per 1,000 discharges increased and then decreased for Vermont commercial plan members over the baseline period, followed by a sharp increase in the test period in 2015. The comparison group rate declined overall during the baseline period and increased steadily during the test period. The rate calculated from the VT APCD remained stable overall through the baseline and test periods, with a slight decline in 2016. Both the comparison group and APCD-derived rates converged with the Vermont rate in 2014, but the Vermont rate was higher than both rates for the remainder of the test period due to the sharp 2015 increase.

VT APCD = Vermont All Payer Claims Database (VHCURES). MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company. The Green Mountain Care Board (GMCB) and Vermont Health Care Uniform Reporting and Evaluation System (VHCURES) are the source for Vermont APCD data. Analyses are solely those of the SIM Round 1 Evaluation team and not GMCB.

Figure G-40. Total PMPM expenditures, FY 2011–FY 2016, Vermont commercial plan members and the comparison group



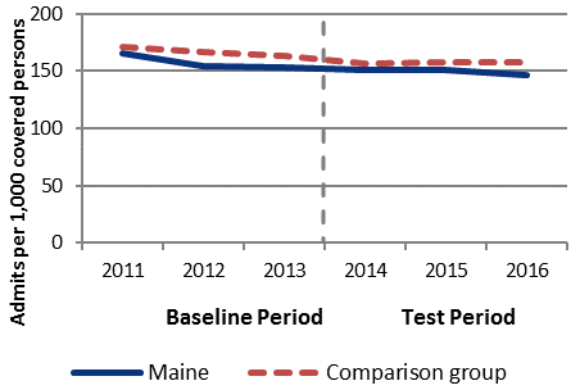
Total PMPM expenditures were relatively stable for both Vermont and the APCD group throughout the baseline period, with increases in both during the test period. Expenditures calculated from the VT APCD increased more rapidly than in the Vermont group, with higher APCD expenditures in the test period. The comparison group had both increases and decreased, but a net increase in expenditures across the baseline and test period.

VT APCD = Vermont All Payer Claims Database (VHCURES). MarketScan is ©2016 Truven Health Analytics Inc., an IBM Company. The Green Mountain Care Board (GMCB) and Vermont Health Care Uniform Reporting and Evaluation System (VHCURES) are the source for Vermont APCD data. Analyses are solely those of the SIM Round 1 Evaluation team and not GMCB.

Medicare

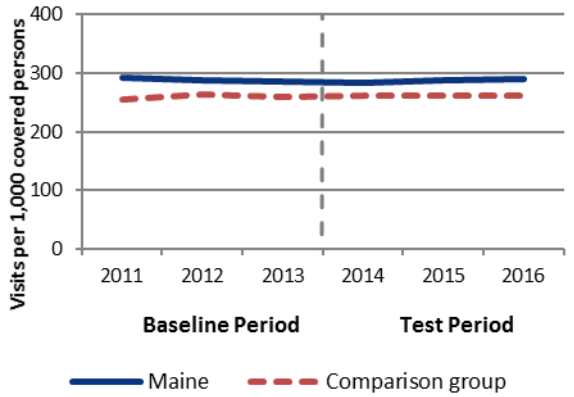
Maine

Figure G-41. Average number of inpatient admissions per 1,000 beneficiaries, FY 2011–FY 2016, Maine Medicare beneficiaries and the comparison group



The average number of **inpatient admissions per 1,000 beneficiaries** declined from 2011 to 2016 for Medicare beneficiaries in both Maine and the comparison group.

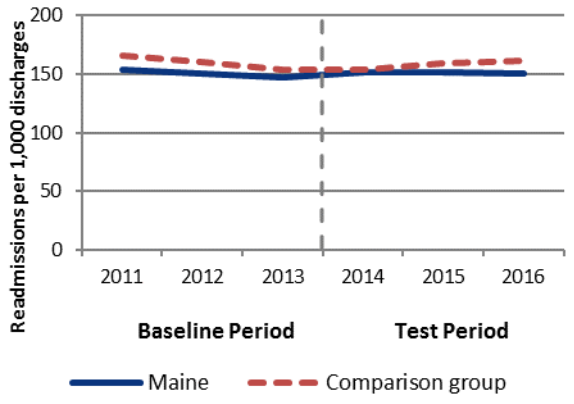
Figure G-42. Average number of emergency department visits that did not lead to a hospitalization per 1,000 beneficiaries, FY 2011–FY 2016, Maine Medicare beneficiaries and the comparison group



The average number of **outpatient ED visits per 1,000 beneficiaries** remained relatively stable from 2011 to 2016.

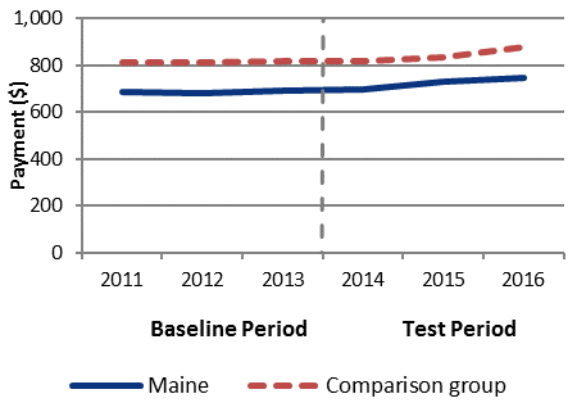
ED = emergency department.

Figure G-43. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2016, Maine Medicare beneficiaries and the comparison group



The average number of **30-day readmissions** per 1,000 discharges was stable for Maine Medicare beneficiaries over the baseline period while declining for the comparison group. As a result, the rate for the two groups converged in 2013. In the test period, the rate for Maine Medicare beneficiaries continued to remain stable while the rate for the comparison group increased.

Figure G-44. Total PMPM expenditures, FY 2011–FY 2016, Maine Medicare beneficiaries and the comparison group

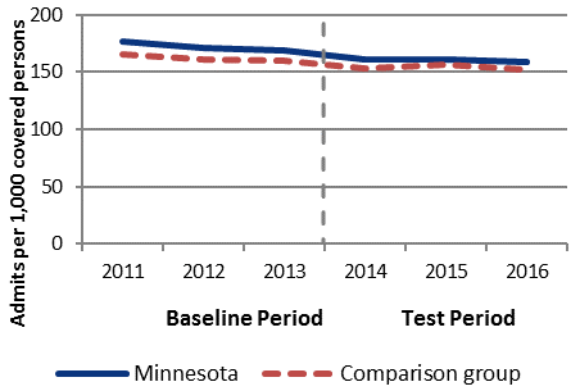


Total PMPM expenditures were lower for Maine Medicare beneficiaries relative to the comparison group throughout the baseline period and test period. For both groups, the total PMPM expenditures remained relative stable until 2015, when expenditures increased.

PMPM = per member per month.

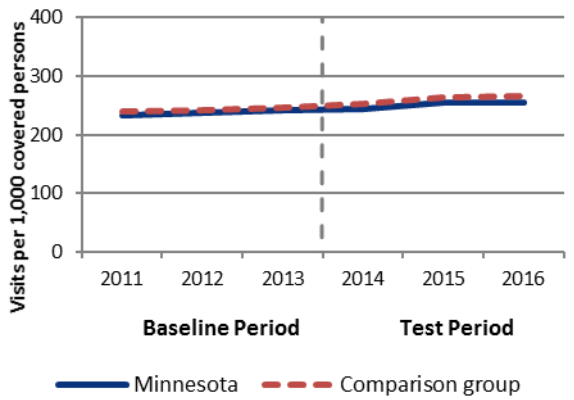
Minnesota

Figure G-45. Average number of inpatient admissions per 1,000 beneficiaries, FY 2011–FY 2016, Minnesota Medicare beneficiaries and the comparison group



The average number of **inpatient admissions per 1,000 beneficiaries** declined from 2011 to 2014 and then remained relatively stable from 2015 to 2016 for Medicare beneficiaries in both Minnesota and the comparison group.

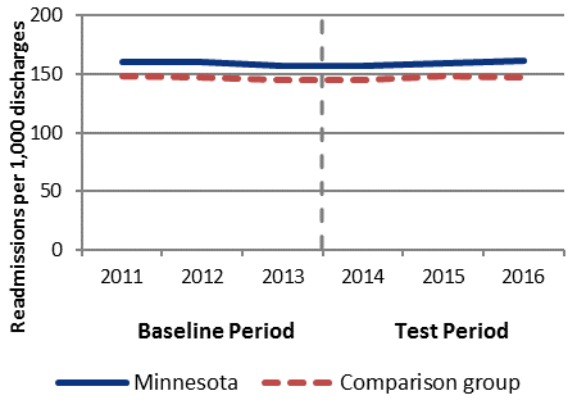
Figure G-46. Average number of emergency department visits that did not lead to a hospitalization per 1,000 beneficiaries, FY 2011–FY 2016, Minnesota Medicare beneficiaries and the comparison group



The average number of **outpatient ED visits per 1,000 beneficiaries** remained relatively stable in the baseline period and slightly increased in the test period.

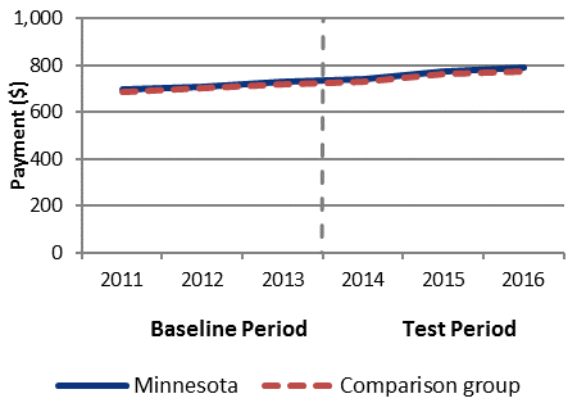
ED = emergency department.

Figure G-47. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2016, Minnesota Medicare beneficiaries and the comparison group



The average number of **30-day readmissions** per 1,000 discharges remained stable from 2011 to 2016 for Minnesota Medicare and comparison group beneficiaries. The rate was consistently higher among Minnesota Medicare beneficiaries relative to the comparison group in each year.

Figure G-48. Total PMPM expenditures, FY 2011–FY 2016, Minnesota Medicare beneficiaries and the comparison group

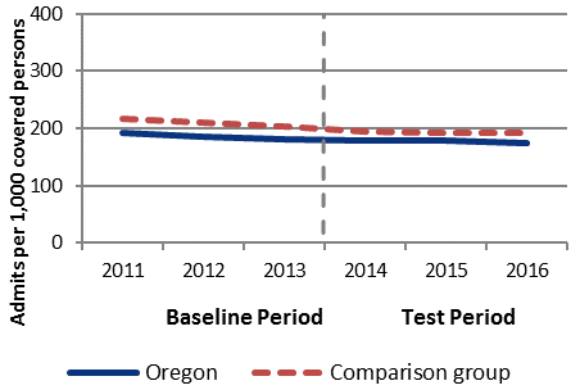


Total PMPM expenditures steadily increased for Minnesota Medicare and comparison group beneficiaries from 2011 to 2016 with similar rates in both groups in each year.

PMPM = per member per month.

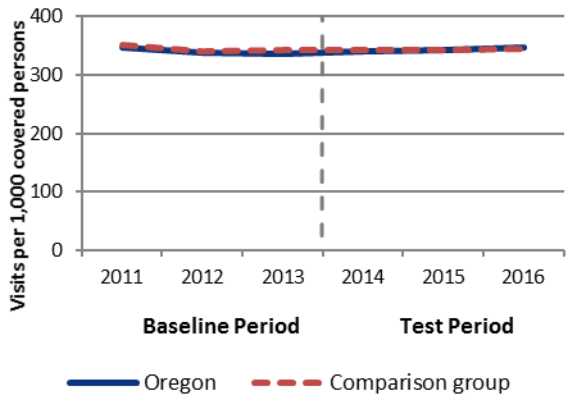
Oregon

Figure G-49. Average number of inpatient admissions per 1,000 beneficiaries, FY 2011–FY 2016, Oregon Medicare beneficiaries dually enrolled in Medicaid and the comparison group



The average number of **inpatient admissions per 1,000 beneficiaries** slightly deceased from 2011 to 2016 for Medicare beneficiaries dually enrolled in Medicaid in both Oregon and the comparison group. The rate was consistently lower among Oregon Medicare beneficiaries dually enrolled in Medicaid relative to the comparison group in each year.

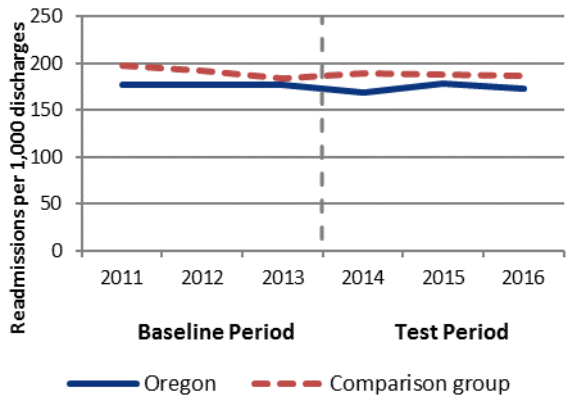
Figure G-50. Average number of emergency department visits that did not lead to a hospitalization per 1,000 beneficiaries, FY 2011–FY 2016, Oregon Medicare beneficiaries dually enrolled in Medicaid and the comparison group



The average number of **outpatient ED visits per 1,000 beneficiaries** remained relatively stable from 2011 to 2016 with similar rates in both groups in each year.

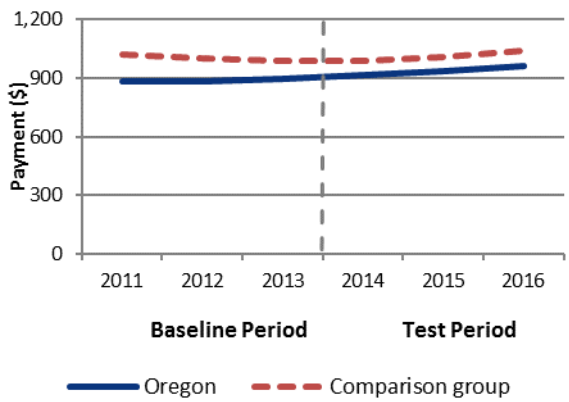
ED = emergency department.

Figure G-51. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2016, Oregon Medicare beneficiaries dually enrolled in Medicaid and the comparison group



The average number of **30-day readmissions** per 1,000 discharges remained stable for Oregon Medicare beneficiaries dually enrolled in Medicaid over the baseline period while declining for the comparison group. As a result, the rate for the two groups converged in 2013. The rate for Oregon Medicare beneficiaries dually enrolled in Medicaid fluctuated in the test period, decreasing from 2013 to 2014, increasing from 2014 to 2015, and decreasing from 2015 to 2016. The rate for the comparison group, however, remained relatively stable in the test period.

Figure G-52. Total PMPM expenditures, FY 2011–FY 2016, Oregon Medicare beneficiaries dually enrolled in Medicaid and the comparison group

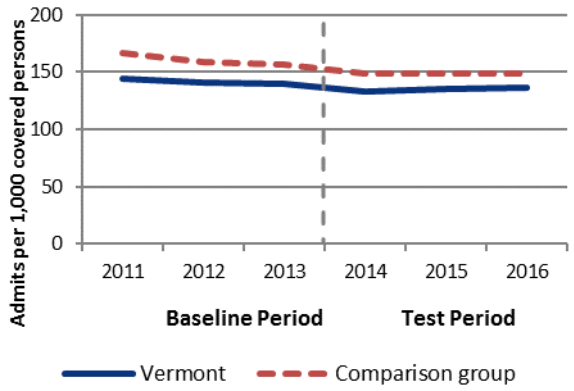


Total PMPM expenditures were lower for Oregon Medicare beneficiaries dually enrolled in Medicaid relative to the comparison group throughout the baseline period and test period. For Oregon beneficiaries, the total PMPM expenditures increased from 2011 to 2016, whereas expenditures remained relatively stable for the comparison group from 2011 to 2016.

PMPM = per member per month.

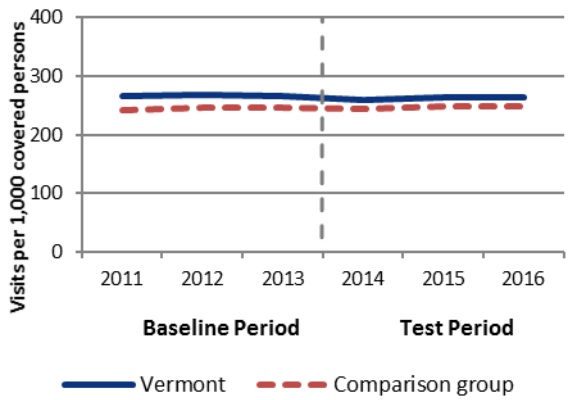
Vermont

Figure G-53. Average number of inpatient admissions per 1,000 beneficiaries, FY 2011–FY 2016, Vermont Medicare beneficiaries and the comparison group



The average number of **inpatient admissions per 1,000 beneficiaries** declined from 2011 to 2014 and remained relatively stable from 2015 to 2016 for Medicare beneficiaries in both Vermont and the comparison group. The rate was lower for Vermont Medicare beneficiaries relative to the comparison group throughout the baseline period and test period.

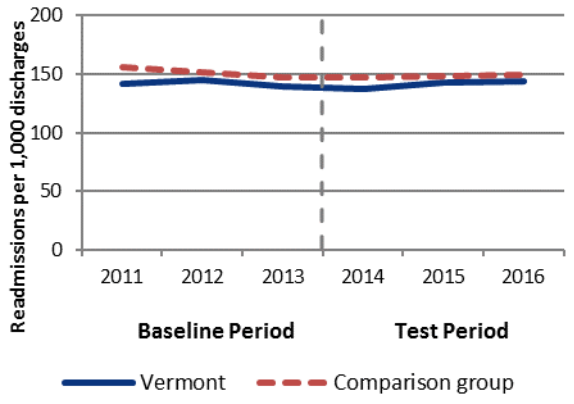
Figure G-54. Average number of emergency department visits that did not lead to a hospitalization per 1,000 beneficiaries, FY 2011–FY 2016, Vermont Medicare beneficiaries and the comparison group



The average number of **outpatient ED visits per 1,000 beneficiaries** remained relatively stable from 2011 to 2016 for Vermont Medicare beneficiaries and the comparison group. The rate was higher for Vermont Medicare beneficiaries relative to the comparison group throughout the baseline period and test period.

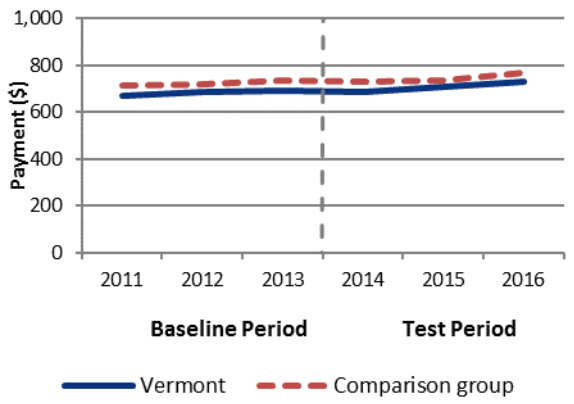
ED = emergency department.

Figure G-55. Average number of 30-day readmissions per 1,000 discharges, FY 2011–FY 2016, Vermont Medicare beneficiaries and the comparison group



The average number of **30-day readmissions** per 1,000 discharges decreased slightly in the baseline period for Vermont Medicare beneficiaries and the comparison group and then remained relatively stable during the test period for both groups.

Figure G-56. Total PMPM expenditures, FY 2011–FY 2016, Vermont Medicare beneficiaries and the comparison group



Total PMPM expenditures steadily increased from 2011 to 2016 for Vermont Medicare beneficiaries and the comparison group. Expenditures was lower for Vermont Medicare beneficiaries relative to the comparison group throughout the baseline period and test period.

PMPM = per member per month.

Regression model

As described above, we use an alternative D-in-D model, which is presented below in ***Equation G.1***. We use the following notation: Y_{ijt} is the outcome for individual i in state j in year t; I_{ij} (= 0,1) is a test indicator equal to 1 if the individual is in a Test state and 0 if the individual is in its comparison group; Time is a linear time trend ranging from 1 to 6, where Time=1 is the first analytic year (2010) and Time=6 is the last analytic year (2016) available. The term that interacts the Test state indicator and time trend (I_{ij}*Time) in ***Equation G.1*** captures differences in trends between a Test state and its comparison group during the baseline period. In ***Equation G.1***, the Q_t terms are equal to 1 in analytic years 4, 5, or 6. The interaction of the test state indicator and each of the Q_t terms ($I_{ij}*Q_t$) measures the difference in the pre-post change

between the Test state and its comparison group above and beyond any differences in baseline trends.

$$Y_{ijt} = \alpha_0 + \beta_1 I_{ij} + \alpha_1 \text{Time} + \beta_2 I_{ij} * \text{Time} + \Sigma\, \alpha_{2,k} Q_k + \Sigma\, \gamma_k I_{ij} * Q_k + \lambda X_{ijt} + \varepsilon_{ijt} \qquad \text{(G.1)}$$

The vector X_{ijt} of individual characteristics includes the following covariates for the commercial population in MarketScan: gender, age and age squared, drug coverage, mental health coverage, relationship to the policyholder (spouse or child), plan type indicator (CDHP), the individual's Hierarchical Condition Categories (HCC) risk score. The models also include county-level controls: population density, percent white, percent of the under 65 population uninsured, percent in poverty, the number of hospital beds per capita, and the median age. The models for the Medicare population include the following covariates: indicators for the urban status of the individual's county of residence, gender, Medicare-Medicaid eligibility, disability status, race (white vs non-white), age and age squared, and HCC risk score (quartiles). The models also include county-level covariates from the Area Resource File, including percent without health insurance, median age, percent of persons in poverty, unemployment rate, and hospital beds per population. The Medicaid models included age, gender, race, Medicaid eligibility category, continuous enrollment, Chronic Illness and Disability Payment System (CDPS) risk score as covariates. We chose to include the available covariates for each payer that could be associated with both the outcomes and residence in a Test State. The last variable ε_{ijt} is a residual term that represents unobserved heterogeneity in the outcome unexplained by X_{ijt} or being in a Round 1 Test state.

The coefficient β_1 in ***Equation G.1*** is the expected difference in the outcome measure between individuals in the Test state and the comparison group at the start of the baseline period, holding constant other variables in the equation. For individuals in the comparison group, the baseline time trend is captured by α_1; for individuals in the Test state, it is $(\alpha_1 + \beta_2)$. Thus, if β_2 is not zero, then the baseline trends are different between the Test state and the comparison group. Similarly, for individuals in the comparison group, the test-period time trend is captured by $(\alpha_1 + \alpha_{2,k})$; for individuals in the Test state, it is $(\alpha_1 + \beta_2) + (\alpha_{2,k} + \gamma_k)$. Furthermore, $(\beta_2 + \gamma_k)$ capture differences in the test-period time trends across the Test state and the comparison group. Therefore, if we net out differences in baseline trends from the difference in test-period trends, the γ_k parameter shows whether the expected outcome increased ($\gamma_k > 0$) or decreased ($\gamma_k < 0$) significantly more than could be explained by baseline differences in trends across states after the SIM Initiative was implemented.

Clustering

The data sources for the analyses contain repeated observations for individuals. Consequently, observations will be clustered at the individual level. Clustering effects are present if, even after controlling for observed characteristics, the outcomes over time for a given individual are correlated. To account for the loss of information in the sample that occurs due to

clustering, we inflate the standard errors at the individual level. This adjustment reduces the probability of a type I error for hypothesis testing—that is, the probability of a statistically significant but spurious effect estimate—but at the same time reduces the power of the test (i.e., the ability to detect a non-zero effect).

Estimation

For the utilization outcomes, we present estimates for the regression adjusted difference in probability of any service use (i.e., γ in ***Equation G.1***). The ATET is interpreted as a change in the probability of a person having any service use (e.g., an inpatient admission) in the Test state relative to its comparison group during the given year, holding all else constant. The adjusted difference is the average change in the probability of any service use in the test years relative to the baseline years for the Test state relative to its comparison group. The adjusted difference was multiplied by 1,000 to scale the result for interpretation of a pre-post change in the rate of any service use per 1,000 members.

For expenditure outcomes, we present the pre-post change in payments for a Test state relative to its comparison group. We again present the coefficient of the interaction of POST and the Test state dummy variable. This coefficient is interpreted as the difference in the change in the dollar amount from the baseline period to the test period in the Test state relative to the comparison group, holding all else constant.

G.3 Population Health Analysis

In addition to claims-based analyses, we examined trends in population health in Round 1 Test states using the Behavioral Risk Factor Surveillance System (BRFSS) survey. The BRFSS is a state-based telephone survey conducted by state health departments and guided by the Centers for Disease Control and Prevention (Centers for Disease Control and Prevention, 2013). The survey is used to collect data from U.S. residents 18 and older on health insurance coverage, health risk behaviors, health status, and preventive health practices. Each year, the survey is made up of a core set of questions that are asked either every year or every other year, optional question modules, and state-added questions (CDC, 2013). Core questions are asked by all states each year. Each state can also choose to add optional modules and approved state-added questions to its survey for a given year. The analyses reported here rely on core questions related to health status, health care access and receipt of preventive services over the 2013 to 2016 period. All of the questions that are examined are asked in each survey year, with the exception of the dental care question, which is only asked in even years (i.e., 2014 and 2016).

The sample frame for the BRFSS relies on a random-digit-dial survey of landline and cellular telephone numbers. Landline telephone interviews are conducted with one randomly selected adult per household (CDC, 2013). Cell-phone interviews are conducted with an adult who answers the phone, who is then treated as a one-person household for weighting purposes, regardless of actual household size. Since the weighting strategy varies across states and over

time, we have imposed a consistent reweighting of the BRFSS samples across states and over time based on individual and family characteristics drawn from the American Community Survey, as described in Appendix B of the State Innovation Models: Model Test Awards Round One, Third Annual Report.

As part of the analysis we examine trends over time for low-income adults, defined as adults with family income at or below 138 percent of the federal poverty level to match the ACA's target population. Since the BRFSS provides a measure of household income rather than family income, we impute family income in the BRFSS using a regression-based imputation strategy that estimated a model of the statistical relationship between family income and household income using the American Community Survey, using the parameter estimates from that model to predict family income for respondents in the BRFSS.[43] We then estimated linear regression models of each outcome, controlling for individual and family characteristics (sex, age, race and ethnicity, educational attainment, marital status, family and household size, employment status, family income, and home ownership), and tested for differences over time relative to the base year (2013).

G.4 References

CDC (Centers for Disease Control and Prevention) (2013). *The BRFSS data user guide.* Atlanta. Available at http://www.cdc.gov/brfss/data_documentation/PDF/UserguideJune2013.pdf

Dowd, B. E., Greene, W. H., & Norton, E. C. (2014). Computation of standard errors. *Health Services Research*, *49*(2), 731–750. See the accompanying appendix as well. http://dx.doi.org/10.1111/1475-6773.12122

Puhani, P. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters*, *115*(1), 85–87. http://dx.doi.org/10.1016/j.econlet.2011.11.025

Puhani, P. A. (2008, April*). The treatment effect, the cross difference, and the interaction term in non-linear 'difference-in-differences' models*. Discussion Paper Series: Forschungsinstitut zur Zukunft der Arbeit, Institute for the Study of Labor. Available at: http://ftp.iza.org/dp3478.pdf.

[43] This imputation approach, which relies on an external data source to impute values for a missing variable in the BRFSS, is most common in microsimulation models, where there is often a need to supplement existing data sources with additional measures to support policy analyses. For example, the Congressional Budget Office uses similar regression-based imputation strategy that relies on the Survey of Income and Program Participation, the Health and Retirement Survey and the Current Population Survey to impute missing variables in the primary database used in its microsimulation model (Schwabish and Topoleski, 2013).

Schwabish, J. A., & Topoleski, J. H. (2013, June). *Modeling individual earnings in CBO's long-term microsimulation model*. CBO Working Paper Series. Working Paper 2013-04. Washington, DC: CBO.

[this page intentionally left blank]