Insurance Coverage for Pregnant Women: Assessing Patterns, Policy Impacts and Methods for Evaluation

Insurance Coverage for Pregnant Women: Assessing Patterns, Policy Impacts and Methods for Evaluation

A dissertation presented

by

Jamie Roberta Daw

to

The Committee on Higher Degrees in Health Policy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Health Policy

Harvard University

Cambridge, Massachusetts

May 2018

Dissertation Advisor: Katherine Swartz Jamie Roberta Daw

Insurance Coverage for Pregnant Women: Assessing Patterns, Policy Impacts and Methods for Evaluation

Abstract

This dissertation includes two applied studies on health insurance coverage for pregnant women in the United States and one methodological contribution to the field of health services research.

Chapter one describes patterns of insurance coverage for pregnant women using national longitudinal survey data from 2005 to 2013. First, we estimate rates of insurance coverage, insurance transitions and insurance lapses in the twelve calendar months before and the six calendar months after childbirth. We find that the period surrounding childbirth is characterized by frequent gaps and changes in coverage that may compromise timely access to recommended care and the continuity and quality of care. Second, we use logistic regression models to identify risk factors associated with insurance lapses before and after pregnancy. We find that risk factors associated with insurance loss after delivery include not speaking English at home, having Medicaid coverage at delivery, living in the South and having a family income of 100 to 185 percent of the poverty level. The findings of this study emphasize the need to develop policies to improve continuity of insurance coverage for women of reproductive-age, particularly low-income women who are eligible for pregnancy-related Medicaid.

Chapter two estimates the association between the Affordable Care Act's dependent coverage provision, which allowed young adults to enroll in their parent's plan until age 26, and payment for birth, prenatal care use, and infant birth outcomes among unmarried and married

women. Drawing on birth certificate data for nearly 3 million births, we use difference-in-differences analysis to compare outcomes among eligible women (ages 24-25) to a control group of slightly older women (ages 27-28) before and after the implementation of the provision. We find that the policy was associated with a decline in Medicaid payment and a 20% increase in private payment for delivery among unmarried women. We also find an association between the policy and modest improvements in early prenatal care, cesarean delivery and preterm birth among unmarried women. We do not find an association between the policy and payment for birth, adequate prenatal care or birth outcomes among married women, nor do we any find changes in low birthweight or NICU admission. The findings of this study suggest that the ACA's dependent coverage provision shifted unmarried pregnant women from Medicaid to their parent's private plans and that this shift was associated with neutral to small positive changes in prenatal care and birth outcomes.

Chapter three uses a Monte Carlo simulation to estimate the bias that can be introduced by applying matching to difference-in-differences. Our results show that matching can have an important impact on estimated intervention effects, particularly when matching on pre-period levels of the outcome itself or on time-varying covariates with low serial correlation. We find that the bias introduced by regression-to-the mean increases with pre-period differences between the treatment and control group and with decreasing serial correlation in the matching covariates. The findings of this study suggest that researchers should exercise caution when matching on pre-period variables in study designs that estimate effects based on changes over time. Based on our results, we provide guidance for selecting matching variables in difference-in-differences analysis.

Table of Contents

Acknowledgements

It is with great appreciation that I acknowledge the guidance, support, and encouragement that enabled this dissertation. First and foremost, I am grateful to my dissertation committee: Kathy Swartz, Laura Hatfield, and Ben Sommers. I have been truly lucky to have mentors that are equally kind as they are brilliant.

Kathy: thank you for being a bedrock of support during my dissertation years. Your expertise has bettered my research and your wise and compassionate voice has helped me to navigate both personal and professional obstacles.

Laura: thank you for believing in my ideas from day one and for persistently pushing me outside of my comfort zone. Your remarkable ability to cut to the heart of problems and effectively communicate complex concepts has had a profound influence on my thinking.

Ben: thank you for setting high expectations, being so generous with your time, and making sure I never lost momentum in my work. Your commitment to making a real-world impact for vulnerable populations has inspired the research in this dissertation and will continue to have an influence on my future career.

Gratitude is also due for Debbie Whitney, as well as Jessica Livingston and Colleen Yout, for always having an open door and putting students first. You are the everyday heroes for so many health policy students and I am no exception.

I also want to acknowledge the encouragement and friendship of my fellow PhD students. When I moved to Boston, I never expected to find such a supportive and caring community. To all of you who have shared celebrations, tears, and beers with me over the last five years, thank you. I sincerely hope that we will continue to be in each other's personal and professional lives for years to come.

None of this would be possible without the love and unconditional support of my parents, Ruth Ann and James. Thank you for instilling in me the belief that I could pursue my dreams, and the work ethic to make them a reality. Knowing that you will always stand by me has given me the confidence to take big leaps, even when I've been unsure of where I might land.

And finally, I am grateful to Ben Oseroff, who has been a daily source of joy throughout the past five years. Thank you for making the ordinary anything but.

List of Tables

List of Figures

CHAPTER 1: Women In The United States Experience High Rates Of Coverage 'Churn' In Months Before And After Childbirth

Authors and affiliations
Jamie R. Daw, M.Sc.; Harvard PhD Program in Health Policy, Cambridge, MA

Laura A. Hatfield, Ph.D.; Department of Health Care Policy, Boston, MA

Katherine Swartz, Ph.D.; Department of Health Policy & Management, Harvard T.H. Chan School of Public Health, Boston, MA

Benjamin D. Sommers, M.D., Ph.D.; Department of Health Policy & Management, Harvard T.H. Chan School of Public Health, Boston, MA

Jamie R. Daw, Laura A. Hatfield, Katherine Swartz and Benjamin D. Sommers. "Women In The United States Experience High Rates Of Coverage 'Churn' In Months Before And After Childbirth." *Health Affairs (*Millwood) 2017, Vol. 36, No.4, 598-606, doi: 10.1377/hlthaff.2016.1241. The published article is archive and available online at www.healthaffairs.org

1.1 Abstract

Insurance transitions before and after childbirth – sometimes referred to as “churning” - can adversely affect continuity and quality of care. Yet, little is known about coverage patterns or changes for women giving birth in the U.S. Using nationally representative survey data from 2005 to 2013, we found high rates of insurance transitions before and after delivery. Half of women who were uninsured nine months before delivery acquired Medicaid coverage by delivery, but 55% of women with Medicaid at delivery experienced a coverage gap in the ensuing six months. Risk factors associated with insurance loss after delivery include not speaking English at home, being unmarried, having Medicaid at delivery, living in the South, and having an income between 100-185% of the poverty level. To minimize the adverse effects of coverage disruptions, states should consider policies that promote continuity of coverage for childbearing women, particularly those with pregnancy-related Medicaid eligibility.

1.2 Introduction

Health insurance facilitates women's access to timely prenatal and postpartum care, which improves birth outcomes and supports the long-term health of women and newborns.[1] However, women are vulnerable to insurance disruptions before and after childbirth because of changes in employment, income, and program eligibility that commonly accompany childbirth.

Under federal law, states must provide Medicaid coverage for pregnancy-related medical services for women with incomes under 138% of the federal poverty level (FPL). States can extend coverage to higher-income pregnant women via Medicaid or the State Children's Health Insurance Program (CHIP) (as of 2016, state income eligibility ceilings for pregnancy-related Medicaid/CHIP varied from 138% to 380% FPL).[2]

Coverage options are limited for women who become pregnant while uninsured or lose private insurance during pregnancy but do not qualify for Medicaid. Under the Affordable Care Act (ACA), all non-group and group insurance plans must include maternity benefits. However, pregnancy is not a qualifying event for special enrollment in a qualified health plan (QHP) on the ACA's insurance marketplaces - though delivering a child is.

Even when women gain Medicaid as a result of pregnancy, coverage is time-limited: pregnant women are only eligible from conception to sixty days postpartum, after which they must either re-qualify for Medicaid as parents/adults, obtain private insurance, or become uninsured. In all states, there is a gap between Medicaid eligibility thresholds for pregnant women and those for parents/adults (as of 2016, state income eligibility ceilings for parents vary from 18% to 221%

FPL).[3] In the District of Columbia and 31 states that expanded Medicaid under the ACA, women with incomes below 138% retain Medicaid eligibility after childbirth. In non-expansion states, many will have incomes that fall in the gap between Medicaid eligibility for adults/parents and eligibility for marketplace subsidies (>100% FPL). The median Medicaid eligibility cutoff for parents in these 19 states is just 44% FPL and only three non-expansion states (ME, TN, and WI) cover parents at or above poverty.[3] Even women who continue to be Medicaid eligible after delivery may face administrative or knowledge-based barriers to postpartum Medicaid enrollment.[4]

While no study has examined the effect of insurance transitions before and after childbirth on health outcomes, research on other populations has shown that coverage disruptions – sometimes called health insurance "churning" - increase the odds of delaying needed care, reduce use of preventive services such as mammography and pap tests, and worsen the perceived quality of care.[5–7] Further, changing plans, even without a gap in coverage, has been associated with a 65% increase in the likelihood of delaying care because of cost and a 37% decrease in the odds of having a usual source of care.[8]

Published reports about insurance coverage of childbearing women in the U.S. are limited to specific states and use cross-sectional data, measuring payment for delivery or coverage at one point in pregnancy, rather than tracking coverage longitudinally across the period before and after childbirth.[9–12] To address this gap, we estimated monthly rates of insurance coverage and insurance transitions for women before and after childbirth using nationally representative longitudinal survey data from 2005 to 2013. We also identified risk factors for coverage lapses.

Our results serve as a valuable baseline from which to evaluate the potential effects of the ACA, as well as any potential major Medicaid reforms under the new administration, on childbearing women. Understanding typical patterns of insurance related to pregnancy is critical for designing policies that promote continuity of coverage across Medicaid eligibility categories and between Medicaid and private insurance.

1.3 Methods

1.3.1 Data

We used pooled panels of the Medical Expenditure Panel Survey – Household Component (MEPS-HC) covering 2005 to 2013. Each MEPS-HC panel is a nationally representative sample of civilian, noninstitutionalized households; the survey collects detailed information on respondents' demographics, health care utilization, health expenditures, and monthly insurance status over a two-year period.

1.3.2 Sample

We identified 2,948 births to 2,843 unique women from inpatient hospital files using ICD-9-CM and Clinical Classification codes for labor/delivery (for births 2005-2007) and the delivery indicator variable reported in the MEPS (available for births 2008-2013).[13] For the 210 women with two deliveries, we randomly selected one delivery to maintain independence of the observations. We excluded deliveries where an infant could not be found in the household or where the infant month of birth and delivery record month differed by more than one month. The final sample included 2,726 women (95.9%). This approach excludes out-of-hospital births, which represented just 2% of U.S. births in 2013.[14]

1.3.3 Insurance Variables

Relative to the month of delivery, we classified monthly insurance status for each calendar month starting three months prior to conception through six months postpartum.

Insurance status was coded as one of three mutually exclusive categories: Medicaid/CHIP, Private/Other Insurance, or Uninsured. Medicaid/CHIP included women who reported *any* coverage by Medicaid, CHIP, or "any other public insurance or state program;" of note, 0.4% of our sample reported Medicare coverage, but always in combination with Medicaid, so they were included in this category as well. Private/Other insurance included those who only reported employer-sponsored coverage, TRICARE/CHAMPVA coverage, or "other private insurance – source unknown." In other words, we applied a hierarchy where women with any Medicaid/CHIP coverage were classified as Medicaid/CHIP regardless of other coverage sources. We defined a lapse in coverage as any month in the Uninsured category.

1.3.4 Analysis

We used logistic regression to estimate the likelihood of an insurance lapse during two periods: nine months before delivery and six months after delivery. Covariates were age, education, race, Hispanic ethnicity, English language spoken at home, marital status, family income, Census region, and insurance type at delivery. We adjusted for the wave of the survey to control for changes over time. Demographics were those reported in the calendar year of the delivery. All estimates used MEPS-HC longitudinal sampling weights.

1.3.5 Limitations

This study has several limitations. First, we identified deliveries using inpatient records provided by hospitals as part of the MEPS. Procedure codes are commonly used to identify obstetric events from hospital discharge data.[15] We suspect that a high proportion of the events we identified were true deliveries, particularly since we only included those where a newborn was subsequently identified in the household. Though the MEPS asks extensively about household composition, there is a small possibility that we did not identify all deliveries to women in the sample due to underreporting by women or inaccurate reporting by hospitals. Second, without information on gestational age, we could not assign a pregnancy status to each calendar month. In our model of prenatal insurance lapses, we analyzed coverage during the nine calendar months up to and including the delivery month. Of course, not all pregnancies last nine months, but the bias from this assumption is likely minimal: among all singleton births in the U.S. in 2013, 89% had a gestational age of 37 weeks or longer and only 3% were less than 34 weeks.[14] Finally, similar to other national surveys, uninsurance tends to be over-reported in the MEPS.[16] A validation study found that 10% of MEPS respondents who report being uninsured hold private insurance.[17] A synthesis of validation studies of insurance reporting in large surveys found 74-88% of Medicaid beneficiaries correctly report Medicaid coverage and approximately 90% report being insured.[18]

1.4 Results

1.4.1 Overall Trends

We calculated the proportion of women insured or not by calendar month leading up to and following delivery (Figure 1.1). The majority of our sample had private/other insurance, and the overall proportion of women with private insurance was relatively stable before, during, and after pregnancy. Leading up to delivery, the proportion of women who were uninsured decreased

and the proportion with Medicaid/CHIP increased. The lowest rate of uninsurance was in the month of delivery (13% of women). However, the uninsurance rate rose rapidly after delivery, nearly returning to the uninsurance rates observed in the months before pregnancy (25% in the tenth month before delivery compared to 23% six months after delivery). These overall patterns did not change substantively after the implementation of the dependent coverage provision of the ACA in 2010 (Figure A1.1).[19]

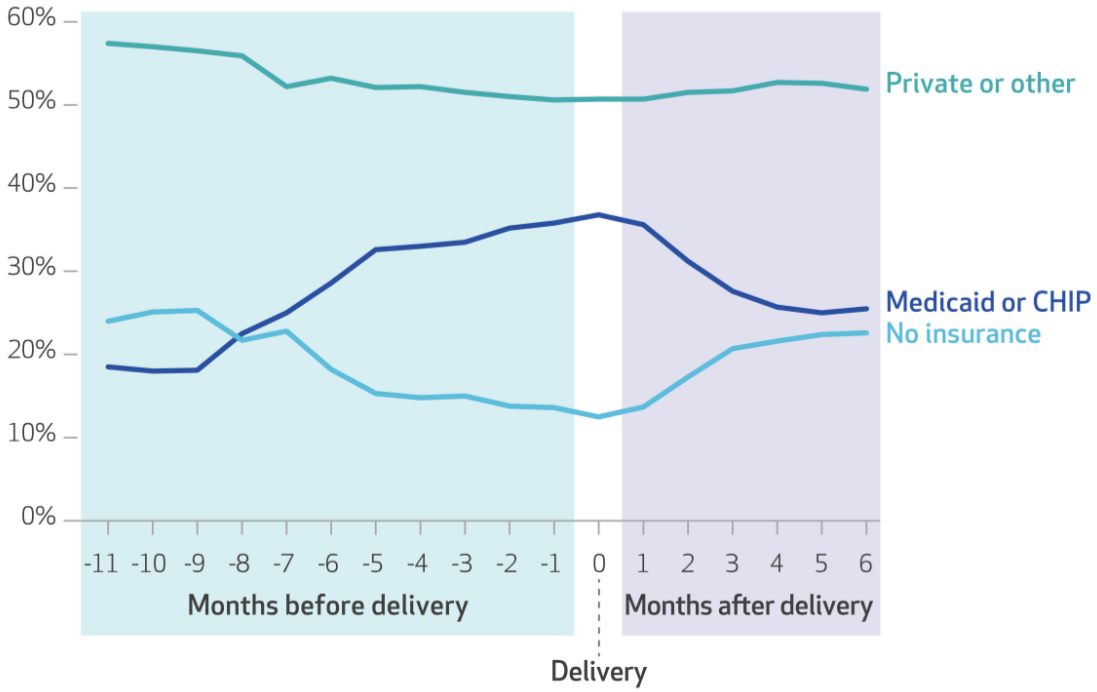


Figure 1.1 Percentages of women who gave birth in the period 2005-13, by health insurance type and month before or after delivery

Notes: Authors' analysis of data for 2005-13 from panels 10-17 of the Medical Expenditure Panel Survey-Household Component. For each month, only women in the survey sample for that month are included (n = 2,727). CHIP is Children's Health Insurance Program.

1.4.2 Insurance Changes Before Childbirth

Overall, in the nine months leading up to and including the delivery month, 58% of women experienced at least one change in insurance status and 62% were uninsured at least one month (Table 1.1). Rates of coverage changes (49%) and uninsurance (49%) before delivery were

lowest among women with private insurance in the month of delivery. In contrast, nearly three-quarters of women with Medicaid in the month of delivery experienced an insurance coverage change during the nine months before delivery, and 65% were uninsured at least one month. Switching from uninsurance to Medicaid was the most common change: 51% of women who were uninsured before pregnancy acquired Medicaid by delivery. See Table A1.1 and Table A1.2 for more detailed statistics on month-to-month changes.[19] Family income was the only measured risk factor significantly associated with an insurance lapse in the nine months up to and including the delivery month. Higher-income women were much less likely to become uninsured during pregnancy (Table A1.3).[19]

Table 1.1 Percentages of women with insurance changes and lapses before and after delivery, by type of insurance status in delivery month

Type of Insurance in Delivery Month	Nine Months Up To and Including Delivery[a]		Six Months After Delivery[b]	
	Any Month-to-Month Change	Any Month Uninsured	Any Month-to-Month Change	Any Month Uninsured
Private or other	49%	49%	36%	35%
Medicaid or CHIP	73	65	59	55
No insurance	52	100	43	70
All types	**58**	**62**	**45**	**47**

Notes: Authors' analysis of data for 2005–13 from panels 10–17 of the Medical Expenditure Panel Survey–Household Component. Only women in the survey sample for all months in each period are included. CHIP is Children's Health Insurance Program. [a]n = 1,751. [b]n = 2,036.

1.4.3 Insurance Changes After Childbirth

Sixty-five percent of women with private insurance in the month of delivery continuously held insurance coverage for six months after delivery (Figure 1.2). In contrast, only 45% of those covered by Medicaid at time of delivery had continuous insurance six months later, and they were more likely to experience month-to-month changes in insurance status. The number of uninsured months following delivery was also much higher among women on Medicaid at

delivery, with nearly 55% experiencing at least one uninsured month and 25% experiencing two or more uninsured months over the next half year (Figure 1.2). For women who were uninsured at delivery, the majority (57%) remained uninsured for the entire six months postpartum.

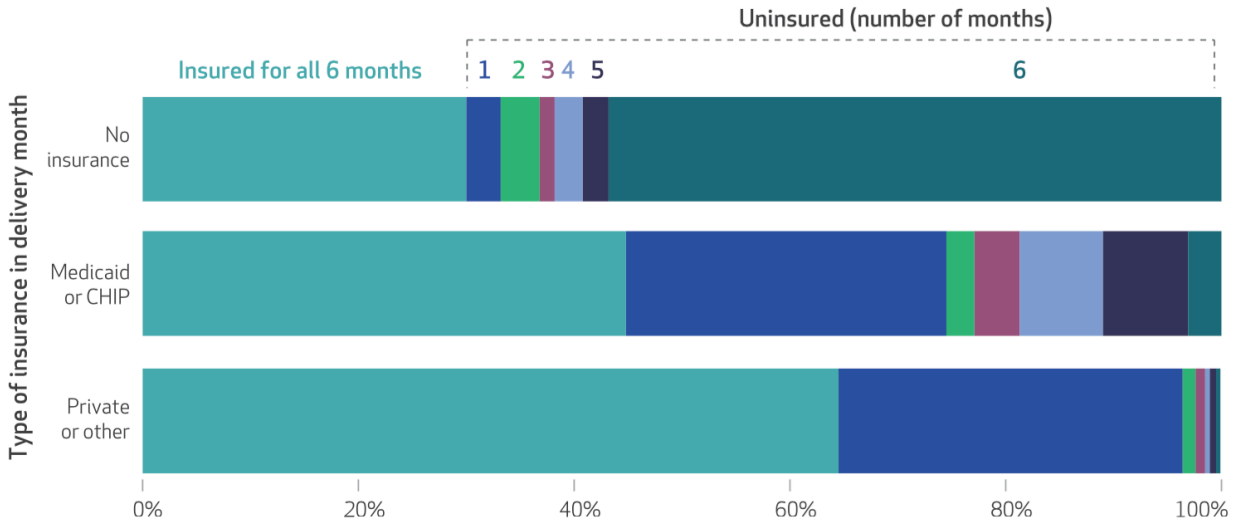


Figure 1.2 Percentages of women who were insured or not insured for any of the six months following delivery, by type of insurance in the month of delivery

Notes: Authors' analysis of data for 2005-13 from panels 10-17 of the Medical Expenditure Panel Survey-Household Component. For each month, only women in the survey sample for the six months after delivery are included (n = 2,036). CHIP is Children's Health Insurance Program.

Risk factors significantly associated with an insurance lapse include income between 100 and 185% FPL (compared to those with incomes <50% of FPL), not speaking English at home, being unmarried, having Medicaid coverage at delivery (compared to Private/Other coverage), and living in the South (compared to the Northeast) (Figure 1.3). Adjusted odds ratios are shown in Table A1.3.[19] Holding other maternal characteristics constant, the predicted probability of an insurance lapse after delivery was 18% lower if women had private insurance in the month of delivery compared to Medicaid (Figure 1.3).

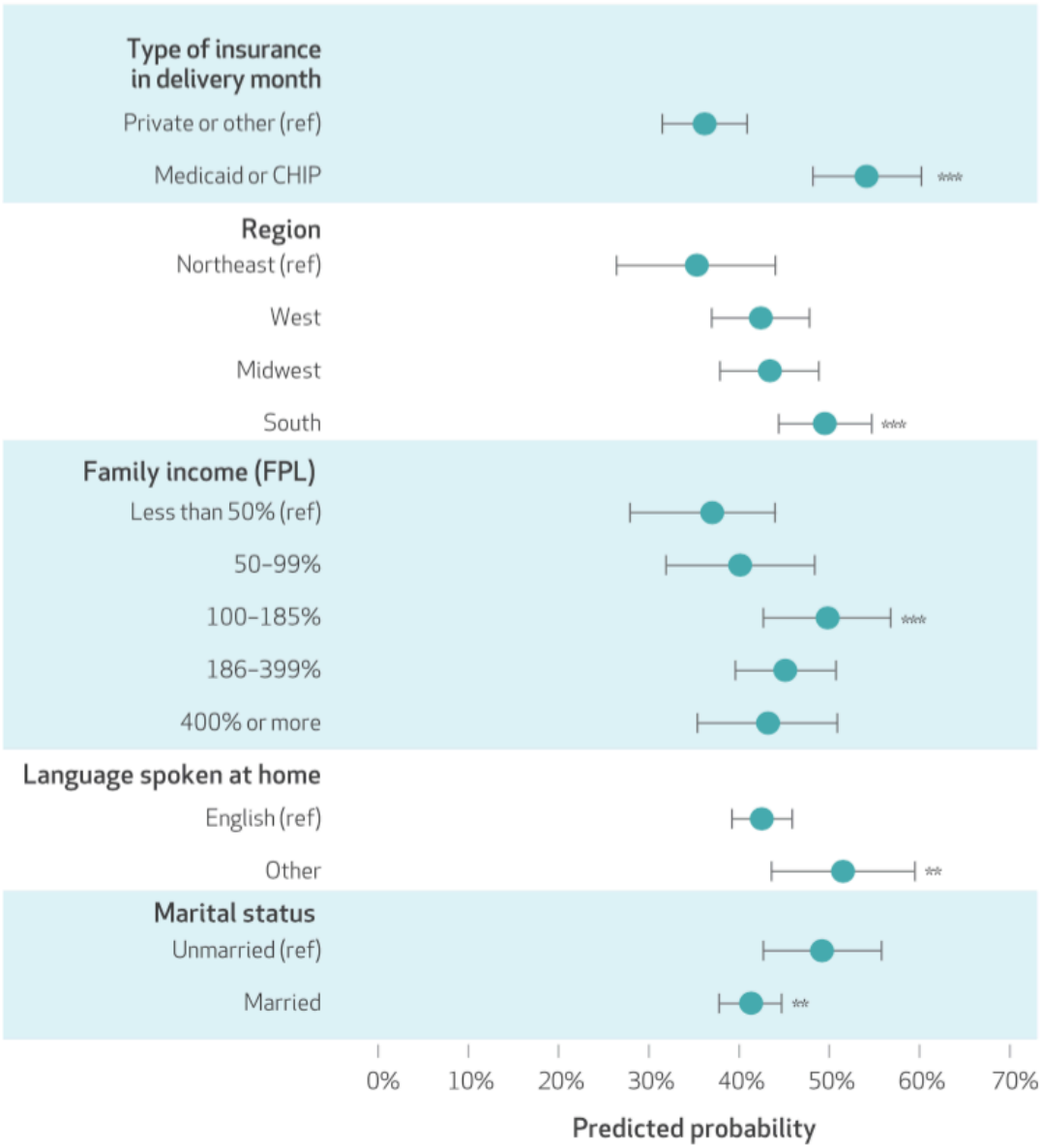


Figure 1.3 Predicted probability of having any insurance lapse in the six months following delivery for women with insurance at delivery

Notes: Authors' analysis of data for 2005–13 from panels 10–17 of the Medical Expenditure Panel Survey–Household Component. Only women in the survey sample for the six months after delivery who reported having insurance in the month of delivery are included (n =1,678). The predicted probabilities are marginal effects averaged across the observed covariates in the sample, based on logistic regression that used any insurance lapse as the outcome and adjusted for age, education, race, Hispanic ethnicity, language spoken at home, marital status, family income, region, insurance type at delivery, and survey panel. Predicted probabilities are shown for maternal characteristics where the adjusted odds ratio from the logistic model for one or more levels was significant relative to the reference category. Error bars show 95 percent confidence intervals. CHIP is Children's Health Insurance Program. FPL is federal poverty level. Statistical significance is of the test of the difference in the predicted probability relative to the reference category. **p < 0.05 ***p < 0.01

1.5 Discussion

Ensuring that pregnant women have insurance coverage has been a goal of policy efforts for decades, including the federal expansion of Medicaid to cover pregnant women in the 1980s, the CHIP Reauthorization Act of 2009, and the ACA. Using nationally representative longitudinal data, this study makes two important contributions to the scarce literature on insurance dynamics for childbearing women. It delivers a granular view of both insurance churning and disruptions that occur on a month-to-month basis. We find that higher rates of coverage for the costs of labor and delivery mask considerable dynamics, with high transition rates in the prenatal and postpartum months, especially for women with Medicaid coverage in the month of delivery. Ours is also the first study to track insurance rates in the six months following childbirth. We find that maintaining postpartum coverage is a particular challenge: nearly half of all births are followed by a period of uninsurance in the six months after delivery, affecting an estimated 1.8 million families in 2013. Among women with Medicaid in the month of delivery, 55% became uninsured at some point during the six months after giving birth. This suggests that many women have no other accessible source of coverage when pregnancy-related Medicaid coverage ends 60 days after delivery.

1.5.1 Comparison to Prior Research

Measuring changes in insurance coverage before and after childbirth is challenging because few surveys allow researchers to identify deliveries or pregnant women and monthly changes in coverage. The only other estimates of insurance transitions before childbirth use data from the Pregnancy Monitoring Assessment and Reporting System (PRAMS).[9,10] An analysis of the 2009 PRAMS, which sampled women in 29 states, found that 30% of women had a different source of

payment for their delivery than their coverage source in the month before pregnancy.[10] Women with a change in coverage source were more likely to be low-income (≤200% FPL), Hispanic, less educated, unmarried, and have an unintended pregnancy.[10]

Other cross-sectional analyses have only estimated uninsurance among pregnant women and payment sources for delivery. A study using the National Health Interview Survey from 2000 to 2009 found that 10% of pregnant women reported being uninsured.[12] The 2009 PRAMS analysis found that only 1.5% of women report self-paying for delivery; however, this estimate is likely an underestimate. The PRAMS sample excludes many states in the South as well as 8.9% of women who did not respond or responded "other" to insurance questions.[10] Widely-cited analyses based on National Vital Statistics birth certificate data from 33 states in 2010 show that 44.9% of deliveries were paid by Medicaid, 50.8% by private/other insurance, and 4.4% were to uninsured women/paid out-of-pocket.[11]

There are several possible reasons why our finding that 12.5% of women reported uninsurance in the month of delivery is higher than other studies. First, some women who receive Medicaid payment for delivery expenses may reasonably consider themselves uninsured in the calendar month of delivery. Uninsured Medicaid-eligible women may enroll in pregnancy-related Medicaid coverage during the delivery episode, and some women, including undocumented immigrants, may receive emergency Medicaid that only covers their delivery expenses; this is consistent with the low-income and disproportionately Hispanic population described in the PRAMS above. Births to undocumented parents represent 7% of all births in the U.S. and a significant proportion of Medicaid-funded births; for example, in North Carolina, 16% of

Medicaid-paid births are funded through emergency Medicaid.[20,21] Second, unlike PRAMS and the Vital Statistics birth certificate data, the MEPS samples all states, including many in the South that have higher rates of uninsurance and lower Medicaid eligibility thresholds for pregnant women. The set of included states matters: the estimated share of self-paid deliveries ranges widely across states (for example, 0.8% in D.C, 8.2% in Texas, 9.0% in Florida, and 16.4% in Nevada based on Vital Statistics data).[11]

Finally, an advantage of the MEPS is that it enables us to measure insurance dynamics over time. However, as stated in the limitations, uninsurance tends to be over-reported in national surveys including the MEPS.[16] To calculate an upper-bound estimate, if we assumed 20% of those who report being uninsured actually had private insurance and 20% had Medicaid, we would overestimate the level of uninsurance by five percentage points (i.e., we would estimate 7.5% uninsurance in the month of delivery rather than 12.5%).[17] We expect reporting errors to be relatively constant over time and thus to minimally affect our estimates of insurance transitions.

1.5.2 Effects of Coverage Changes and Lapses

Lack of insurance and coverage discontinuity have important implications for both women and infants from preconception to postpartum. Before delivery, interventions can be introduced to modify health behaviors and conditions that contribute to adverse pregnancy outcomes, such as smoking or prescription drug use; uninsurance is a barrier to such timely care .[22] Preconception uninsurance may also decrease the effectiveness of coverage gained during pregnancy: women without a regular provider before pregnancy are less likely to receive timely prenatal care.[23]

Insurance gaps during pregnancy place women at elevated risk of delaying prenatal care initiation and receiving fewer than the advised 7-10 prenatal visits.[24] Inadequate prenatal care is associated with adverse birth outcomes including prematurity, stillbirth, and neonatal death.[1] Uninsurance during pregnancy among low-income women may also be detrimental from a cost perspective: studies have found that spending on prenatal care results in savings in future Medicaid costs.[25,26]

After delivery, coverage disruptions may present a barrier to accessing postpartum care, especially the recommended six-week postpartum visit, which only 40% of American women attend.[27] Uninsured women may struggle to manage existing chronic health issues and common pregnancy-related conditions such as pain and urinary incontinence, many of which persist through the year after birth.[28] Postpartum depression, which affects an estimated 13-19% of women, could go unidentified or untreated without proper follow up care – negatively affecting a woman's health and her relationships with family members.[29] Our finding that a quarter of women with Medicaid coverage at delivery experience two or more uninsured months in the six months after delivery is particularly concerning given the evidence that low-income and Medicaid coverage during pregnancy are associated with higher rates of postpartum depressive symptoms.[30] Coverage gaps after delivery may also inhibit access to family planning services, including counselling on adequate spacing between pregnancies, which is associated with improved birth outcomes.[31] Indeed, postpartum loss of Medicaid coverage may partly explain why the odds of a short inter-pregnancy interval are 41% higher among women who have a Medicaid-paid delivery compared to those who do not.[32]

The wellbeing of infants can also be negatively affected by maternal uninsurance after delivery. Poor management of maternal mental health adversely affects a child's cognitive, behavioral, and socio-emotional development.[33,34] Moreover, having an uninsured parent is associated with a lower likelihood of a child having insurance and receiving recommended pediatric care.[35,36]

1.5.3 Policy Implications

We analyzed data collected prior to the implementation of the ACA Medicaid expansions and insurance marketplaces in 2014 and found that one-quarter of women were uninsured nine months before childbirth. Coverage gains under the ACA likely have resulted in a significant decrease in this share. From 2013 to 2015, the uninsurance rate among reproductive-age women decreased from 19.9% to 12.8% nationwide and from 33.9% to 23.2% among women living below poverty.[37] Early evidence also suggests that the dependent coverage provision was associated with a small increase in the number of deliveries paid by private insurance among women 19 to 25 years of age, though this was mostly driven by a decrease in Medicaid-funded rather than self-paid deliveries.[38]

However, women who are uninsured when they become pregnant and who are not eligible for Medicaid will continue to face challenges under the ACA. Pregnancy is not a qualifying event that allows special enrollment in a QHP, raising concerns that these women will be left without a coverage option.[39] Paying out-of-pocket for pregnancy-related services represents a substantial financial burden; the average payment for prenatal care and delivery among commercial beneficiaries was $18,329 for vaginal delivery and $27,866 for a caesarean section in 2010.[40] New York and Vermont, which included pregnancy as a qualifying event for enrollment in a

QHP as of 2016, will serve as examples of the implications of expanding special enrollment to pregnant women.

By covering more women regardless of their pregnancy or parental status, state Medicaid expansions under the ACA are poised to improve continuity of coverage for low-income women. From 2013 to 2015, the percentage of reproductive-age women with Medicaid coverage increased by 33% in expansion states (17.7% to 23.7%) compared to 7% in non-expansion states (14.0% to 15.0%).[37] In the nineteen states that have not expanded Medicaid, continuity of coverage for low-income women will remain a critical concern, and if Medicaid expansion is rolled back under the new administration, these concerns will likely worsen again in those states that have expanded. We found that women in the South were more likely to experience uninsurance after delivery, likely driven in part by low-income thresholds to qualify for Medicaid as a parent in Southern states. Differential Medicaid expansion is likely to further increase state disparities in access to coverage for childbearing women, especially preconception and postpartum.

States should consider reforms to Medicaid/CHIP eligibility criteria that explicitly recognize the transition from pregnancy to parenthood. States could also consider expanding outreach efforts, streamlining enrollment processes, or introducing special navigators to smooth the transition for low-income women from pregnancy-related Medicaid to adult/parental Medicaid or a QHP. Such efforts could be particularly effective for women whose primary language is not English, whom we found to be at greater risk of being uninsured after delivery.

While the future of the ACA under the Trump administration remains unclear, it is likely that Medicaid will continue to be a major source of coverage for pregnant women. Thus, the coverage discontinuities identified in this study will remain a concern, particularly for women who become Medicaid eligible because of their pregnancy and reside in states with limited Medicaid options for parents and low-income adults. If states are granted increased flexibility in designing Medicaid programs, they should consider how policy proposals that affect reproductive-age women will affect coverage continuity, and thus access to care, in the months surrounding childbirth. Indeed, although data are not yet available to evaluate how the ACA impacted coverage transitions before and after childbirth, such evidence is needed to inform federal and state policies. It is likely that the ACA's Medicaid and QHP expansions – while in some cases complicating the potential for insurance transitions – has overall significantly improved coverage rates for women both before and after childbirth. Further evidence in this area could help inform the debate over repeal and any future changes to the law.

1.6 Conclusion

The months surrounding childbirth are characterized by frequent gaps and changes in coverage that may compromise women's timely access to recommended care, and continuity and quality of care. While the ACA coverage expansions improved coverage levels among reproductive-age women, challenges remain for maintaining continuity of coverage in the period surrounding childbirth, especially for women who qualify for pregnancy-related Medicaid coverage. State and federal policymakers should consider policies that help to smooth coverage transitions from preconception to the postpartum period by aligning coverage options for women of reproductive-age, pregnant women, and parents.

1.7 References

1. Partridge, S., Balayla, J., Holcroft, C. A. & Abenhaim, H. A. Inadequate Prenatal Care Utilization and Risks of Infant Mortality and Poor Birth Outcome: A Retrospective Analysis of 28,729,765 U.S. Deliveries Over 8 Years. *Am. J. Perinatol.* **29,** 787–794 (2012).

2. Kaiser Family Foundation. Medicaid and CHIP Income Eligibility Limits for Pregnant Women, January 1, 2016. (2016).

3. Kaiser Family Foundation. Medicaid Income Eligibility Limits for Parents, 2002-2017. (2017). Available at: http://kff.org/medicaid/state-indicator/medicaid-income-eligibility-limits-for-parents/. (Accessed: 14th June 2016)

4. Stuber, J. & Bradley, E. Barriers to Medicaid Enrollment: Who Is at Risk? *Am. J. Public Health* **95,** 292–298 (2005).

5. Federico, S. G., Steiner, J. F., Beaty, B., Crane, L. & Kempe, A. Disruptions in Insurance Coverage: Patterns and Relationship to Health Care Access, Unmet Need, and Utilization Before Enrollment in the State Children's Health Insurance Program. *Pediatrics* **120,** e1009–e1016 (2007).

6. Sudano, J. J. & Baker, D. W. Intermittent Lack of Health Insurance Coverage and Use of Preventive Services. *Am. J. Public Health* **93,** 130–137 (2003).

7. Sommers, B. D., Gourevitch, R., Maylone, B., Blendon, R. J. & Epstein, A. M. Insurance Churning Rates For Low-Income Adults Under Health Reform: Lower Than Expected But Still Harmful For Many. *Health Aff. (Millwood)* **35,** 1816–1824 (2016).

8. Lavarreda, S. A. M., Gatchell, M., Ponce, N., Brown, E. R. & Chia, Y. J. Switching Health Insurance and Its Effects on Access to Physician Services. *Med. Care* **46,** 1055–1063 (2008).

9. Adams, E. K., Gavin, N. I., Handler, A., Manning, W. & Raskind-Hood, C. Transitions In Insurance Coverage From Before Pregnancy Through Delivery In Nine States, 1996–1999. *Health Aff. (Millwood)* **22,** 219–229 (2003).

10. D'Angelo, D. *et al. Patterns of Health Insurance Coverage Around the Time of Pregnancy Among Women with Live-Born Infants — Pregnancy Risk Assessment Monitoring System, 29 States, 2009*. (Centers for Disease Control and Prevention, 2015).

11. Centers for Disease Control and Prevention. Source of Payment for the Delivery: Births in a 33-state and D.C. Reporting Area, 2010. *National Vital Statistics Reports* (2013). Available at: http://www.cdc.gov/nchs/data/nvsr/nvsr62/nvsr62_05.pdf.

12. Kozhimannil, K. B., Abraham, J. M. & Virnig, B. A. National Trends in Health Insurance Coverage of Pregnant and Reproductive-Age Women, 2000 to 2009. *Womens Health Issues* **22,** e135–e141 (2012).

13. To identify births from 2005-2007, we used Clinical Classification codes (195, 196, 218) and ICD-9-CM 3-digit codes (644, 646, 648, 669, V22-V23, V30-V39). This list was developed based on the codes included on records flagged by the MEPS delivery indicator from 2008-2013.

14. Centers for Disease Control and Prevention. Final Birth Data 2013. *National Vital Statistics Reports* (2015). Available at: http://www.cdc.gov/nchs/data/nvsr/nvsr64/nvsr64_01.pdf.

15. Kuklina, E. V. *et al.* An Enhanced Method for Identifying Obstetric Deliveries: Implications for Estimating Maternal Morbidity. *Matern. Child Health J.* **12,** 469–477 (2008).

16. Lewis, K., Ellwood, M. R. & Czajka, J. L. Counting the Uninsured. *Urban Institute* (1998). Available at: http://www.urban.org/research/publication/counting-uninsured.

17. Hill, S. C. The Accuracy of Reported Insurance Status in the MEPS. *Inquiry* **44,** 443–68 (2007).

18. Call, K. T., Davern, M. E., Klerman, J. A. & Lynch, V. Comparing Errors in Medicaid Reporting Across Surveys: Evidence to Date. *Health Serv. Res.* **48,** 652–664 (2013).

19. *See Appendix.*

20. Passel, J. S. & Cohn, D. Number of Babies Born To Unauthorized Immigrants in U.S. Continues to Decline. *Pew Research Center* (2016). Available at: http://www.pewresearch.org/fact-tank/2016/10/26/number-of-babies-born-to-unauthorized-immigrants-in-u-s-continues-to-decline/. (Accessed: 16th August 2017)

21. DuBard, C. A. & Massing, M. W. Trends in Emergency Medicaid Expenditures for Recent and Undocumented Immigrants. *JAMA* **297,** 1085–1092 (2007).

22. Atrash, H. K., Johnson, K., Adams, M., Cordero, J. F. & Howse, J. Preconception Care for Improving Perinatal Outcomes: The Time to Act. *Matern. Child Health J.* **10,** S3-11 (2006).

23. Braveman, P, Marchi, K, Egerter, S, Pearl, M & Neuhaus, J. Barriers to Timely Prenatal Care Among Women With Insurance: The Importance of Prepregnancy Factors. *Obstet. Gynecol.* **95,** 874–880 (2000).

24. Braveman P, Bennett T, Lewis C, Egerter S & Showstack J. Access To Prenatal Care Following Major Medicaid Eligibility Expansions. *JAMA* **269,** 1285–1289 (1993).

25. Schramm, W. F. Weighing Costs And Benefits Of Adequate Prenatal Care for 12,023 Births In Missouri's Medicaid Program, 1988. *Public Health Rep.* **107,** 647–652 (1992).

26. Hueston, W. J., Quattlebaum, R. G. & Benich, J. J. How Much Money Can Early Prenatal Care for Teen Pregnancies Save?: A Cost-Benefit Analysis. *J. Am. Board Fam. Med.* **21,** 184–190 (2008).

27. American College of Obstetricians and Gynecologists. *Committee Opinion: Optimizing Postpartum Care*. (2016).

28. Borders, N. A Critical Review of Postpartum Health Relative to Method of Delivery. *J Midwifery Womens Health* **51,** 242–248 (2006).

29. O'Hara, M. W. & McCabe, J. E. Postpartum Depression: Current Status and Future Directions. *Annu. Rev. Clin. Psychol.* **9,** 379–407 (2013).

30. Lancaster, C. A. *et al.* Risk Factors for Depressive Symptoms During Pregnancy: A Systematic Review. *AJOG* **202,** 5–14 (2010).

31. Conde-Agudelo A, Rosas-Bermúdez A & Kafury-Goeta A. Birth Spacing And Risk Of Adverse Perinatal Outcomes: A Meta-Analysis. *JAMA* **295,** 1809–1823 (2006).

32. Gemmill, A. & Lindberg, L. D. Short Interpregnancy Intervals in the United States. *Obstet. Gynecol.* **122,** 64–71 (2013).

33. Field, T. Postpartum Depression Effects on Early Interactions, Parenting, and Safety Practices: A Review. *Infant Behav. Dev.* **33,** 1–6 (2010).

34. Grace, S. L., Evindar, A. & Stewart, D. E. The Effect of Postpartum Depression on Child Cognitive Development and Behavior: A Review and Critical Analysis of the Literature. *Arch. Women's Ment. Health* **6,** 263–274 (2003).

35. Sommers, B. D. Insuring Children Or Insuring Families: Do Parental And Sibling Coverage Lead To Improved Retention Of Children In Medicaid And CHIP? *J. Health Econ.* **25,** 1154–1169 (2006).

36. DeVoe, J. E., Tillotson, C. J. & Wallace, L. S. Children's Receipt of Health Care Services and Family Health Insurance Patterns. *Ann. Fam. Med.* **7,** 406–413 (2009).

37. Guttmacher Institute. Uninsured Rate Among Women of Reproductive Age Has Fallen More Than One-Third Under the Affordable Care Act. (2016). Available at: https://www.guttmacher.org/article/2016/11/uninsured-rate-among-women-reproductive-age-has-fallen-more-one-third-under. (Accessed: 9th December 2016)

38. Akosa Antwi, Y., Ma, J., Simon, K. & Carroll, A. Dependent Coverage under the ACA and Medicaid Coverage for Childbirth. *N. Engl. J. Med.* **374,** 194–196 (2016).

39. Garro, N., Hernandez, B. & Pellegrini, C. HHS Must Remove Barriers To Coverage For Pregnant Women. *Health Affairs Blog* (2016). Available at: http://healthaffairs.org/blog/2016/02/19/hhs-must-remove-barriers-to-coverage-for-pregnant-women/. (Accessed: 16th December 2016)

40. Truven Health. The Cost of Having a Baby in the United States. (2013). Available at: http://transform.childbirthconnection.org/reports/cost/.

CHAPTER 2: Association of the Affordable Care Act Dependent Coverage Provision With Prenatal Care Use and Birth Outcomes

Authors and affiliations
Jamie R. Daw, M.Sc.; Harvard PhD Program in Health Policy, Cambridge, MA

Benjamin D. Sommers, M.D., Ph.D.; Department of Health Policy & Management, Harvard T.H. Chan School of Public Health, Boston, MA

Jamie R. Daw and Benjamin D. Sommers. Association of the Affordable Care Act Dependent Coverage Provision With Prenatal Care Use and Birth Outcomes. *JAMA* 319, no.6 (2018):579-587 doi:10.1001/jama.2018.0030

2.1 Abstract

Importance: The impact of the Affordable Care Act's (ACA) dependent coverage provision on pregnancy-related health care and health outcomes is unknown.

Objective: To determine whether the dependent coverage provision was associated with changes in payment for birth, prenatal care, and birth outcomes among married and unmarried women.

Design, Setting, and Participants: Retrospective cohort study, analyzed using a differences-in-differences analysis of individual-level birth certificate data comparing live births among U.S. women aged 24-25 and women aged 27-28 before (2009) and after the dependent coverage provision (2011-2013). Results were stratified by marital status.

Main Exposure: The ACA's dependent coverage provision, which allowed young adults to stay on their parent's health insurance until age 26.

Main Outcomes(s) and Measure(s): Primary outcomes were payment source for birth, early prenatal care (first visit in first trimester), and adequate prenatal care (defined by the APNCU index). Secondary outcomes were cesarean delivery, premature birth, low birthweight, and infant NICU admission.

Results: The study population included 1 379 005 births among women aged 24 to 25 years (exposure group; 299 024 in 2009; 1 079 981 in 2011-2013), and 1 551 192 births among women aged 27 to 28 years (control group; 325 564 in 2009; 1 225 628 in 2011-2013). Overall, the

dependent coverage provision was associated with a 1.9 percentage-point increase in private insurance payment for birth (95% CI: 1.6, 2.1), 1.4 percentage-point decrease in Medicaid (95% CI: -1.7, -1.2), and -0.3 percentage-point decrease in self-pay/uninsurance (95% CI: -0.4, -0.1) among 24-25 year-old women compared to 27-28 year-olds. Among unmarried women, the policy was associated with a 3.8 percentage-point increase in private insurance (95% CI: 3.5, 4.2), 0.8 percentage-point increase in early prenatal care (95% CI: 0.4, 1.3), 0.6 percentage-point increase in adequate prenatal care (95% CI: 0.1, 1.0), 0.6 percentage-point decrease in cesarean deliveries (95% CI: -1.1, -0.2), and 0.3 percentage-point decrease in preterm births (95% CI: -0.5, 0.002). Among married women, the policy was associated with a 0.7 percentage-point increase in early prenatal care (95% CI: 0.4, 1.0). The policy was not associated with statistically significant changes in adequate prenatal care, cesarean deliveries or preterm births among married women. There were no significant changes in low birthweight or NICU admission.

Conclusions and Relevance: In this study of nearly 3 million births among women aged 24 to 25 years compared with those aged 27 to 28 years, the Affordable Care Act dependent coverage provision was associated with increased private insurance payment for birth, increased use of prenatal care, and modest reduction in preterm births and cesarean deliveries among unmarried women. The policy was not associated with changes in low birthweight or NICU admission.

2.2 Introduction

Beginning in September 2010, the Affordable Care Act (ACA) required private health insurers to allow young adults to remain on their parent's plan until their 26th birthday. Nearly one-third of U.S. births are to women in the age range affected by the policy (19-25). While studies have found that this dependent coverage provision expanded insurance, increased preventive care, and improved self-reported health among young adults, little attention has been given to the impact of the provision on pregnant women.[1–6] The only study among pregnant women found that the policy was associated with an increase in privately insured births and a decrease in Medicaid-paid births, concentrated among unmarried women.[7]

Insurance changes among reproductive-age and pregnant women associated with the provision could lead to improvements in prenatal care use and birth outcomes (Figure A2.1). Compared to women who are insured prior to pregnancy, uninsured women who become eligible for pregnancy-related Medicaid may face delays accessing prenatal care because of late pregnancy recognition, enrollment barriers, and difficulty finding a clinician. Coverage before pregnancy is associated with earlier initiation of prenatal care, which is associated with reduced adverse outcomes such as prematurity and neonatal death.[8–10] Expansion of coverage to reproductive-age women may also improve outcomes by increasing access to care before conception, in turn potentially improving chronic-disease management, reducing tobacco use, increasing access to contraception, and improving quality and continuity of care.

The objective of this study was to investigate whether the dependent coverage provision's expansion of health insurance was associated with changes in payment for birth, prenatal care, and birth outcomes, among married and unmarried women.

2.3 Methods

2.3.1 Data

We used individual-level data from the Centers for Disease Control and Prevention public-use natality files, which contain a census of U.S. birth certificate data. We limited our sample to states that had adopted the "2003 revised U.S. Standard Certificate of Live Birth" format in each year, since this format contains consistent outcome definitions and information on payment for birth. Birth certificate data is entered by the birth facility based on clinical records and maternal surveys. During our study period, the number of states using the 2003 format increased from 28 in 2009 to 40 (plus Washington D.C.) by 2013, but the public dataset does not identify individual states over time.[11] This study was deemed non-human subjects research by the Harvard University IRB.

2.3.2 Study Design

Similar to previous evaluations of the dependent coverage provision[1–5], we conducted a retrospective cohort study analyzed using difference-in-differences analysis, comparing younger adults eligible for the provision to a control group of slightly older adults. We stratified our results by marital status because marriage has the potential to modify the effect of the policy. Married women are more likely to have private insurance than unmarried women (due partly to spousal insurance), decreasing the probability that they will receive coverage through a parent's plan.[12] This hypothesis is supported by previous research on the dependent coverage provision, including a study on payment for birth, which found larger increases in private insurance among unmarried women.[1,7]

The study period was 2009-2013, since our dataset only includes source of payment for delivery starting in 2009, and in 2014, the ACA's Medicaid and Marketplace expansions took effect, changing coverage options for all ages. While dependent coverage was not mandated until September 23, 2010, some insurers voluntarily implemented it earlier in 2010. Accordingly, we excluded births in 2010 as a "washout period," similar to several prior studies.[3,5,7] We classified 2011-2013 as the post-policy period.

2.3.3. Main Exposure

The primary assumption of difference-in-differences analysis is that the trends observed in the "control" group are a valid counterfactual for what would have occurred in the "exposure" group if not for the policy. Following published best practices for this study design, we inspected pre-policy trends before conducting the assessment of the policy's effects, in order to select our comparison groups.[13] Based on an analysis of trends in our study outcomes before 2010, we defined the "exposure" group as births to women aged 24 to 25 years and the "control" group as births to women aged 27 to 28 years. While many studies of this provision used broader age groups (e.g. 19-25 and 27-34), such wide age-bands may yield less appropriate comparison groups, and in our case violated the parallel trends assumption for the primary payment outcomes; thus, we chose narrower age bands for a more appropriate analytical comparison (Table A2.1).[14] Pre-period monthly trends were more similar within marital strata than for the overall sample (Table A2.2). We detected small but significant differential linear trends prior to 2010 for prenatal care and NICU admission overall and among unmarried women only (Table A2.2). Estimating our primary difference-in-differences regression as if the policy took effect six months earlier in July 2009 (*post hoc* "placebo testing"), we identified a similar pattern, with

significant results for prenatal care and NICU admission overall and among unmarried women (Table A2.3).

2.3.4 Outcomes

Consistent with numerous studies of health insurance expansions, including the dependent coverage provision, we examined outcomes in three domains: insurance status, access to care, and health.[1,3,15–17] This approach is supported by evidence from a variety of settings showing that coverage changes can lead to changes in utilization and health outcomes.[18]

As a proxy for insurance status, we analyzed payment for birth – Medicaid, private insurance, or self-pay. Payment for birth does not necessarily capture coverage status before and during pregnancy; however, most women with private insurance at delivery have it throughout pregnancy, unlike Medicaid, which is often acquired during pregnancy or at delivery.[19] To measure access to care, we examined early prenatal care (first prenatal visit in the first trimester) and adequate prenatal care (defined by the Adequacy of Prenatal Care Utilization [APNCU] index)[20]. The APNCU index defines adequate prenatal care as having a visit in the first trimester and at least 80 percent of the expected number of visits based on gestational age and clinical guidelines. Finally, we selected a set of secondary infant birth outcomes for which previous research suggests an association with insurance status and prenatal care: delivery by caesarean section,[21–24] preterm birth (<37 weeks),[10,17,23,25] low birthweight (<2500g),[17,23,25–28] and infant NICU admission.[29]

2.3.5 Statistical Analysis

We fitted linear probability models for each outcome as a function of exposure group status, post-implementation status (2011-2013), and their interaction. See Appendix A2 for additional details on the model specification. The interaction term is the difference-in-differences estimate of the association between the provision and the outcome, i.e. the differential change before and after the dependent coverage provision among 24-25 year olds compared to 27-28 year olds. We used linear probability models, which are standard in difference-in-differences analyses for their ease of interpretation of the model's key interaction terms.[30] We used robust standard errors, but also conducted a sensitivity analysis with standard errors clustered at the age group-month level, and this did not substantively change our results (Table A2.4). Multivariable models controlled for potential confounders, including month of delivery; maternal marital status, age, race, ethnicity, and education; whether the birth was a woman's first live birth; multiple delivery; and paternal age. Race/ethnicity data was drawn from birth certificate data; states typically use this information to track population-level health and disparities. Since private health insurance rates are correlated with employment rates, which could differentially affect the exposure and control groups, we also adjusted for monthly sex- and age-specific unemployment rates.[31] To assess whether changes in the outcomes were potentially accounted for by changes in payment for birth, we conducted a sensitivity analysis estimating the association between the provision and prenatal care and birth outcomes adjusting for payment for birth (Table A2.5).

Observations with missing or unreported covariates were included in the analysis and coded using an indicator variable for observations in which the covariate was not reported. We did not adjust for multiple comparisons and thus our results should be considered exploratory. A two-

sided P value of <.05 was considered statistically significant. All analyses were performed using STATA (version 13.1; StataCorp).

2.4 Results

2.4.1 Study Population

The final sample included 1,379,005 births to 24-25 year-olds (the exposure group) and 1,551,192 births to 27-28 year-olds (the control group). Prior to the policy, compared to the control group, the exposure group had a younger paternal age and a higher proportion of women who were Hispanic, Black, unmarried or without post-secondary education (Table 2.1).

Table 2.1. Maternal and paternal characteristics before and after the dependent coverage provision

Characteristic	Control Group (27-28 year olds)		Exposure Group (24-25 year olds)		Differential Change Exposure-Control[a]	*P* Value
	Pre-Policy (95% CI)	Post-Policy (95% CI)	Pre-Policy (95% CI)	Post-Policy (95% CI)		
Total sample size, No. Births	325,564	1,225,628	299,024	1,079,981		
Maternal age, mean (years)	27.5 (27.5, 27.5)	27.5 (27.5, 27.5)	24.5 (24.5, 24.5)	24.5 (24.5, 24.5)	0.02 (-0.3, 0.3)	0.89
Married	68.7 (68.6, 68.9)	67.9 (67.8, 67.9)	54.4 (54.2, 54.6)	51.2 (51.1, 51.3)	-2.3 (-2.6, -2.0)	<0.001
Hispanic ethnicity	26.3 (26.1, 26.4)	21.9 (21.9, 22.0)	29.6 (29.4, 29.7)	25.2 (25.2, 25.3)	0.03 (-0.2, 0.3)	0.84
Race						
White	80.2 (80.1, 80.4)	78.8 (78.7, 78.8)	79.6 (79.4, 79.7)	76.5 (76.5, 76.6)	-1.6 (-1.8, -1.4)	<0.001
Black	12.5 (12.4, 12.6)	13.7 (13.6, 13.8)	15.1 (14.9, 15.2)	18.0 (17.9, 18.0)	1.7 (1.5, 1.9)	<0.001
Other	7.3 (7.2, 7.4)	7.5 (7.5, 7.6)	5.3 (5.3, 5.4)	5.5 (5.5, 5.5)	-0.1 (-0.2, 0.04)	0.15
Education						
Less than HS	15.7 (15.5, 15.8)	12.7 (12.6, 12.7)	19.9 (19.8, 20.0)	16.6 (16.6, 16.7)	-0.3 (-0.5, -0.1)	0.01
High School	45.6 (45.4, 45.8)	45.0 (44.9, 45.1)	58.2 (58.0, 58.4)	59.9 (59.8, 60.0)	2.3 (2.1, 2.6)	<0.001
Any Post-Secondary	38.8 (38.6, 38.9)	42.3 (42.2, 42.4)	21.9 (21.8, 22.1)	23.4 (23.3, 23.5)	-2.1 (-2.3, -1.8)	<0.001
First live birth	37.6 (37.4, 37.7)	39.4 (39.3, 39.5)	41.5 (41.3, 41.7)	42.2 (42.1, 42.3)	-1.1 (-1.4, -0.8)	<0.001
Multiple delivery	3.2 (3.2, 3.3)	3.2 (3.2, 3.3)	2.7 (2.7, 2.8)	2.7 (2.6, 2.7)	-0.1 (-0.2, 0.003)	0.06
Paternal age, mean (years)	30.3 (30.2, 30.3)	30.3 (30.3, 30.3)	27.7 (27.7, 27.7)	27.7 (27.7, 27.7)	-0.03 (-0.1, -0.001)	0.04

Notes: Estimates are given as percentage points except where noted. Estimates are based on an unadjusted model with robust standard errors analyzing data before the policy (2009) and after the policy (2011-2013) excluding 2010 as the policy-implementation period. Differential change represents the difference in the change in each characteristic from pre- to policy in the exposure group relative to the control group, for example, the mean maternal age increased by 0.02 more years in the exposure group relative to the control group.

2.4.2 Payment for Birth

Figure 2.1 shows unadjusted trends in payment for birth before and after the policy for all women and by marital status. Table 2.2 shows unadjusted and adjusted difference-in-differences estimates for changes in each outcome. In the adjusted difference-in-differences analyses, the dependent coverage provision was associated with a 1.9 percentage-point increase in private insurance payment for birth (95% CI: 1.6, 2.1; $p<0.001$; 5% increase from a baseline of 37%), a 1.4 percentage-point decrease in Medicaid payment (95% CI: -1.7, -1.2; $p<0.001$; 3% decrease from a baseline of 52%), and a -0.3 percentage-point change in self-payment/uninsurance (95% CI: -0.4, -0.1; $p<0.001$; 6% decrease from a baseline of 5%).

Among unmarried women, the provision was associated with a 3.8 percentage-point increase in private insurance (95% CI: 3.5, 4.2; $p<0.001$; 20% increase from a baseline of 19%), as well as a 3.6 percentage-point decline in Medicaid among unmarried women (95% CI: -4.0, -3.2; $p<0.001$; 5% decrease from a baseline of 71%) (Table 2.3). The provision was not associated with significant changes in source of payment for married women. The differential changes in private payment (-4.2, 95% CI: -4.6, -3.7; $p<0.001$) and Medicaid payment (3.8, 95% CI: 3.3, 4.3; $p<0.001$) between unmarried and married women were statistically significant.

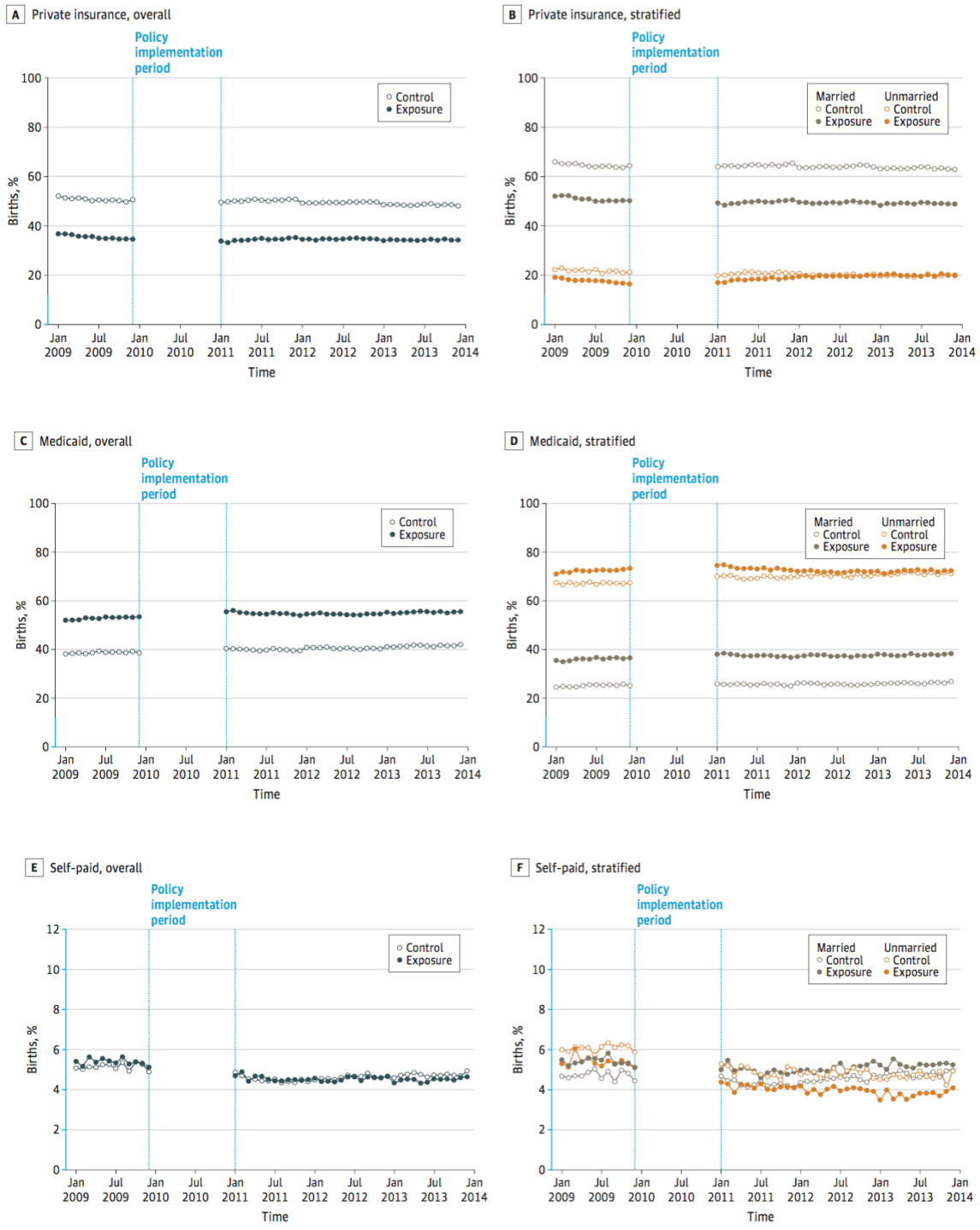


Figure 2.1 Primary source of payment for birth for women aged 24 to 25 years (exposure group) and women aged 27 to 28 years (control group) overall and by marital status

Notes: The blue y-axis scale indicates range from 0% to 12%. The policy implementation period (January 2010-December 2010) was excluded from the analysis. Linear regression modeling with the payment type as the outcome was used to calculate the monthly season adjustments as coefficients on monthly dummy variables. The seasonal adjustments were subtracted from means calculated at the group and month level.

Table 2.1 Estimated changes in the outcomes associated with the ACA dependent coverage provision

Outcome	Pre-Policy Baseline		Unadjusted Difference-in-Differences Estimate (95% CI)	*P* Value	Adjusted Difference-in-Differences Estimate (95% CI)	*P* Value
	Control	Exposure				
Payment for Birth						
Private	52.4	36.9	0.3 (0.01, 0.6)	0.05	1.9 (1.6, 2.1)	<0.001
Medicaid	37.4	51.6	0.1 (-0.2, 0.4)	0.50	-1.4 (-1.7, -1.2)	<0.001
Self-Pay	4.9	5.2	-0.3 (-0.5, -0.2)	<0.001	-0.3 (-0.4, -0.1)	<0.001
Prenatal Care						
Early Prenatal Care	75.7	70.0	0.6 (0.3, 0.8)	<0.001	1.0 (0.7, 1.2)	<0.001
Adequate Prenatal Care	77.5	73.5	0.04 (-0.2, 0.3)	0.74	0.4 (0.2, 0.6)	<0.001
Birth Outcomes						
Cesarean Delivery	32.1	30.1	0.2 (-0.1, 0.4)	0.17	0.005 (-0.3, 0.3)	0.97
Preterm Birth	9.1	9.4	-0.1 (-0.3, 0.04)	0.15	-0.2 (-0.3, -0.03)	0.02
Low Birthweight	7.2	7.5	0.1 (-0.1, 0.2)	0.30	-0.01 (-0.1, 0.1)	0.91
NICU Admission	6.6	6.7	-0.1 (-0.2, 0.1)	0.39	-0.1 (-0.3, 0.3)	0.11

Notes: Estimates are given in percentage-points. All models analyze data from 2009-2013 excluding 2010 as the policy-implementation period and use robust standard errors (N=2,930,197). Adjusted models are adjusted for month of delivery; a monthly linear time trend; maternal marital status, age, race, ethnicity, and education; whether the birth was a woman's first live birth; multiple delivery; paternal age; and monthly unemployment rates.

2.4.3 Prenatal Care

In adjusted analyses, the policy was associated with a 1.0 percentage-point increase in early prenatal care (95% CI: 0.7, 1.2; p<0.001; 1.4% increase from a baseline of 70%), and a 0.4 percentage-point increase in adequate prenatal care (95% CI: 0.2, 0.6; p<0.001; 0.6% increase from a baseline of 74%). After adjusting for payment for birth, the association with early prenatal care was still significant but roughly 30% smaller in magnitude, and the association with adequate prenatal care was no longer significant (Table A2.5).

Significant increases in early prenatal care were identified for both unmarried (0.8 percentage-points, 95% CI: 0.4, 1.3; p<0.001; 1.3% increase from a baseline of 64%) and married women (0.7 percentage-points, 95% CI: 0.4, 1.0; p<0.001; 0.9% increase from a baseline of 75%). Significant increases in adequate prenatal care were only present among unmarried women (0.6 percentage-points, 95% CI: 0.1, 1.0; p=0.01; 0.9% increase from a baseline of 68%); however,

the differential change in adequate prenatal care between unmarried and married women was not statistically significant (-0.5, 95% CI: -1.0, 0.1).

2.4.4 Birth Outcomes

In adjusted analyses, the policy was associated with a statistically significant decrease in preterm birth (-0.2 percentage-points, 95% CI: -0.3, -0.03; p=0.02; 2.2% decrease from a baseline of 9.4%), but not with changes in cesarean delivery, low birthweight, or NICU admission. The association with preterm birth was no longer statistically significant after adjusting for payment for birth (Table A2.5).

In stratified analyses, the policy was associated with significant decreases in preterm birth (-0.3 percentage-points, 95% CI: -0.5, 0.002, p=0.05; 2.9% decrease from a baseline of 10.3% and cesarean section (-0.6 percentage-points, 95% CI: -1.0, -0.2; p=0.01; 1.9% decrease from a baseline of 32%) among unmarried women only. The differential change between unmarried and married women was significant for cesarean sections (0.9, 95% CI: 0.4, 1.4) but not for preterm birth (0.1, 95% CI: -0.2, 0.5).

Table 2.2 Estimated changes in the outcomes associated with the ACA dependent coverage provision, by marital status

	Unmarried Women						Married Women							
	Pre-Policy Baseline		Unadjusted Difference-in-Differences Estimate (95% CI)	*P* Value	Adjusted Difference-in-Differences Estimate	*P* Value	Pre-Policy Baseline		Unadjusted Difference-in-Differences Estimate (95% CI)	*P* Value	Adjusted Difference-in-Differences Estimate (95% CI)	*P* Value	Adjusted Difference Married-Unmarried (95% CI)	*P* Value
Outcome	Control	Exposure					Control	Exposure						
Payment for Birth														
Private	22.6	18.9	3.2 (2.9, 3.6)	<0.001	3.8 (3.5, 4.2)	<0.001	65.9	52.0	-0.9 (-1.3, -0.5)	<0.001	-0.3 (-0.6, 0.01)	0.05	-4.2 (-4.6, -3.7)	<0.001
Medicaid	66.9	71.2	-3.0 (-3.5, -2.6)	<0.001	-3.6 (-4.0, -3.2)	<0.001	24.1	35.1	0.8 (0.4, 1.1)	<0.001	0.2 (-0.1, 0.5)	0.15	3.8 (3.3, 4.3)	<0.001
Self-Pay	5.7	5.1	-0.1 (-0.3, 0.1)	0.19	-0.1 (-0.3, 0.1)	0.17	4.5	5.2	-0.1 (-0.3, 0.02)	0.08	-0.1 (-0.3, 0.01)	0.07	0.0 (-0.3, 0.3)	1.00
Prenatal Care														
Early Prenatal Care	65.3	63.5	0.6 (0.2, 1.1)	0.01	0.8 (0.4, 1.3)	<0.001	80.4	75.4	0.5 (0.2, 0.8)	<0.001	0.7 (0.4, 1.0)	<0.001	-0.1 (-0.7, 0.4)	0.66
Adequate Prenatal Care	69.9	68.4	0.4 (-0.04, 0.8)	0.08	0.6 (0.1, 1.0)	0.01	80.9	77.6	-0.04 (-0.3, 0.3)	0.80	0.1 (-0.2, 0.4)	0.47	-0.5 (-1.0, 0.1)	0.08
Birth Outcomes														
Cesarean Delivery	34.0	31.9	-0.5 (-0.9, -0.1)	0.02	-0.6 (-1.1, -0.2)	<0.01	31.3	28.6	0.2 (-0.1, 0.6)	0.17	0.3 (-0.1, 0.6)	0.12	0.9 (0.4, 1.4)	<0.001
Preterm Birth	10.7	10.3	-0.2 (-0.4, 0.1)	0.24	-0.3 (-0.5,0.002)	0.05	8.4	8.6	-0.2 (-0.4,-0.03)	0.03	-0.1 (-0.3, 0.05)	0.14	0.1 (-0.2, 0.5)	0.46
Low Birthweight	9.1	8.8	0.2 (-0.1, 0.4)	0.19	0.1 (-0.2, 0.3)	0.63	6.3	6.4	-0.1 (-0.3, 0.03)	0.10	-0.1 (-0.2, 0.1)	0.42	-0.1 (-0.4, 0.2)	0.40
NICU Admission	7.6	7.4	-0.2 (-0.4, 0.1)	0.14	-0.2 (-0.5, 0.01)	0.06	6.2	6.1	-0.2 (-0.3, 0.02)	0.08	-0.1 (-0.3, 0.1)	0.28	0.1 (-0.2, 0.4)	0.35

Notes: Estimates are given in percentage-points. All models analyze data from 2009-2013 excluding 2010 as the policy-implementation period and use robust standard errors. Unmarried N =1,159,420. Married N=1,770,777. Adjusted models are adjusted for month of delivery; a monthly linear time trend; maternal marital status, age, race, ethnicity, and education; whether the birth was a woman's first live birth; multiple delivery; paternal age; and monthly unemployment rates. The adjusted difference between married and unmarried women is modeled as the interaction between the post-policy period, exposure group, and marital status.

2.5 Discussion

In this national study of prenatal care and birth outcomes among young adults, the ACA's dependent coverage provision was associated with a significant increase in private insurance payment and a reduction in Medicaid payment for childbirth. The provision was also associated with modest but statistically significant increases in early and adequate prenatal care. In subgroup analyses, changes in payment for birth and adequate prenatal care were only significant among unmarried women. For secondary birth outcomes, the provision was associated with modest decreases in preterm births and cesarean deliveries; changes for these outcomes were only significant among unmarried women.

The study findings for changes in payment for birth are consistent with Antwi et al. (2016), which found a 2.5 percentage-point increase in privately paid births, a 2.1 percentage-point decrease in Medicaid paid births, and a 0.3 percentage-point decrease in self-paid births.[7] The small decrease in uninsured births associated with the policy contrasts with studies of the general young adult population, which estimated decreases of uninsurance between 3.0 and 4.7 percentage points, driven almost entirely by shifts from uninsurance to private insurance.[1,2] This difference reflects the special role Medicaid plays in covering pregnant women. Medicaid paid for 45% of U.S. births in 2010,[34] and even if a woman is uninsured for most of her pregnancy, by the time of delivery, there is a high likelihood she will have obtained Medicaid if she is eligible. Hospitals have a strong incentive to enroll eligible women even after delivery in order to obtain payment (Medicaid pays retroactive costs for eligible individuals for up to 3 months prior to enrollment).[35] Thus, the estimates of payment for birth in this study likely understate the role the

policy played in increasing pre- and post-conception coverage and the number of months women were insured during pregnancy.

Unmarried women experienced a 20% increase in private coverage in contrast to no significant change among married women. This is consistent with Antwi et al. (2016) which also found significant changes in payment for birth among unmarried women only.[1,7] Without the option of spousal coverage, unmarried women were less likely to have private insurance before the ACA and more likely to benefit from the provision. The finding that changes in adequate prenatal care, preterm birth, and cesarean sections were only significant among unmarried women is consistent with the interpretation that the observed changes in prenatal care and birth outcomes are related to changes in payment for birth, and direct adjustment for payment attenuated the observed changes in prenatal care and birth outcomes.

The improvement in prenatal care observed in this study is consistent with previous research showing that women with preconception coverage are better able to access clinicians early in pregnancy.[9] Other research has shown that having private insurance at delivery is associated with higher rates of continuous coverage before, during, and after pregnancy, compared to pregnancy-related Medicaid coverage, which only covers women from conception to sixty days postpartum.[19]

The estimated increases in prenatal care utilization (+1.0 percentage points for early prenatal care and +0.4 percentage points for adequate prenatal care) are modest but not negligible when considered in the context of other interventions aimed to improve prenatal care use. Evidence

from pregnancy-related Medicaid expansions in the 1980s found that despite large increases in coverage, gains in prenatal care use were inconsistent.[25] For example, Epstein and Newhouse (1998) observed an increase of 3 percentage-points in early prenatal care in South Carolina but no improvement in California.[23] Similarly modest evidence exists for community-based interventions to improve prenatal care. For example, evidence on home visiting programs for high-risk women have found results ranging from no significant difference to a 4.5 percentage-point increase in adequate prenatal care use, suggesting that this study's estimates for the dependent coverage provision are in the range of changes detected after more intensive interventions.[36]

Given the relatively small coverage and utilization changes associated with the policy, it is not surprising that changes in birth outcomes were small in magnitude. These modest results are similar to previous studies of coverage expansions to pregnant women, which have found small or no changes in premature birth, cesarean section, and low birthweight, despite improvements in prenatal care.[17,25] While prenatal care is not the only mechanism through which expansion of coverage might affect birth outcomes (Figure A2.1), evidence on the effectiveness of routine prenatal care is mixed.[37,38] Regardless, it is likely that insurance coverage is still important to improve access to such interventions.

The dependent coverage provision is one of many components of the ACA with the potential to affect reproductive-age and pregnant women. Future research should examine the impact of other aspects of the law on insurance coverage during pregnancy and resulting impacts on access to care, maternal outcomes, and both short- and long-term children's health outcomes.

2.5.1 Limitations

This study has several limitations. First, difference-in-differences analysis relies on the assumption that the post-implementation trend for the control group is a valid counterfactual for what would have been observed in the exposure group if not for the policy. Based on an analysis of monthly pre-period trends, this study used narrower age bands than previous studies. Small but significant divergent trends and "placebo effects" were detected for early prenatal care and NICU admission among unmarried women, suggesting that results for these outcomes should be interpreted with caution.

Second, while the use of narrow age bands improves the study's internal validity, this approach only estimates the association between the provision and outcomes for births to 24 and 25 year-old women. The provision may have differentially affected women of younger ages, among whom rates of prenatal care are typically lower and risk of adverse birth outcomes higher.[10]

Third, without state identifiers, the analysis could not be restricted to states based on the year they adopted the 2003 birth certificate format. A previous analysis of the dependent coverage provision and payment for childbirth compared the natality data used in this study to the Nationwide Inpatient Sample and found similar results, suggesting this threat to validity is limited for payment-related outcomes.[7] Though regression models controlled for observed covariates, unobserved changes in the composition of the exposure and control group over time could bias the results. While it is difficult to predict the direction of potential bias, changes in the observed covariates in Table 2.1 suggests that, compared to the control group, the exposure group had a greater post-policy increase in demographic risk factors for inadequate prenatal care

and poor birth outcomes (i.e. being unmarried, non-White, and less educated). This pattern suggests that – if anything – this study may underestimate the provision's association with improved prenatal care and birth outcomes.

2.6 Conclusion

In this study of nearly 3 million births among women aged 24 to 25 years compared with those aged 27 to 28 years, the Affordable Care Act dependent coverage provision was associated with increased private insurance payment for birth, increased use of prenatal care, and modest reduction in preterm births and cesarean deliveries among unmarried women. The policy was not associated with changes in low birthweight or NICU admission.

2.7 References

1. Sommers BD, Buchmueller T, Decker SL, Carey C, Kronick R. The Affordable Care Act Has Led To Significant Gains In Health Insurance And Access To Care For Young Adults. *Health Aff (Millwood)*. 2013;32(1):165-174.

2. Sommers BD, Kronick R. The Affordable Care Act and Insurance Coverage for Young Adults. *JAMA*. 2012;307(9):913-914.

3. Chua K-P, Sommers BD. Changes in Health and Medical Spending Among Young Adults Under Health Reform. *JAMA*. 2014;311(23):2437-2439.

4. Han X, Yabroff KR, Robbins AS, Zheng Z, Jemal A. Dependent Coverage and Use of Preventive Care under the Affordable Care Act. *N Engl J Med*. 2014;371(24):2341-2342.

5. Akosa Antwi Y, Moriya AS, Simon K, Sommers BD. Changes in Emergency Department Use Among Young Adults After the Patient Protection and Affordable Care Act's Dependent Coverage Provision. *Ann Emerg Med*. 2015;65(6):664-672.e2.

6. Hernandez-Boussard T, Burns CS, Wang NE, Baker LC, Goldstein BA. The Affordable Care Act Reduces Emergency Department Use By Young Adults: Evidence From Three States. *Health Aff (Millwood)*. 2014;33(9):1648-1654. doi:10.1377/hlthaff.2014.0103

7. Akosa Antwi Y, Ma J, Simon K, Carroll A. Dependent Coverage under the ACA and Medicaid Coverage for Childbirth. *N Engl J Med*. 2016;374(2):194-196.

8. Rosenberg D, Handler A, Rankin KM, Zimbeck M, Adams EK. Prenatal Care Initiation Among Very Low-income Women in the Aftermath of Welfare Reform: does Pre-pregnancy Medicaid Coverage Make a Difference? *Matern Child Health J*. 2007;11(1):11-17.

9. Braveman, P, Marchi, K, Egerter, S, Pearl, M, Neuhaus, J. Barriers to Timely Prenatal Care Among Women With Insurance: The Importance of Prepregnancy Factors. *Obstet Gynecol*. 2000;95(6):874-880.

10. Partridge S, Balayla J, Holcroft CA, Abenhaim HA. Inadequate Prenatal Care Utilization and Risks of Infant Mortality and Poor Birth Outcome: A Retrospective Analysis of 28,729,765 U.S. Deliveries Over 8 Years. *Am J Perinatol*. 2012;29(10):787-794.

11. Osterman M, Martin J, Curtin S, Matthews TJ, Wilson E, Kirmeyer S. *Newly Released Data From the Revised U.S. Birth Certificate, 2011*. Centers for Disease Control and Prevention; 2013:1-22. https://www.cdc.gov/nchs/data/nvsr/nvsr62/nvsr62_04.pdf.

12. Bernstein AB, Cohen RA, Brett KM, Bush MA. Marital status is associated with health insurance coverage for working-age women at all income levels, 2007. *NCHS Data Brief*. 2008;(11):1-8.

13. Ryan AM, Burgess JF, Dimick JB. Why We Should Not Be Indifferent to Specification Choices for Difference-in-Differences. *Health Serv Res*. 2014.

14. Slusky DJG. Significant Placebo Results in Difference-in-Differences Analysis: The Case of the ACA's Parental Mandate. *East Econ J*. 2015:1-24.

15. Wallace J, Sommers BD. Health Insurance Effects on Preventive Care and Health. *Am J Prev Med*. 2016;50(5):S27-S33. doi:10.1016/j.amepre.2016.01.003

16. Miller S, Wherry LR. Health and Access to Care during the First 2 Years of the ACA Medicaid Expansions. *N Engl J Med*. 2017;376(10):947-956. doi:10.1056/NEJMsa1612890

17. Wherry LR, Fabi R, Schickedanz A, Saloner B. State And Federal Coverage For Pregnant Immigrants: Prenatal Care Increased, No Change Detected For Infant Health. *Health Aff (Millwood)*. 2017;36(4):607-615. doi:10.1377/hlthaff.2016.1198

18. Sommers BD, Gawande AA, Baicker K. Health Insurance Coverage and Health — What the Recent Evidence Tells Us. *N Engl J Med*. 2017;377(6):586-593. doi:10.1056/NEJMsb1706645

19. Daw JR, Hatfield LA, Swartz K, Sommers BD. Women In The United States Experience High Rates Of Coverage 'Churn' In Months Before And After Childbirth. *Health Aff (Millwood)*. 2017;36(4):598-606.

20. Kotelchuck M. An Evaluation of the Kessner Adequacy of Prenatal Care Index and a Proposed Adequacy of Prenatal Care Utilization Index. *Am J Public Health*. 1994;84(9):1414-1420.

21. Huesch MD. Association Between Type of Health Insurance and Elective Cesarean Deliveries: New Jersey, 2004–2007. *Am J Public Health*. 2011;101(11):e1-e7. doi:10.2105/AJPH.2011.300333

22. Aron DC, Gordon HS, Digiuseppe DL, Harper DL, Rosenthal GE. Variations in risk-adjusted cesarean delivery rates according to race and health insurance. *Med Care*. 38(1):35-44.

23. Epstein AM, Newhouse JP. Impact of Medicaid Expansion on Early Prenatal Care and Health Outcomes. *Health Care Financ Rev*. 1998;19(4):85-99.

24. Currie J, Gruber J. Public health insurance and medical treatment: the equalizing impact of the Medicaid expansions. *J Public Econ*. 2001;82(1):63-89. doi:10.1016/S0047-2727(00)00140-7

25. Howell EM. The Impact of the Medicaid Expansions for Pregnant Women: A Synthesis of the Evidence. *Med Care Res Rev*. 2001;58(1):3-30. doi:10.1177/107755870105800101

26. Showstack JA, Budetti PP, Minkler D. Factors associated with birthweight: an exploration of the roles of prenatal care and length of gestation. *Am J Public Health*. 1984;74(9):1003-1008.

27. Murray JL, Bernfield M. The differential effect of prenatal care on the incidence of low birth weight among blacks and whites in a prepaid health care plan. *N Engl J Med*. 1988;319(21):1385-1391. doi:10.1056/NEJM198811243192105

28. Currie J, Gruber J. Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women. *J Polit Econ*. 1996;104(6):1263-1296.

29. de Jongh BE, Locke R, Paul DA, Hoffman M. The differential effects of maternal age, race/ethnicity and insurance on neonatal intensive care unit admission rates. *BMC Pregnancy Childbirth*. 2012;12:97. doi:10.1186/1471-2393-12-97

30. Karaca-Mandic P, Norton EC, Dowd B. Interaction Terms in Nonlinear Models. *Health Serv Res*. 2012;47(1 Pt 1):255-274.

31. Bureau of Labor Statistics, U.S. Department of Labor. Labor Force Statistics from the Current Population Survey. https://www.bls.gov/cps/. Accessed May 16, 2017.

32. Sommers BD, Gunja MZ, Finegold K, Musco T. Changes in Self-reported Insurance Coverage, Access to Care, and Health Under the Affordable Care Act. *JAMA*. 2015;314(4):366-374. doi:10.1001/jama.2015.8421

33. Wherry LR, Miller S. Early Coverage, Access, Utilization, and Health Effects Associated With the Affordable Care Act Medicaid Expansions: A Quasi-experimental Study. *Ann Intern Med*. 2016;164(12):795. doi:10.7326/M15-2234

34. Markus AR, Andres E, West KD, Garro N, Pellegrini C. Medicaid Covered Births, 2008 Through 2010, In The Context Of The Implementation of Health Reform. *Womens Health Issues*. 2013;23(5):e273-280. doi:10.1016/j.whi.2013.06.006

35. Centers for Medicare and Medicaid. Eligibility. https://www.medicaid.gov/medicaid/eligibility/. Published 2017. Accessed June 19, 2017.

36. Issel LM, Forrestal SG, Slaughter J, Wiencrot A, Handler A. A Review of Prenatal Home-Visiting Effectiveness for Improving Birth Outcomes. *J Obstet Gynecol Neonatal Nurs*. 2011;40(2):157-165.

37. Alexander GR, Kotelchuck M. Assessing the role and effectiveness of prenatal care: history, challenges, and directions for future research. *Public Health Rep*. 2001;116(4):306-316.

38. Fiscella K. Does prenatal care improve birth outcomes? A critical review. *Obstet Gynecol*. 1995;85(3):468-479.

CHAPTER 3: Matching and Regression-to-the-Mean in Difference-in-Differences Analysis

Authors and affiliations
Jamie R. Daw, M.Sc.; Harvard PhD Program in Health Policy, Cambridge, MA

Laura A. Hatfield, Ph.D.; Department of Health Care Policy, Boston, MA

Jamie R. Daw and Laura A. Hatfield. Matching and Regression-to-the-Mean in Difference-in-Differences Analysis. *Health Services Research,* doi: 10.1111/1475-6773.12993.

3.1 Abstract

Objective. To demonstrate regression-to-the-mean bias introduced by matching on pre-period variables in difference-in-differences studies.

Data Sources. Simulated data.

Study Design. We performed a Monte Carlo simulation to estimate the effect of a placebo intervention on simulated longitudinal data for units in treatment and control groups using unmatched and matched difference-in-differences analyses. We varied the pre-period level and trend differences between the treatment and control groups, and the serial correlation of the matching variables. We assessed estimator bias as the mean absolute deviation of estimated program effects from the true value of zero.

Principal Findings. When pre-period outcome level is correlated with treatment assignment, an unmatched analysis is unbiased, but matching units on pre-period outcome levels produces biased estimates. The bias increases with greater pre-period level differences and weaker serial correlation in the outcome. This problem extends to matching on pre-period level of a time-varying covariate. When treatment assignment is correlated with pre-period trend only, the unmatched analysis is biased, and matching units on pre-period level or trend does not introduce additional bias.

Conclusions. Researchers should be aware of the threat of regression-to-the-mean when constructing matched samples for difference-in-differences. We provide guidance on when to incorporate matching in this study design.

Keywords: observational research, matching, difference-in-differences.

3.2 Introduction

Difference-in-differences is one of the most popular approaches for measuring the effect of health policies and programs.[1] A difference-in-differences analysis can be employed any time one measures an outcome in units that have been exposed to a policy or program (the treatment group) and those that have not (the control group) both before and after the intervention is implemented. Here, "group" means the level at which treatments are applied (e.g., states, clinics) and "unit" means the level at which measurements are made (e.g., counties within states, patients within clinics, etc.). The effect of the treatment is calculated as the difference in the outcome between the treatment and control groups after the intervention minus the difference before the intervention. One concludes that the intervention affected the outcome if the difference between the two groups changes from the pre-period to the post-period.

As the popularity of difference-in-differences has risen, so has the application of matching methods to this study design. The objective of matching is to reduce potential confounding by improving the comparability of units in the treatment and control groups. In the context of difference-in-differences, researchers identify a subset of potential confounders and match units from the treatment and control group on measures of these variables prior to the intervention. The effect of the intervention is then estimated using this matched sample.

Difference-in-differences studies with matched samples have been used to evaluate the impact of a variety of health policies and programs including high-deductible health plans [2,3], team-based and coordinated care programs [4–6], multi-payer medical homes [7], telehealth programs [8], home-visiting programs [9], hospital closures [10], workplace wellness programs [11], and quality

reporting.[12,13] In this paper, we demonstrate how matching in difference-in-differences can introduce a well-known statistical phenomenon, regression-to-the-mean, resulting in biased estimates of intervention effects. We also offer practical guidance to researchers on how to select matching variables to minimize this important threat to validity.

3.2.1 Basics of Difference-in-Differences

We focus on the widely-used "micro-level" difference-in-differences model in which the treatment is assigned to a group (e.g., health plans adopt benefit design changes) and observations are available for units within groups (e.g., health spending of enrollees within health plans) before and after an intervention.[1]

A difference-in-differences study is usually analyzed with a regression model such as,

$$Y_{ijt} = \beta_0 + \beta_1 treatment_j + \beta_2\, post_t + \delta\big(treatment_j * post_t\big) + \varepsilon_{ijt}\,, \quad (1)$$

where *i* indexes units, *j* indexes groups, *t* indexes time, *treatment* is an indicator for whether a group was treated, *post* is an indicator for whether a measurement was taken in the post-treatment period, and ε is the random error term. We consider the simplest case, in which there are only two groups: one treatment group and one control group. The difference-in-differences estimate of the average treatment effect on the treated is $\hat{\delta} = \big(\bar{Y}_{treatment,post} - \bar{Y}_{treatment,pre}\big) - (\bar{Y}_{control,post} - \bar{Y}_{control,pre})$.

3.2.2 Confounding in Difference-in-Differences

One advantage of difference-in-differences, relative to cross-sectional designs, is that it does not require treatment and control groups to have similar baseline means, often referred to as pre-

period "levels", in the outcome or other covariates. This is because the design measures the effect of an intervention as the *relative change* in the outcomes between units in the treatment and control groups over time. As a result, the definition of confounding differs from cross-sectional studies. A confounder of a difference-in-differences study is any variable related to both treatment assignment and the *change in the outcome over time* (i.e., the trend). Contrast this to a confounder in a cross-sectional study, which is any variable related to both treatment assignment and the *level of the outcome at a point in time*. Of course, a variable may be related to both the level and trend in the outcome and therefore a confounder in both senses. The point is that variables related only to treatment assignment and outcome level (not trend) do not bias difference-in-differences studies. They are not confounders and therefore not a useful target of matching that is intended to reduce bias due to observable confounders.

Like other observational designs, difference-in-differences still requires a strong assumption to produce unbiased causal estimates of the treatment effect: the change from pre- to post-period in the control group is a valid counterfactual for the change that would have occurred in the treatment group in the absence of the intervention. This is often stated as two assumptions, "common shocks" and "parallel trends", referring, respectively, to the assumption that events during the study period affect the treatment and control groups equally and the assumption that the two groups would have equal trends in the post period if not for the intervention.[14] This is also equivalent to assuming no unobserved confounding, where we emphasize that confounding in difference-in-differences relates only to variables correlated with treatment assignment and outcome trends. Assessment of these assumptions, which we refer to collectively as the "counterfactual assumption", would require observing an alternative reality (i.e., the

counterfactual change in the treatment group outcomes in the absence of intervention), and is therefore impossible.

3.2.3 Matching in Difference-in-Differences

Even though differences between treatment and control units at baseline are not a threat to validity *per se*, researchers often match units from the treatment and control groups on pre-period measures of the outcome or other variables. Doing so attempts to correct for confounding bias by balancing on variables that are different in treatment and control groups. Matching uses pre-period measurements of three kinds of variables: covariates, outcome levels, and outcome trends.

In this paper, we do not consider the possibly beneficial application of matching on covariates that differ between the groups and are correlated with future outcome *trends* (i.e., matching on confounders in the difference-in-differences sense). Instead we consider matching on covariates that are correlated with *levels* of the outcome (or are the outcome level itself), which is the source of the regression-to-the-mean bias we discuss here.

We also consider matching on pre-period outcome trends. Difference-in-differences does not strictly require the treatment and control groups to have similar trends in the outcome prior to the intervention. However, divergence in pre-period trends is usually seen as a strong indication that the counterfactual assumption is violated. Thus, researchers may match on pre-period outcome trends when they suspect that pre-period trends are correlated with the change in the outcome from the pre-period to post-period. The hope is that balance on pre-period trends strengthens the

plausibility of the counterfactual assumption. In this paper, we show that matching on pre-period trends may have limited benefit.

3.2.4 Regression to the Mean and Bias in Matched Difference-in-Differences Analysis

The statistical phenomenon of regression-to-the-mean (aka the "regression fallacy") occurs when a group is selected for extreme values of one variable and then another variable is measured for that group. In the longitudinal setting relevant to difference-in-differences studies, it occurs between repeated measures of the same variable; for example, if we measure weight for a sample of individuals today and select a subset of individuals with higher than average weight, those individuals will have a mean weight that is closer to average upon subsequent measurement. Two factors determine the magnitude of regression-to-the-mean effects.[15] First, regression effects increase as units are selected further away from their group mean. Second, regression effects decrease with increased correlation between the sample selection variable and the other measured variable. In longitudinal settings, this means that regression effects decrease with increased correlation between measures over time (i.e., the serial correlation). Selecting units that have extreme values of a variable that is unstable over time will produce large regression-to-the-mean effects. Variables that do not vary over time, such as sex or region, are not subject to longitudinal regression-to-the-mean effects.

The vulnerability of matched samples to regression-to-the-mean has been known since the work of McNemar in the 1940s, yet receives little contemporary attention in guidance on matching methods or in discussions of the validity of published evaluations using matched samples.[16,17] The simple idea is that matching is a sample selection technique; it selects units that are extreme

relative to their respective group means to achieve balance in the matched sample. For example, when the treatment group mean is larger than the control group mean, matching will select control units that are higher than average (relative to all control units) and treatment units that are lower than average (relative to all treatment units). If the variables on which units are matched vary over time, matched units will "regress back" toward the means of the groups from which they were selected. More precisely, this phenomenon might be called "regression to the means" as matched units from treatment and control groups regress back to their respective means over time. For simplicity, we use the idiomatic expression "regression to the mean".

Longitudinal studies such as difference-in-differences may reach biased conclusions in the presence of regression-to-the-mean effects. Recall that a confounder of a difference-in-differences study is any variable related to both treatment assignment and the *change in the outcome over time* (i.e., the trend). The process of matching *introduces* such a confounder—the indicator of whether a unit is selected into the sample by matching—that is (inadvertently) correlated with the change in the outcome over time because of regression-to-the-mean. As an illustration, consider a voluntary program that increases cost-sharing for enrollees in employer-sponsored health plans. Suppose that a firm that adopts the program does so in response to higher mean enrollee health spending compared to a firm that does not adopt the program. If we match enrollees on baseline health spending, we will select control enrollees with higher-than-average spending (relative to the control firm mean) and treatment enrollees with lower-than-average spending (relative to the treatment firm mean) to achieve a balanced sample. Even if the program has no true effect, on subsequent measurement, average spending will decrease among the matched control enrollees and increase among matched treatment enrollees purely because of

regression toward their respective group means. The divergent trends between the two matched groups violates the critical counterfactual assumption of difference-in-differences. This could lead to the false conclusion that the program increased health spending. This threat to validity is a particular concern because it results from a researcher applying matching to what otherwise would be an unbiased analysis.

The magnitude of the bias introduced in a matched difference-in-differences analysis will vary with the magnitude of the regression-to-the-mean, which depends on (1) how extreme the measures of the matching variable are in the matched sample relative to those of the unmatched sample (e.g., the magnitude of the pre-period differences between the treatment and control groups), and (2) the correlation between the matching variable and the post-period outcome (e.g., the serial correlation between measures of the outcome).

3.3 Methods

We use a Monte Carlo simulation to demonstrate the bias resulting from regression-to-the-mean in matched difference-in-differences analysis. We generate data under different causal scenarios, create both matched and unmatched samples, and analyze them with standard difference-in-differences regression models to estimate the effect of a null intervention, for which the true treatment effect is zero.

3.3.1 Data Generation

We generate data under four causal scenarios, detailed in Appendix A3 and summarized here. Three of the scenarios are unconfounded (for difference-in-differences) because treatment assignment is 1) completely random (i.e., unrelated to group mean levels or trends), 2) correlated

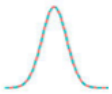
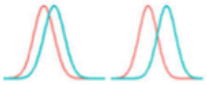
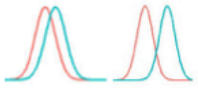
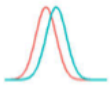
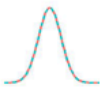
only with group mean outcome level, or 3) correlated with group mean level of a time-varying covariate that is in turn correlated only with the group mean outcome level. The fourth scenario is confounded because treatment assignment is correlated with group mean outcome trend, thus violating the counterfactual assumption. In each scenario, there are no additional unobserved variables that determine treatment assignment.

For each scenario, we generate data for 1,000 units in the treatment group and 1,000 units in the control group. For each unit, we generate eight observations of a normally distributed outcome, centered around the respective group mean levels and trends. Because we generate under a null intervention, the group mean levels and trends of the outcome are constant over the pre and post periods. The unit measurements are assumed to be equally spaced and could be conceptualized, for example, as eight quarterly measurements over two years. The repeated measures follow an autoregressive covariance structure of order 1 [i.e., AR(1)] with constant variance and a single correlation parameter (see Figure A3.1). The AR(1) structure implies that a unit's measurements from one time period to the next are positively correlated, and that this correlation decays exponentially for measurements that are further apart in time.

For the third causal scenario, we also generate eight observations of a normally distributed covariate for each unit, centered around the respective means of the treatment and control groups. These repeated covariate measures also follow an AR(1) covariance structure and have an additional correlation parameter that controls the strength of the relationship with the unit outcome. Note that this covariate is not a confounder because it is only correlated with outcome level, not with changes in the outcome over time.

Within each causal scenario, we also vary factors that affect the magnitude of regression-to-the mean effects: serial correlation across observations of both the outcome and the covariate, strength of correlation between the outcome and covariate, and pre-period level and trend difference between the treatment and control groups. Table 3.1 summarizes the causal data-generating scenarios.

Table 3.1 Summary of simulation scenarios

Group-Level Treatment Assignment	Mean Pre-Period Difference Between Treatment and Control Groups			Confounded?	Pre-Period Matching Variable	Results Figure
	Outcome Level	*Outcome Trend*	*Covariate Level*			
Random	 *No difference*	*No difference*	n/a	No	None Outcome level Outcome trend	3.1a 3.1b 3.1c
Correlated with outcome level	 *1 or 2 SD*	*No difference*	n/a	No	None Outcome level Outcome trend	3.1a 3.1b 3.1c
Correlated with covariate level	 *Depends on X,Y correlation (weak, moderate, or strong)*	*No difference*	 *1 SD*	No	None Covariate level	3.2a 3.2b
Correlated with outcome trend	 *No difference*	 *0.01, 0.05, or 0.1 SD*	n/a	Yes	None Outcome level Outcome trend	3.3a 3.3b 3.3c

Notes: The serial correlation of the matching variable varies across simulation iterations within each scenario. SD refers to standard deviation of the covariate or outcome.

3.3.2 Analysis

Of the eight observations for each unit, we assume four occur in the pre-period and four in the post-period. The unmatched samples are simply all the simulated units. We generate matched samples by matching on (1) the pre-period level of the outcome, (2) the pre-period level of a

covariate correlated with the outcome (for causal scenario 3), or (3) the pre-period trend in the outcome, calculated using linear regression estimates of the pre-period slope for each unit. We use one-to-one, nearest-neighbor matching (without replacement) with a caliper of 0.2 SD of the matching variable.

For each unmatched and matched sample, we estimate the effect of the intervention using the "micro-level" difference-in-differences regression estimator (Equation 1) with standard errors clustered by unit. We summarize bias using the mean absolute deviation between the estimated effect and the true value of zero. We average the absolute deviation over the 1000 simulation replicates and scale the result in terms of standard deviations of the outcome variable. Table 3.1 indicates the figures that display the corresponding results.

3.3.3 Limitations

Since our objective is to demonstrate a well-known statistical phenomenon (regression-to-the-mean) in the setting of a popular study design (difference-in-difference with matched samples), we apply relatively simple scenarios, which do not reflect all the potential causal scenarios in which difference-in-differences analysis could be applied. However, the underlying mechanism of treatment assignment in real data is rarely known and thus, these simple cases are often assumed to hold. The magnitude of the bias estimated in our study is specific to data-generating processes that resemble our simulation scenarios. For instance, we simulate only under null intervention effects because we are interested in bias. In addition, while the AR(1) correlation structure is reasonable, other correlation structures are possible and this may have implications for the magnitude of the results. Similarly, we focus exclusively on normally distributed

variables and OLS regression, which is the standard modeling approach used for difference-in-differences analysis. While regression-to-the-mean effects do not depend on the specific distribution of the outcomes and covariates, the magnitude of the results may vary for non-Normal variables and other treatment estimators.

3.4 Results

3.4.1 Treatment Randomly Assigned or Correlated With Pre-Period Level Only

When treatment is correlated with pre-period levels of the outcome, matching units on baseline outcome measures can introduce bias in an otherwise unbiased analysis. Figure 3.1 shows the results for scenarios where treatment is randomly assigned or correlated with pre-period outcome level only. The figure displays the estimator bias for three matching strategies across different strengths of serial correlation in the outcome and different baseline level differences between the treatment and control groups (ranging from zero to two standard deviations). The estimates from the unmatched samples are unbiased across all strengths of serial correlation and for all mean pre-period level differences (see Figure 3.1(a)). This is what we expect: pre-period differences in level do not bias difference-in-differences analyses. However, when we apply difference-in-differences analysis to samples matched on baseline level, bias is introduced. As shown in Figure 3.1(b), the form of the bias is a classic manifestation of regression-to-the-mean: regression effects increase with decreasing serial correlation in the outcome and increasing baseline level differences between the treatment and control group. In the most extreme case, where the serial correlation in the outcome is zero, the two groups regress entirely back to their original baseline differences, resulting in large, spurious treatment effects. As shown in Figure 3.1(c), when treatment is not correlated with trends in the outcome, matching on pre-period trend does not

introduce bias. As in the case where treatment is assigned randomly, this is because matching does not result in a selection of units that are extreme relative to their group means.

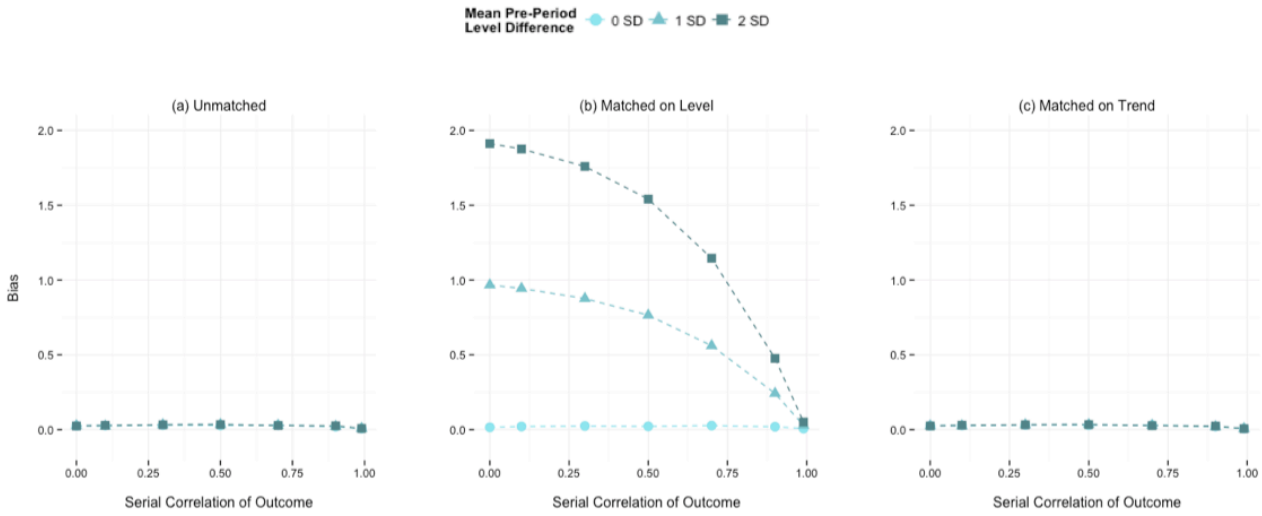


Figure 3.1 Bias of matching strategies for group-level treatment *randomly assigned* or correlated with pre-period *level only*

Notes: The serial correlation of the outcome refers to the autoregressive parameter of the AR(1) correlation structure. Bias is measured as the absolute deviation of the treatment estimate from zero in standard deviations of the outcome. SD is standard deviations of the outcome. When the mean pre-period level difference is 0 SD, group-level treatment is randomly assigned. When the mean pre-period level difference is 1 or 2 SD, group-level treatment is correlated with pre-period level.

Matching to reduce baseline differences in a time-varying covariate that is correlated with the outcome can produce the same bias problems. Figure 3.2(b) shows the bias of matching on the pre-period level of a covariate that is correlated with the outcome. The bias behaves similarly to matching on the outcome itself, increasing with decreasing serial correlation in the covariate, as well as with increasing pre-period difference between the two groups. However, compared to Figure 3.1(b), the magnitude of the bias is proportional to the correlation between the covariate and the outcome. In other words, the stronger the relationship between the covariate and the outcome, and the greater the mean difference between the two groups at baseline, the greater the bias. We find that matching on pre-period covariates does not introduce bias when (1) the

covariates are not correlated with both outcome and treatment, and (2) the covariates are fixed or very highly correlated over time. As before, the unmatched analysis is unbiased, see Figure 3.2(a).

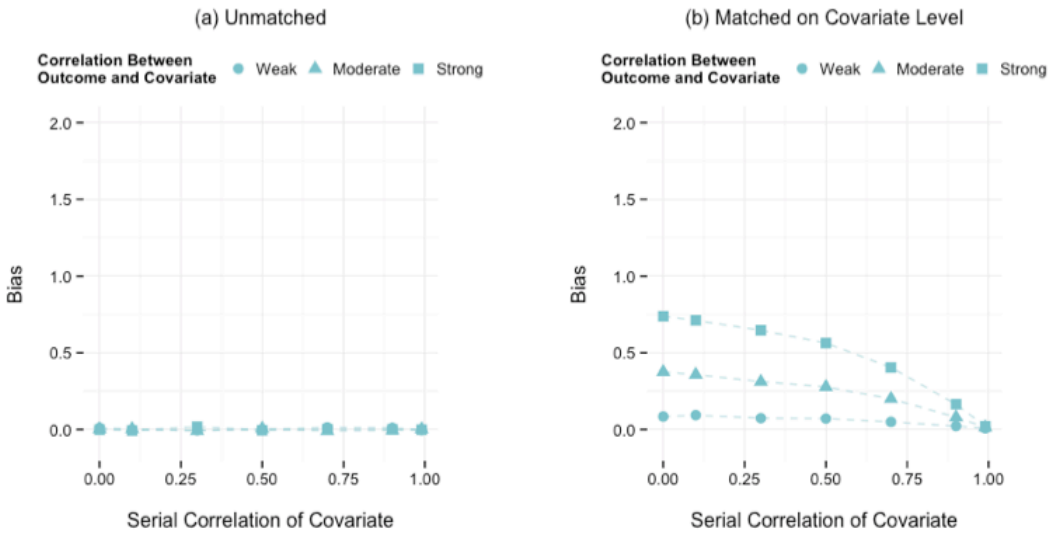


Figure 3.2 B Bias of matching strategies for group-level treatment *randomly assigned* or correlated with pre-period *level only*

Notes: The serial correlation of the outcome and covariate refers to the autoregressive parameter of the AR(1) correlation structure. The mean pre-period level difference in the covariate between the treatment and control groups is 1 standard deviation across all iterations. Bias is measured as the absolute deviation of the treatment estimate from zero in standard deviations of the outcome.

3.4.2 Treatment Correlated With Pre-Period Trends

When the counterfactual assumption is violated, we find that matching is insufficient to overcome the bias. Figure 3.3 shows the results for unmatched and matched samples for scenarios where treatment assignment is correlated with pre-period trend only. In this scenario, the causal assumption is violated because the pre-period differences in trend persist in the post period, making the control group's change over time an invalid counterfactual for the treatment group's change. As expected, this violation results in biased estimates when using the unmatched samples. As shown in Figure 3.3(a), the bias is proportional to the differences in trends between

the treatment and control group. Matching on pre-period level does not introduce additional bias (see Figure 3.3(b)) because the samples were generated with no correlation between treatment and pre-period outcome trend, so matching does not select extreme units and there is no regression to the mean. However, as shown in Figure 3.3(c), matching on trend neither introduces additional bias nor fully corrects for the violation. We find a small decrease in bias in the presence of very high serial correlation in the outcome. When serial correlation is low, pre-period trend is a poor predictor of the pre- and post-period difference; however, when serial correlation is high, pre-period trends are more stable and more predictive of future trends, resulting in a modest reduction in bias in samples matched on trend (see Figure A3.2).

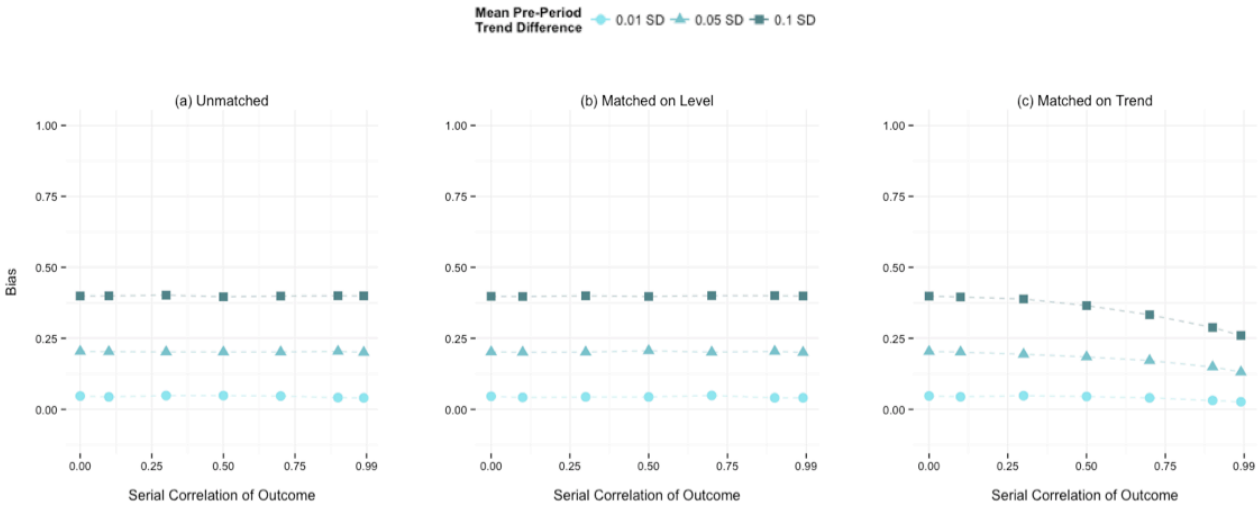


Figure 3.3 Bias of matching strategies for group-level treatment *randomly assigned* or correlated with pre-period *trend only*

Notes: The serial correlation of the outcome refers to the autoregressive parameter of the AR(1) correlation structure. Bias is measured as the absolute deviation of the treatment estimate from zero in standard deviations of the outcome. The maximum serial correlation in the outcome is 0.99 because no units can be matched on trend if the outcome is perfectly correlated over time. SD is standard deviations of the outcome.

3.5 Discussion

Longitudinal study designs using matched samples have long been known to be vulnerable to the problem of regression-to-the-mean. Our results emphasize that health services researchers should

be aware of this threat when constructing matched samples for difference-in-differences analysis. When treatment group assignment is correlated with outcome level in the pre-period, we find that matching units on pre-period outcome level can introduce bias to an otherwise unbiased analysis. The magnitude of the bias increases with 1) greater pre-period level differences between the treatment and control group, and 2) lower serial correlation of the outcome. This problem extends to samples matched on pre-period level of a time-varying covariate when the covariate is correlated both with treatment and the outcome. The bias is minimal when pre-period level differences are small or serial correlation is very high, and is absent when there are no pre-period level differences or the matching variable is constant.

These results imply a challenging paradox. The greater the pre-period level differences between the treatment and control group in the outcome or a covariate that is correlated with the outcome, the more inclined researchers may be to match on pre-period level. However, the greater the pre-period level difference, the larger the bias that can result from matching due to regression-to-the mean. Thus, researchers should not attempt to 'match-away' level differences in time-varying variables, as doing so can introduce bias. Indeed, this practice is particularly unnecessary given that pre-period level differences *per se* are not a violation of the causal assumption of differences-in-differences.

Researchers often want to exercise prudence and include all potential confounders in their matching model. As a result, matching on baseline levels, and increasingly baseline trends, has become a mainstream practice. Common pre-period matching variables include time-varying individual-level variables, such as number of primary care visits, and time-varying practice- and

regional-level variables such as rates of emergency department visits or risk-adjusted mortality. In many studies, matching models include pre-period measures of the study outcome itself, for example, matching on baseline levels of health care spending in a study of the effect of a policy on spending. To give context to the strengths of serial correlation shown in our results, Table A3.3 shows the year-to-year correlation for a selection of variables measured on people and hospital service areas.

Our results caution against the popular "kitchen sink" approach to including pre-period variables in matching models. We would argue that the inclusion of matching variables ought to be considered carefully given the potential for bias due to regression effects. Our results also show when matching units on baseline measures does not introduce a risk of regression-to-the-mean: when there is good pre-period balance between the two groups, strong or perfect serial correlation in the matching variable, or when the association between the matching variable and the outcome is weak.

3.5.1 Guidance for Selecting Matching Variables in Difference-in-Differences Analysis

Based on previous research and the results of our simulations, we present a flowchart for selecting matching variables for difference-in-differences analysis (Figure 3.4). We assume a process whereby researchers consider matching approaches in response to inspections for pre-period differences such as those that have been recommended in other general checklists for this design (for example, those proposed by Ryan, Burgess, and Dimick 2014).[1]

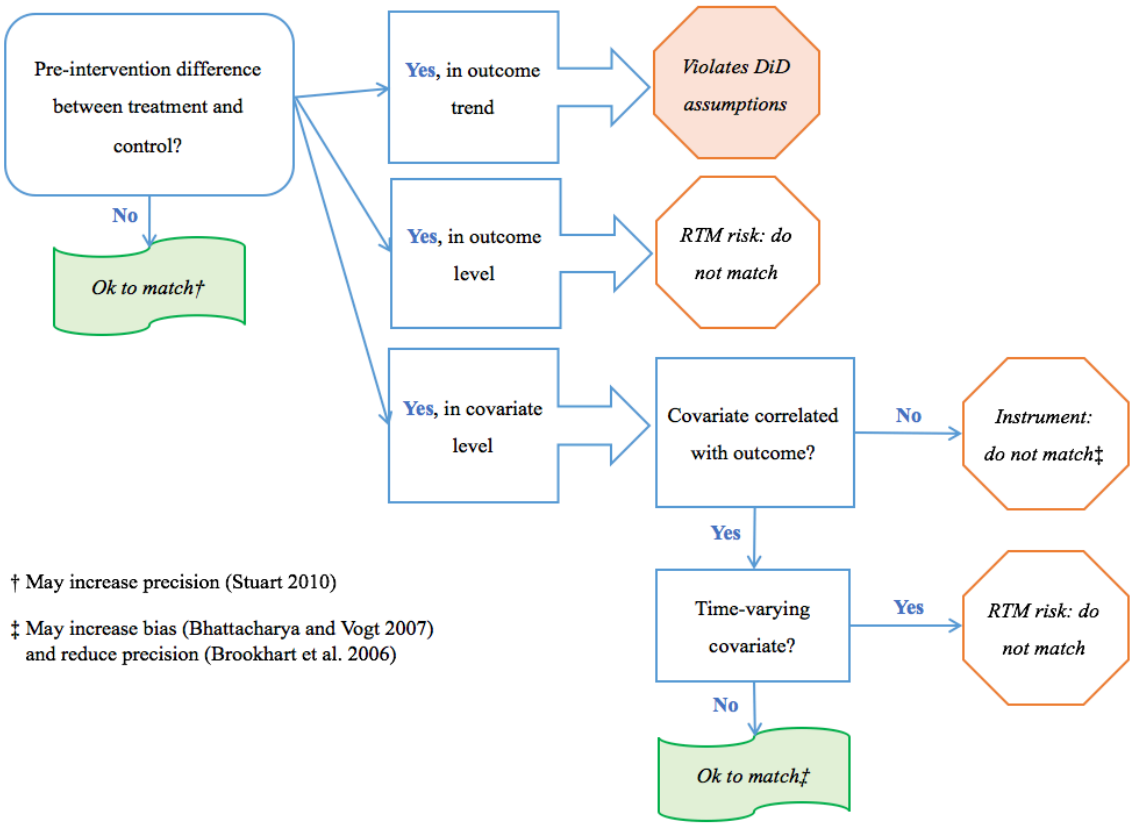


Figure 3.4 Selecting matching variables for difference-in-differences analysis

3.5.1.1 Pre-Period Differences in Outcome Level

While a baseline difference in outcome level between the treatment and control group is not itself a threat to validity of a difference-in-differences analysis, in practice, it may raise concerns that the control group is not a valid counterfactual. In response to finding a pre-period difference in the outcome level, researchers should consider whether pre-period level is likely to be correlated with the pre-post change, possibly with empirical tests using the control group data. If there is no or a weak relationship between pre-period level and the pre-post change (and thus no threat to validity), researchers should proceed without matching on pre-period level. As we

demonstrate, in this scenario, matching units on pre-period level will unnecessarily introduce bias when the pre-period level difference is large and the serial correlation is low.

3.5.1.2 Pre-Period Differences in Covariate Level

Again, while pre-period covariate differences between treatment and control groups are not a threat to validity for difference-in-differences, they may undermine confidence in the control group as a valid counterfactual. As for differences in pre-period levels of the outcome, this concern is justified if pre-period levels of the covariate are predictors of the pre-post change in the outcome. In the face of pre-period covariate differences, researchers should draw on scientific understanding and empirical evidence for the relationship between pre-period covariate levels and pre-post outcome changes. A covariate that differs between the two groups at baseline and is correlated with pre-post outcome changes is an appropriate matching variable if it is stable over time (e.g. sex, race, or region). Empirical estimates of the control group's repeated measures correlation matrix may be informative. If the covariate is fixed or highly correlated over time, researchers may proceed with matching on pre-period levels of the covariate without risk of regression-to-the-mean. If it is weakly or moderately correlated over time, researchers should avoid matching on the variable or consider stabilizing transformations. For example, one could transform a continuous variable with moderate serial correlation (such as a hospital quality score) into a categorical variable (such as low, medium, and high).

A covariate that differs between the two groups at baseline but is not associated with the outcome is an instrument and should not be used as a matching variable. Doing so can increase bias and decrease the precision of treatment estimates.[18,19]

Finally, baseline covariates that are not substantially different between treatment and control but are correlated with the outcome may be good candidates for matching, as their inclusion may yield more precise treatment estimates.[20] As we show, matching on baseline variables that do not differ between treatment and control does not introduce the threat of regression-to-the-mean.

3.5.1.3 Pre-Period Differences in Outcome Trend

There is no empirical test of the validity of the counterfactual assumption of a difference-in-differences analysis. However, researchers try to bolster confidence that the assumption holds by looking for evidence of differences in pre-period trends, which may indicate that the control group is not a valid counterfactual. However, similar pre-period trends are no guarantee that the similarities would persist in the post-period in the absence of the intervention. More important, in our data-generating scenarios, matching on pre-period trend does not address violations of the counterfactual assumption, although matching may provide small bias reductions when the trends are highly stable. Matching on trend does not introduce additional bias (relative to unmatched samples) because the regression-to-the-mean effect simply pulls the post-period trends back to the original violation. However, matching on trend can lead to misleading assessments of the validity of the counterfactual assumption because plots of pre-period trends in matched samples will appear parallel (as matching forces them to be so); yet unless the trends are highly stable, this comparability will break down in the post period due to regression-to-the-mean trend in each group (see Figure A3.3).

It is possible that other matching techniques could address violations of the counterfactual assumption. Methodologists and applied researchers are experimenting with new approaches such as cross-temporal matching algorithms.[21,22] The synthetic control method is yet another alternative for constructing a comparable control group based on pre-period measures.[23] There is a need for further simulation studies to compare the performance of techniques researchers are applying in the field–including the potential for regression-to-the-mean bias–under varying data-generating and treatment assignment scenarios.

3.6 Conclusion

Matching is an increasingly popular approach for improving the comparability of treatment and control groups in difference-in-differences analysis. Limited attention has been given to how matching, as a sample selection technique, can introduce bias due to regression-to-the-mean. Our results show that matching can have an important impact on estimated intervention effects, particularly when matching on pre-period levels of the outcome itself or on time-varying covariates with low serial correlation. We provide guidance on when to incorporate matching in difference-in-differences based on observed evidence of potential violations of the assumptions of the design.

3.7 References

1. Ryan, A. M., Burgess, J. F. & Dimick, J. B. Why We Should Not Be Indifferent to Specification Choices for Difference-in-Differences. *Health Serv. Res.* (2014).

2. Wharam, J. F. *et al.* Emergency Department Use and Subsequent Hospitalizations Among Members of a High-Deductible Health Plan. *JAMA* **297,** 1093–1102 (2007).

3. Waters, T. M., Chang, C. F., Cecil, W. T., Kasteridis, P. & Mirvis, D. Impact of High-Deductible Health Plans on Health Care Utilization and Costs. *Health Serv. Res.* **46,** 155–172 (2011).

4. Scanlon, D. P. *et al.* Financial and Clinical Impact of Team-Based Treatment for Medicaid Enrollees With Diabetes in a Federally Qualified Health Center. *Diabetes Care* **31,** 2160–2165 (2008).

5. Xing, J., Goehring, C. & Mancuso, D. Care Coordination Program For Washington State Medicaid Enrollees Reduced Inpatient Hospital Costs. *Health Aff. (Millwood)* **34,** 653–661 (2015).

6. Dale, S. B. *et al.* Two-Year Costs and Quality in the Comprehensive Primary Care Initiative. *N. Engl. J. Med.* **374,** 2345–2356 (2016).

7. Rosenthal, M. B. *et al.* A Difference-in-Difference Analysis of Changes in Quality, Utilization and Cost Following the Colorado Multi-Payer Patient-Centered Medical Home Pilot. *J. Gen. Intern. Med.* **31,** 289–296 (2015).

8. Baker, L. C., Johnson, S. J., Macaulay, D. & Birnbaum, H. Integrated Telehealth And Care Management Program For Medicare Beneficiaries With Chronic Disease Linked To Savings. *Health Aff. (Millwood)* **30,** 1689–1697 (2011).

9. Mattke, S., Han, D., Wilks, A. & Sloss, E. Medicare Home Visit Program Associated With Fewer Hospital And Nursing Home Admissions, Increased Office Visits. *Health Aff. (Millwood)* **34,** 2138–2146 (2015).

10. Joynt, K. E., Chatterjee, P., Orav, E. J. & Jha, A. K. Hospital Closures Had No Measurable Impact On Local Hospitalization Rates Or Mortality Rates, 2003–11. *Health Aff. (Millwood)* **34,** 765–772 (2015).

11. Caloyeras, J. P., Liu, H., Exum, E., Broderick, M. & Mattke, S. Managing Manifest Diseases, But Not Health Risks, Saved PepsiCo Money Over Seven Years. *Health Aff. (Millwood)* **33,** 124–131 (2014).

12. Hollenbeak, C. S. *et al.* Reductions in Mortality Associated With Intensive Public Reporting of Hospital Outcomes. *Am. J. Med. Qual.* **23,** 279–286 (2008).

13. Osborne, N. H., Nicholas, L. H., Ryan, A. M., Thumma, J. R. & Dimick, J. B. Association of Hospital Participation in a Quality Reporting Program With Surgical Outcomes and Expenditures for Medicare Beneficiaries. *JAMA* **313,** 496–504 (2015).

14. Angrist, J. & Pischke, J.-S. *Mostly Harmless Econometrics*. (Princeton University Press, 2009).

15. Barnett, A. G., Pols, J. C. van der & Dobson, A. J. Regression To The Mean: What It Is And How to Deal With It. *Int. J. Epidemiol.* **34,** 215–220 (2005).

16. McNemar, Q. Sampling in Psychological Research. *Psychol. Bull.* **37,** 331–365 (1940).

17. Althauser, R. P. & Rubin, D. Measurement Error and Regression to the Mean in Matched Samples. *Soc. Forces* **50,** 206–214 (1971).

18. Bhattacharya, J. & Vogt, W. B. *Do Instrumental Variables Belong in Propensity Scores?* (Social Science Research Network, 2007).

19. Brookhart, M. A. *et al.* Variable selection for propensity score models. *Am. J. Epidemiol.* **163,** 1149–1156 (2006).

20. Stuart, E. A. Matching Methods for Causal Inference: A Review and a Look Forward. *Stat. Sci.* **25,** 1–21 (2010).

21. Stuart, E. A. *et al.* Using Propensity Scores in Difference-in-Differences Models to Estimate The Effects of a Policy Change. *Health Serv. Outcomes Res. Methodol.* **14,** 166–182 (2014).

22. Gozalo, P., Plotzke, M., Mor, V., Miller, S. C. & Teno, J. M. Changes in Medicare Costs with the Growth of Hospice Care in Nursing Homes. *N. Engl. J. Med.* **372,** 1823–1831 (2015).

23. Abadie, A., Diamond, A. & Hainmueller, J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *J. Am. Stat. Assoc.* **105,** 493–505 (2010).

A1. Appendix for Chapter 1

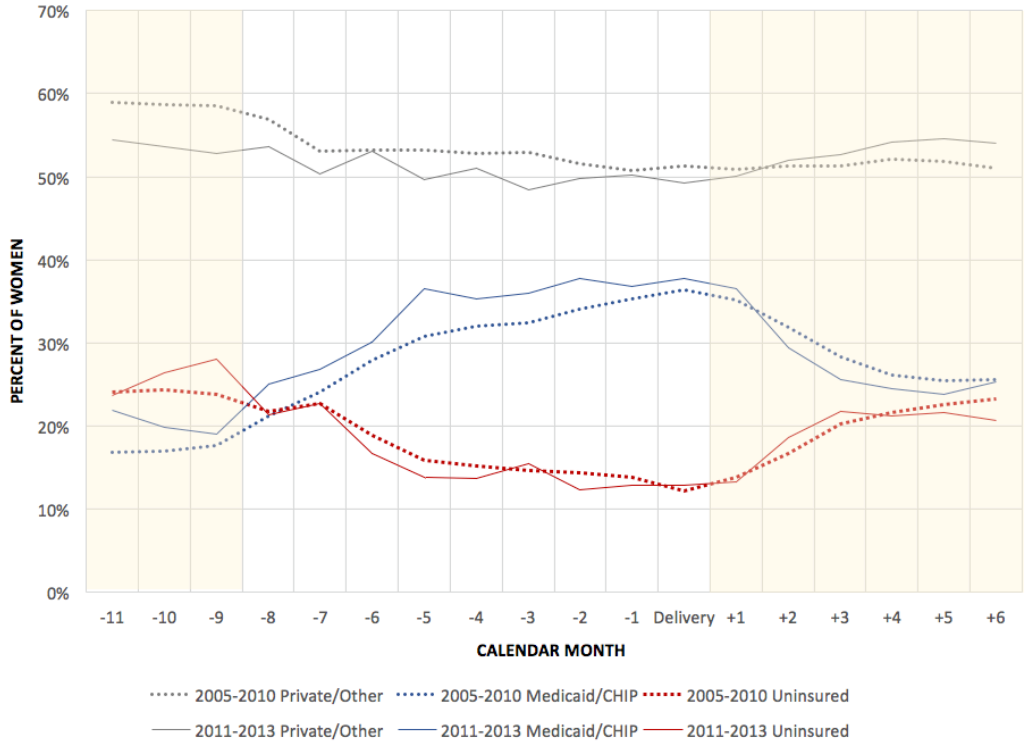


Figure A1.1 Percentage of women insured by coverage type and calendar month 2005-2010 compared to 2011-2013

Notes: Authors' analysis of data for 2005-13 from panels 10-17 of the Medical Expenditure Panel Survey-Household Component. 2005-2010 n=1,891; 2011-2013 n=835.

Table A1.1 Overall insurance status changes from nine months before delivery to delivery and delivery to six months after delivery

BEFORE DELIVERY	**Insurance at Delivery Month**		
Insurance 9 Calendar Months Before Delivery Month	Private or other	Medicaid or CHIP	No Insurance
Private or other	83%	9%	8%
Medicaid or CHIP	1%	91%	8%
No Insurance	11%	51%	38%

AFTER DELIVERY	**Insurance 6 Months After Delivery Month**		
Insurance at Delivery Month	Private or other	Medicaid or CHIP	No Insurance
Private or other	92%	<0.5%	8%
Medicaid or CHIP	5%	66%	29%
No Insurance	27%	12%	61%

Notes: Authors' analysis of data for 2005-13 from panels 10-17 of the Medical Expenditure Panel Survey-Household Component. Only women in the survey sample for all calendar months in each period are included (before delivery, n = 1,751; after delivery, n = 2,036).

Table A1.2 Month-to-month insurance transitions by type

PRENATAL	-9	-8	-7	-6	-5	-4	-3	-2	-1	Delivery Month
Any Transition From Previous Month	--	8%	12%	15%	13%	11%	13%	11%	10%	9%
Private - Medicaid	--	1%	1%	1%	1%	1%	<0.5%	<0.5%	<0.5%	<0.5%
Private - Uninsured	--	1%	4%	2%	3%	3%	3%	3%	3%	2%
Medicaid - Private	--	<0.5%	<0.5%	<0.5%	<0.5%	<0.5%	<0.5%	<0.5%	<0.5%	<0.5%
Medicaid - Uninsured	--	1%	2%	2%	2%	3%	3%	2%	2%	1%
Uninsured-Private	--	1%	<0.5%	4%	3%	3%	3%	3%	3%	3%
Uninsured-Medicaid	--	4%	4%	5%	4%	3%	3%	3%	2%	2%

POSTPARTUM	Delivery Month	+1	+2	+3	+4	+5	+6
Any Transition From Previous Month	--	9%	13%	13%	11%	11%	10%
Private - Medicaid	--	<0.5%	<0.5%	<0.5%	<0.5%	<0.5%	<0.5%
Private - Uninsured	--	2%	3%	3%	3%	3%	3%
Medicaid - Private	--	<0.5%	1%	1%	1%	<0.5%	<0.5%
Medicaid - Uninsured	--	3%	4%	5%	3%	2%	2%
Uninsured-Private	--	2%	2%	3%	3%	3%	3%
Uninsured-Medicaid	--	1%	2%	2%	2%	2%	2%

Notes: Authors' analysis of data for 2005-13 from panels 10-17 of the Medical Expenditure Panel Survey-Household Component. For each month-to-month change period, only women in the survey sample for those months are included (n = 2,726).

Table A1.3 Logistic regression of any insurance lapse before and after delivery

Maternal Characteristic	**Before Delivery** *Nine Months Prior to and Including Delivery (all women)*			**After Delivery** *Six Months After Delivery (women with **any insurance** at delivery)*		
	OR	95% CI	p-value	OR	95% CI	p-value
Age						
<25	Ref		-	Ref		-
25-34	1.21	(0.86 - 1.72)	0.28	0.99	(0.74 - 1.35)	0.98
≥35	1.05	(0.68 - 1.61)	0.83	0.97	(0.64 - 1.49)	0.78
Education						
< High School	Ref		-	Ref		-
High School Diploma	1.06	(0.74 - 1.52)	0.77	0.83	(0.58 - 1.19)	0.31
Post-Secondary Degree	0.82	(0.53 - 1.26)	0.36	0.75	(0.46 - 1.24)	0.60
Race						
White	Ref		-	Ref		-
Black	0.91	(0.64 - 1.30)	0.61	0.71*	(0.49 - 1.02)	0.06
Other	1.30	(0.80 - 2.12)	0.29	0.68	(0.43 - 1.08)	0.16
Hispanic Ethnicity	1.03	(0.70 - 1.51)	0.89	0.86	(0.61 - 1.21)	0.38
Non-English Spoken at Home	1.27	(0.86 - 1.89)	0.23	1.46**	(1.00 - 2.15)	0.05
Married	1.08	(0.76 - 1.53)	0.67	0.71**	(0.51 - 0.98)	0.04
Family Income %FPL						
< 50%	Ref		-	Ref		-
50-99%	0.75	(0.44 - 1.28)	0.29	1.15	(0.76 - 1.76)	0.51
100-185%	0.59**	(0.36 - 0.98)	0.04	1.76***	(1.16 - 2.66)	0.008
186-399%	0.45***	(0.27 - 0.75)	0.002	1.44	(0.88 - 2.34)	0.15
≥400%	0.33***	(0.19 - 0.59)	<0.001	1.32	(0.73 - 2.38)	0.36
Region						
Northeast	Ref		-	Ref		-
Midwest	0.83	(0.53 - 1.31)	0.42	1.44	(0.90 - 2.3)	0.13
South	1.04	(0.67 - 1.63)	0.85	1.87***	(1.16 - 3.01)	0.01
West	0.73	(0.47 - 1.16)	0.18	1.38	(0.86 - 2.21)	0.19
Medicaid/CHIP at Delivery				2.13***	(1.45 - 3.12)	<0.001
n	1,751			1,678		

Notes: Authors' analysis of data for 2005-13 from panels 10-17 of the Medical Expenditure Panel Survey-Household Component. Only women in the survey sample for all calendar months in each period are included. Adjusted odds ratios are from logistic regression using any insurance lapse as the outcome. Each regression adjusted for survey panel. OR is odds ratio. CI is confidence interval. *** p<0.01, ** p<0.05, * p<0.1

A2. Appendix for Chapter 2

Description of Regression Models

Difference-in-Differences Analysis

$$Outcome_{igt} = \beta_0 + \beta_1 Exposure_g + \beta_2 Post_t + \beta_3 Exposure_g * Post_t + \beta_4 UnemploymentRate_t + \beta_x X_i + \Omega\, Month_t + \partial TimeTrend_t + \varepsilon_{igt}$$

(Equation 1)

where *i* indexed birth, *g* group, and *t* date. *Exposure* was an indicator for whether a birth was in the exposure group (maternal age 24-25 years) or the control group (maternal age 27-28 years). Post was an indicator for whether a birth occurred in the period after the implementation of the dependent coverage provision (January 2011 – December 2013). X_i was a vector of control variables (age, race, ethnicity, education, paternal age, marital status, first-live birth, and multiple delivery). *UnemploymentRate* was the age-month specific unemployment rate, from the U.S. Bureau of Labor Statistics. *Month* was the calendar month of delivery. *TimeTrend* was a linear variable measuring the number of months since the beginning of the study period (January 2009). β_3 was the difference-in-differences estimate of the relative change in the outcome from pre- to post-policy in the exposure group relative to the control group.

Difference-in-Differences Analysis with Marital Status Interaction

$$Outcome_{igt} = \beta_0 + \beta_1 Exposure_g + \beta_2 Post_t + \beta_3 Exposure_g * Post_t + \beta_4 Married_i + \beta_5 Post_t * Married_i + \beta_6 Exposure_g * Married_i + \beta_7 Exposure_g * Post_t * Married_i * + \beta_8 UnemploymentRate_t + \beta_x X_i + \Omega\, Month_t + \partial TimeTrend_t + \varepsilon_{igt}$$

(Equation 2)

Married was an indicator of whether a birth was to a married mother. The remaining variables were defined as in Equation 1. β_7 was the estimate of the difference in the difference-in-

differences estimate between married and unmarried women (i.e. the relative difference in the change in the outcome from pre- to post-policy in the exposure and treatment group between married and unmarried women).

Pre-Policy Trend Comparison for the Exposure Group and the Control Group:

$$Outcome_{igt} = \beta_0 + \beta_1 Time_t + \beta_2 Exposure_g * TimeTrend_t + \beta_3 UnemploymentRate_{gt} + \beta_x X_i + \Omega Month_g + \varepsilon_{igt}$$

(Equation 3)

A key assumption in a difference-in-differences analysis is that the trends in the pre-policy period between the two comparison groups are similar. This analysis (presented in Table A2.2) tested the trends prior to the implementation of the dependent coverage provision in the exposure group (24-25 year olds) and the control group (27-28 year olds). Using monthly data limited to the pre-policy period (January 2009 to December 2009), we modeled each outcome as a function of a monthly time trend and an interaction term for the monthly time trend and exposure group status. β_2 identified any diverging pre-policy monthly trend in the exposure group compared to the control group.

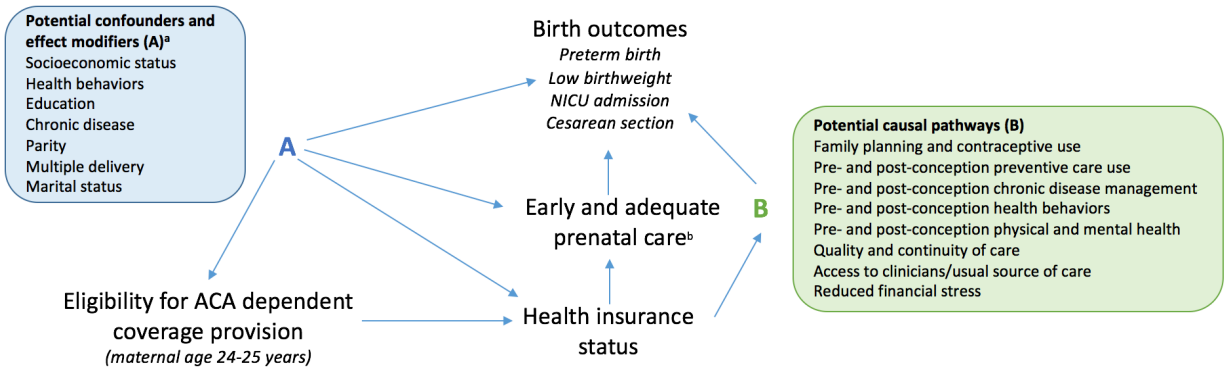


Figure A2.1 Directed acyclic graph

Notes: [a]In a difference-in-differences design, variables are only confounders if they change differentially in the exposure group (ages 24-25) and control group (ages 27-28) over time. For example, education would be a confounder if education level is correlated with birth outcomes or prenatal care and education levels change differentially in the exposure and control group from the pre-policy to post-policy period.

Table A2.1 Pre-period trend differences between the exposure group and control group for different exposure and control group definitions

	Broader Age Bands		Narrower Age Bands (Selected Definition)	
Exposure Group Age Range	19-25 year olds		24-25 year olds	
Control Group Age Range	27-29 year olds		27-28 year olds	
Outcome	Trend Difference	*P* value	Trend Difference	*P* value
Payment for Birth				
Private	0.05	0.02	0.002	0.94
Medicaid	-0.03	0.23	0.01	0.66
Self-Pay	-0.04	<0.001	-0.03	0.06
Prenatal Care				
Early Prenatal Care	0.12	<0.001	0.09	0.01
Adequate Prenatal Care	0.09	<0.001	0.08	0.01
Birth Outcomes				
Cesarean Delivery	0.01	0.73	0.05	0.14
Preterm Birth	0.03	0.08	0.02	0.43
Low Birthweight	0.01	0.42	0.002	0.90
NICU Admission	-0.01	0.75	-0.02	0.17

Notes: Trend difference represents the interaction between exposure group status and a monthly linear time trend calculated from a multivariate regression before the implementation of the dependent coverage provision. Trend differences are given in percentage-points.

Table A2.2 Pre-period trend differences between the exposure group and control group, by marital status

	Overall		Unmarried		Married	
Outcome	Trend Difference	*P* value	Trend Difference	*P* value	Trend Difference	*P* value
Payment for Birth						
Private	0.002	0.94	0.02	0.70	-0.07	0.10
Medicaid	0.01	0.66	0.02	0.77	0.05	0.22
Self-Pay	-0.03	0.06	-0.04	0.18	-0.02	0.40
Prenatal Care						
Early Prenatal Care	0.09	0.01	0.15	0.01	0.01	0.82
Adequate Prenatal Care	0.08	0.01	0.09	0.12	0.04	0.33
Birth Outcomes						
Cesarean Delivery	0.05	0.14	0.0002	1.00	0.07	0.09
Preterm Birth	0.02	0.43	0.04	0.21	0.003	0.91
Low Birthweight	0.002	0.90	-0.01	0.77	0.01	0.53
NICU Admission	-0.02	0.17	-0.06	0.05	0.004	0.86

Notes: Trend difference represents the interaction between exposure group status and a monthly linear time trend calculated from a multivariate regression before the implementation of the dependent coverage provision. Trend differences are given in percentage-points.

Table A2.3 Placebo tests, by marital status

Outcome	Overall		Unmarried		Married	
	Estimated Change in Outcome (Adjusted)	*P* value	Estimated Change in Outcome (Adjusted)	*P* value	Estimated Change in Outcome (Adjusted)	*P* value
Payment Source for Birth						
Private	0.11	0.61	0.32	0.34	-0.50	0.09
Medicaid	0.02	0.94	0.00	0.99	0.30	0.30
Self-Pay	-0.20	0.08	-0.37	0.06	-0.04	0.79
Prenatal Care						
Early Prenatal Care	0.61	0.01	0.93	0.03	0.11	0.71
Adequate Prenatal Care	0.48	0.04	0.28	0.49	0.30	0.28
Birth Outcomes						
Cesarean Delivery	0.15	0.54	-0.21	0.61	0.33	0.28
Preterm Birth	0.04	0.78	0.19	0.45	-0.01	0.97
Low Birthweight	-0.03	0.80	-0.21	0.38	0.12	0.43
NICU Admission	-0.11	0.39	-0.47	0.04	0.17	0.30

Notes: Estimated change in outcome represents the interaction between Exposure group status and a policy indicator variable indicating implementation of a placebo policy assumed to be implemented six months prior to 2010 (July 2009). Estimates are calculated using multivariate regressions in the period before the dependent coverage provision that are otherwise identical to those used to produce the main results. Estimates are given in percentage-points.

Table A2.4 Sensitivity analysis with clustered standard errors

Outcome	Estimated Change in Outcome (Adjusted)	Robust Standard Errors		Cluster-Robust Standard Errors	
		95% Confidence Interval	*P* Value	95% Confidence Interval	*P* Value
Payment for Birth					
Private	1.9	1.6, 2.1	<0.001	1.3, 2.4	<0.001
Medicaid	-1.4	-1.7, -1.2	<0.001	-1.8, -1.1	<0.001
Self-Pay	-0.3	-0.4, -0.1	<0.001	-0.4, -0.1	<0.001
Prenatal Care					
Early Prenatal Care	1.0	0.7, 1.2	<0.001	0.6, 1.4	<0.001
Adequate Prenatal Care	0.4	0.2, 0.6	<0.001	0.1, 0.7	0.02
Birth Outcomes					
Cesarean Delivery	0.005	-0.3, 0.3	0.97	-0.3, 0.3	0.97
Preterm Birth	-0.2	-0.3, -0.03	0.02	-0.3, -0.01	0.04
Low Birthweight	-0.01	-0.1, 0.1	0.91	-0.1, 0.1	0.91
NICU Admission	-0.1	-0.3, -0.3	0.11	-0.3, 0.05	0.17

Notes: Estimates are given in percentage points unless otherwise noted. Cluster-robust standard errors clustered at the age group-month level

Table A2.6 Sensitivity analysis adjusting for payment source

Outcome	Main Regression			Adjusted for Payment for Birth		
	Estimated Change in Outcome (Adjusted)	95% Confidence interval	*P* Value	Estimated Change in Outcome (Adjusted)	95% Confidence interval	*P* Value
Prenatal Care						
Early Prenatal Care	1.0	0.7, 1.2	<0.001	0.7	0.5, 1.0	<0.001
Adequate Prenatal Care	0.4	0.2, 0.6	<0.001	0.2	-0.04, 0.4	0.09
Birth Outcomes						
Cesarean Delivery	0.005	-0.3, 0.3	0.97	0.07	-0.2, 0.3	0.61
Preterm Birth	-0.2	-0.3, -0.03	0.02	-0.1	-0.3, 0.01	0.06
Low Birthweight	-0.01	-0.1, 0.1	0.91	0.03	-0.1, 0.2	0.67
NICU Admission	-0.1	-0.3, -0.3	0.11	-0.07	-0.2, 0.1	0.31

Notes: Estimates are given in percentage-points. Exposure group includes births to 24-25 year-old women and control group includes births to 27-28 year-old women. All models analyze data from 2009-2013 excluding 2010 as the policy-implementation period and use robust standard errors (N=2,930,197). Adjusted models are adjusted for payment source for birth; month of delivery; a monthly linear time trend; maternal marital status, age, race, ethnicity, and education; whether the birth was a woman's first live birth; multiple delivery; paternal age; and monthly unemployment rates

A3. Appendix for Chapter 3

Description of Data Generation Process

The simulation generates data for units (indexed by i) in two groups ($z = 0$ for control, $z = 1$ for treatment) at T time points, $\boldsymbol{y}_{iz} = (y_{i1z}, \dots, y_{iTz})'$. The unit observations are drawn from a multivariate normal distribution, $\boldsymbol{y}_{iz} \sim N(\boldsymbol{\mu}_z, \Sigma_{\mathrm{Y}})$, where $\boldsymbol{\mu}_z = (\mu_{1z}, \dots, \mu_{8z})'$ is the corresponding group mean vector. Across scenarios, we fix the control group outcome mean at $\mu_{t0} = 0$ and vary the treatment group mean to create differences in outcome and level between the groups.

The TxT variance-covariance matrix Σ_{Y} is the same in the treatment and control groups and has an auto-regressive of order 1, AR(1), structure. The matrix is completely determined by two parameters. First, a variance parameter controls the variance of observations at each time point, and we assume this is constant over time and fixed across scenarios $Var(y_{itz}) = \sigma_Y^2 = 1$. Second, a correlation parameter ρ_Y governs the relationship between outcomes at different time points, which decays exponentially with increasing time separation, $Cov(y_{itz},\ y_{it\prime z}) = \sigma_Y^2 \rho_Y^{|t-t\prime|}$.
For these simulations, we vary the parameters shown in Table A1 and the outcome correlation over $\rho_Y = \{0, .1, .3, .5, .7, .9, .99\}$.

For some scenarios, we also generate a time-varying covariate, $\boldsymbol{x}_{iz} = (x_{i1z}, \dots, x_{iTz})'$. The unit observations are drawn from a multivariate normal distribution, $\boldsymbol{x}_{iz} \sim N(\boldsymbol{\gamma}_z, \Sigma_{\mathrm{x}})$, where $\boldsymbol{\gamma}_z = (\gamma_{1z}, \dots, \gamma_{Tz})'$ is the corresponding group mean vector. Across scenarios, we fix the control group

covariate mean at $\gamma_{t0} = 0$ and vary the treatment group mean to create the differences in covariate level between the groups.

The TxT variance-covariance matrix Σ_X is analogous to the outcome matrix: it is the same in both groups and has AR(1) structure determined by two parameters, a constant variance parameter at each time point $Var(x_{itz}) = \sigma_X^2 = 1$ and a correlation parameter ρ_X such that $Cov(y_{itz},\ y_{it\prime z}) = \ \sigma_X^2 \rho_X^{|t-t\prime|}$.

To generate the outcome observations in these covariate-dependent scenarios, we draw outcomes from a multivariate normal distribution, $\boldsymbol{y}_{iz} \sim N(\boldsymbol{\alpha}_z + \beta \boldsymbol{x}_{iz}, \Sigma_Y)$. We fix $\boldsymbol{\alpha}_0 = \mathbf{0_T}$ and $\boldsymbol{\alpha}_1 = \mathbf{1_T}$ (T-vectors of zeros and ones, respectively) and vary β across scenarios to create weak, moderate, and strong relationships between the covariate and the outcome. For these, we fix $\rho_Y = 0.6$ and vary the parameters shown in Table A2 and the covariate correlation over $\rho_X = \{0, .1, .3, .5, .7, .9, .99\}$.

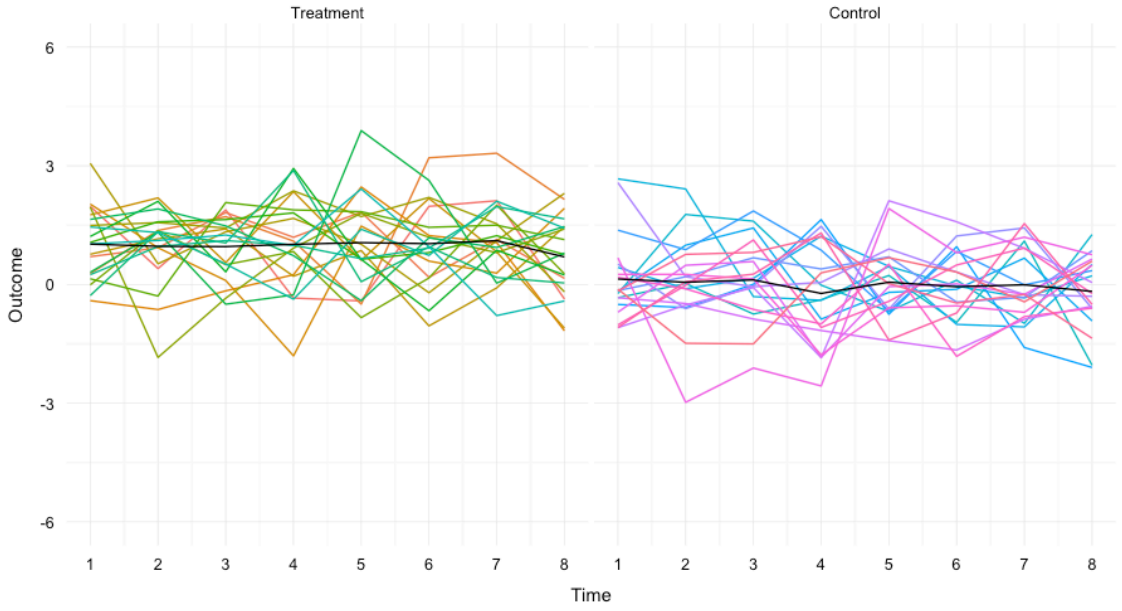
Table A3.1 Parameters of simulations with no covariate

Treatment assignment mechanism	**Time periods**	**Treatment group outcome mean**	**Results Figure**
Random	$t = 1, \dots, 8$	$\mu_{t1} = 0$	1a,b,c
Correlated with outcome level	$t = 1, \dots, 8$	$\mu_{t1} = \{1,2\}$	1a,b,c
Correlated with outcome trend	$t = 1, \dots, 8$	$\mu_{t1} = \begin{Bmatrix} -.025 + 0.01t, \\ -.125 + 0.05t, \\ -.25 + 0.10t \end{Bmatrix}$	3a,b,c

Table A3.2 Parameters of simulations with covariate

Treatment assignment mechanism	**Time periods**	**Treatment group covariate mean**	**Covariate-outcome relationship**	**Results figure**
Correlated with covariate level	$t = 1, \dots, 8$	$\gamma_{t1} = \{1,2\}$	$\beta = \{.1, .4, .8\}$	2a,b

(a) Low serial correlation, $\rho_Y = 0.10$



(b) High serial correlation, $\rho_Y = 0.95$

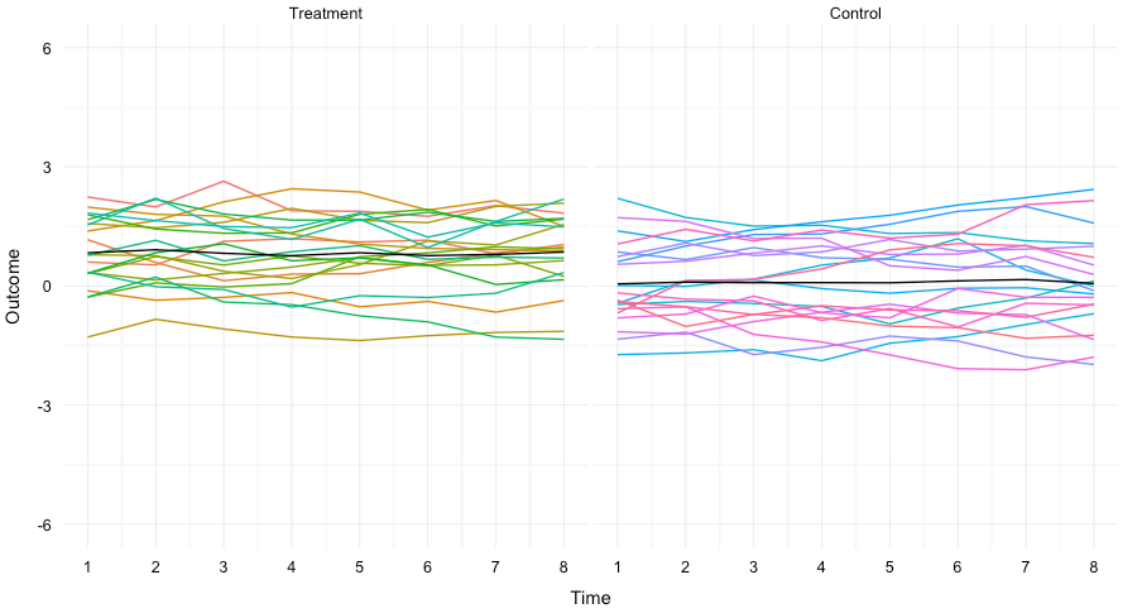


Figure A3.1 Spaghetti plots of the outcome

Notes: The serial correlation of the outcome refers to the autoregressive parameter of the AR(1) correlation structure. Data are from ten randomly selected units from treatment and control where the pre-period difference in level between treatment and control is 1 standard deviation of the outcome and there is no pre-period difference in trend. The black horizontal line on each plot represents the mean of the displayed units.

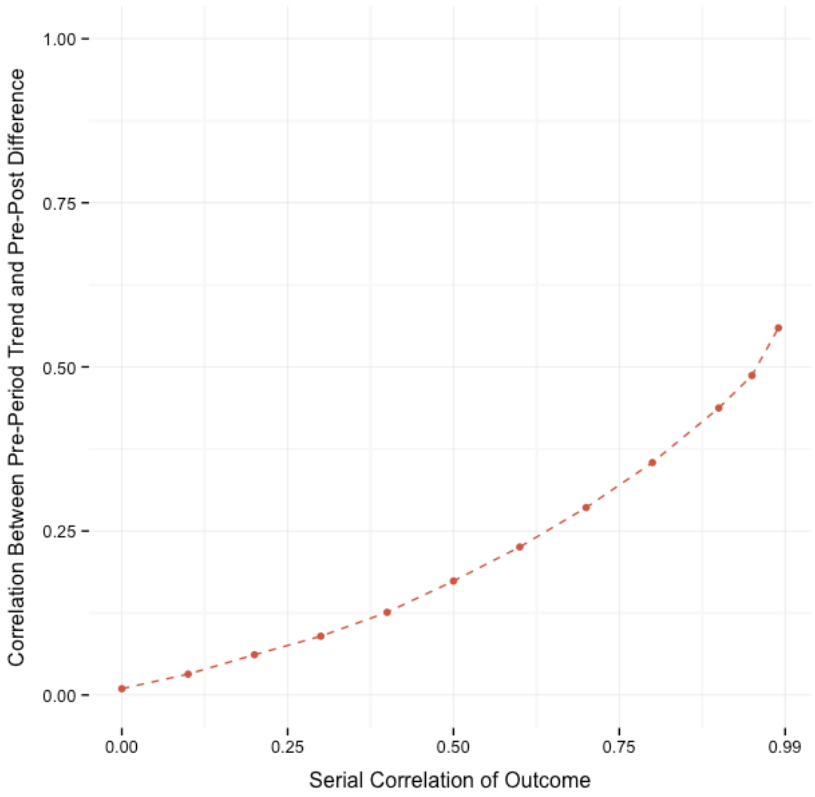
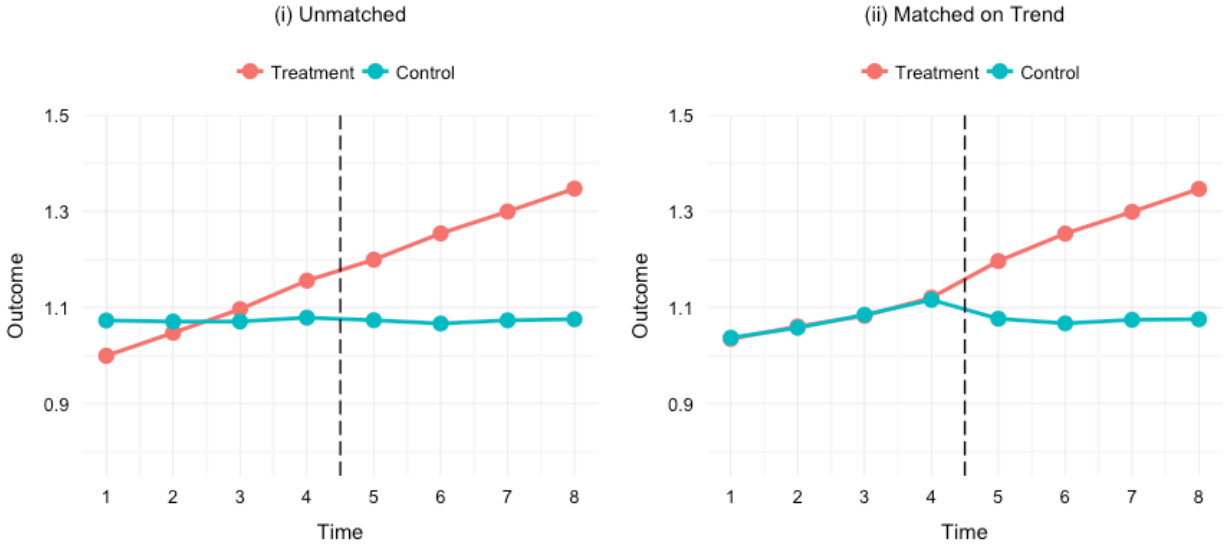


Figure A3.2 Correlation between pre-period trend and pre-post difference

Notes: The serial correlation of the outcome refers to the autoregressive parameter of the AR(1) correlation structure. Data are based on the mean correlation across 100 simulation iterations for each scenario. Calculations are for unmatched samples and the multiple observation pre-post design where the pre-period difference in trend between treatment and control is 0.05 standard deviations of the outcome.

(a) Low serial correlation, $\rho = 0.10$



(b) High serial correlation, $\rho = 0.95$

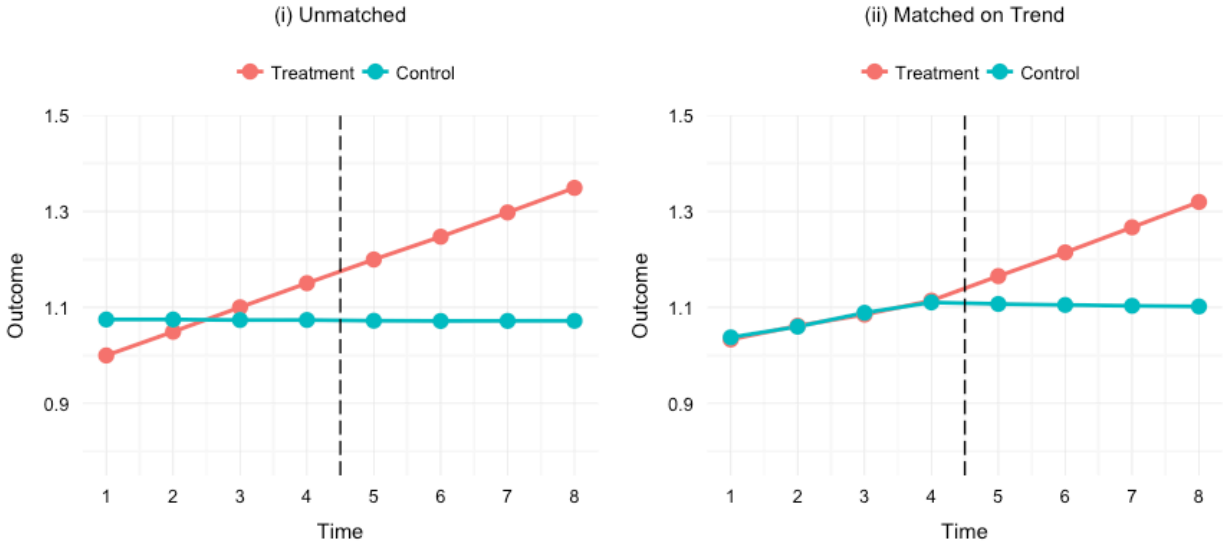


Figure A3.3 Plots of the mean outcome for unmatched samples and samples matched on trend where group-level treatment assignment is correlated with pre-period trend

Notes: The serial correlation of the outcome refers to the autoregressive parameter of the AR(1) correlation structure. Data are based on the mean outcome for each measurement occasion for the treatment and control group across 100 simulation iterations for each scenario. Calculations are for unmatched and matched samples where there is no pre-period difference in level and the pre-period difference in trend between treatment and control is 0.05 standard deviations of the outcome.

Table A3.3 Year-to-year correlation for a selection of health-related variables

Individual-level variable	**Correlation**
Body Mass Index	0.86
Household Income (households <100% FPL in base year)	0.30
Household Income (households >100% FPL in base year)	0.71
Total Health Expenditures	0.40
Total Prescription Drug Expenditures	0.51
Total Outpatient Expenditures	0.23
Total ER Expenditures	0.05
Total Hospital Expenditures	0.17
Total Hospital Nights	0.16
HSA-level variable	**Correlation**
Percent of Medicare patients readmitted within 30 days of discharge following medical admission	0.37
Inpatient hip replacement per 1000 Medicare enrollees	0.71
Percent of Medicare deaths occurring in the hospital	0.79
Percent of diabetic Medicare enrollees receiving HbA1c testing	0.86

Notes: Authors' calculations from the Medical Expenditure Panel Survey Panel 17 and Dartmouth Health Atlas 2009-14.