

ABSTRACT

Title of Dissertation:

THREE ESSAYS ON MARYLAND'S
GLOBAL BUDGET REVENUE PROGRAM
AND HOSPITAL-BASED NEONATAL CARE

Liyang Xie, Doctor of Philosophy, 2020

Dissertation directed by:

Dr. Michel Boudreaux, Assistant Professor,
Department of Health Policy and Management

Health care spending is a major concern in the United States. State and federal governments have been engaged in a number of health care system reform initiatives designed to contain costs by regulating both price and quantity. Comprehensive evaluations of these initiatives are crucial for policymakers reshaping and expanding reforms.

This dissertation evaluates the impact of Maryland's Global Budget Revenue (GBR) program, one of the most innovative statewide hospital payment reforms, on birth-related hospital utilization. The GBR program was designed to provide incentives for hospitals to reduce high-cost services and substitute them for lower-cost population health investments. This is largely accomplished by capping annual budgets. This dissertation evaluated the effects of GBR on high-cost neonatal services,

especially the neonatal intensive care unit (NICU). I examine heterogeneous treatment effects with respect to observable clinical needs and financial incentives.

In Chapter One, I provide an overview of Maryland's GBR program and introduce the conceptual framework. In Chapter Two, I examine the impact of GBR on NICU admissions and infant mortality. I explore the heterogeneity of treatment effects by infant health risk. Chapter Three expands the analysis to broader birth-related hospital services by investigating the impact of GBR on length of stay (LOS), the total cost of care, and utilization of specific high-cost services. Chapter Four departs from GBR and examines NICU utilization related to another critical source of financial incentive – health insurance type. Chapter Five concludes the dissertation.

I find that Maryland's GBR program led to a substantial decline in NICU admissions, which was mainly driven by the decrease in admissions of relatively healthy infants, and there are no changes in the infant or neonatal mortality rate. The GBR program is also associated with declines in LOS and high-cost services used for infants. Finally, I observe that infant, maternal, and state characteristics explain the variations in NICU care across insurance type for high-risk infants but not for relatively low-risk infants. My findings provide positive evidence on implementing global hospital budget programs and shed light on the economic incentives affecting NICU care.

THREE ESSAYS ON MARYLAND'S GLOBAL BUDGET REVENUE PROGRAM
AND HOSPITAL-BASED NEONATAL CARE

by

Liyang Xie

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:

Professor Michel Boudreaux, Chair
Professor Jie Chen
Professor Eva DuGoff
Professor Luisa Franzini
Professor Xin He

© Copyright by
Liyang Xie
2020

Dedication

This dissertation is dedicated to the health care workers all around the world on the frontlines of fighting Covid-19. Your sacrifice, fearlessness, and efforts inspire me to devote myself to making the healthcare system better.

Acknowledgements

First and foremost, I want to thank my dissertation committee – Michel Boudreaux, Luisa Franzini, Jie Chen, Eva DuGoff, and Xin He – for your guidance and support.

I am also grateful for the financial support from the University of Maryland Department of Health Policy and Management, Graduate School, School of Public Health, and Maryland Population Research Center, which afforded me the essential data for my dissertation and the opportunities of attending conferences.

I thank David Goodman and Janice Wilson for answering my pile of questions about NICU and neonatal care.

I also want to thank my friends Jun, Nianyang, Jingkai, Deanna, Daniel, Ellesse, Guanyi, and Xuan for your friendship, encouragement, and insight.

To Michel, you are the best advisor ever. I have been so lucky to be your research assistant over the past five years. You spent countless hours helping me in every possible way to become a better researcher and I am constantly inspired by your passion, focus, and dedication to research in improving healthcare and defeating poverty. I cannot thank you enough.

Finally, to my beloved family: Fang, Ping, Guizhen, Xiao, and Qimei, thank you for always supporting me to pursue my dreams and teaching me how to cook delicious Chinese dishes.

Table of Contents

| | |
|---|-----|
| Dedication | ii |
| Acknowledgements | iii |
| Table of Contents | iv |
| List of Tables | vii |
| List of Figures | ix |
| List of Abbreviations | x |
| Chapter 1 : Introduction | 1 |
| 1.1 Background | 2 |
| 1.1.1 Maryland’s All-Payer Rate-setting System | 2 |
| 1.1.2 Settings of Maryland’s Global Budget Revenue Program..... | 3 |
| 1.1.3 Previous Findings on the Effects of Maryland’s Global Budget Revenue Program on Hospital Utilization | 6 |
| 1.2 Conceptual Framework | 10 |
| 1.3 A Preview of Chapters Two, Three, and Four | 14 |
| Chapter 2 : Changes in NICU Admissions after Maryland’s Global Budget Revenue Program..... | 16 |
| 2.1 Introduction..... | 16 |
| 2.2 Methods..... | 18 |
| 2.2.1 Study Design..... | 18 |
| 2.2.2 Data | 19 |
| 2.2.3 Study Variables | 20 |
| 2.2.4 Statistical Analysis..... | 21 |
| 2.3 Results..... | 23 |
| 2.3.1 Newborn Characteristics..... | 23 |
| 2.3.2 NICU Utilization..... | 25 |
| 2.3.3 Infant Mortality | 29 |
| 2.3.4 Sensitivity Analyses..... | 33 |
| 2.4 Discussion | 33 |
| 2.5 Conclusion | 36 |
| Chapter 3 : The Impact of Maryland’s Global Budget Revenue Program on Birth- Related Hospitalization: Length of Stay, Cost, and Service Utilization..... | 38 |
| 3.1 Introduction..... | 38 |
| 3.2 Methods..... | 39 |
| 3.2.1 Study Design..... | 39 |
| 3.2.2 Data | 40 |
| 3.2.3 Study Variables | 41 |
| 3.2.4 Statistical Analysis..... | 43 |

| | |
|---|-----|
| 3.3 Results..... | 45 |
| 3.3.1 Descriptive Statistics..... | 45 |
| 3.3.2 Pre-trends..... | 48 |
| 3.3.3 Effect of GBR on LOS..... | 54 |
| 3.3.4 Effect of GBR on Cost of Birth..... | 56 |
| 3.3.5 Effect of GBR on Service Utilization..... | 58 |
| 3.3.6 Cost Estimation by Hospital Department..... | 60 |
| 3.3.7 Sensitivity Analyses..... | 63 |
| 3.4 Discussion..... | 64 |
| 3.5 Conclusion..... | 65 |
| Chapter 4 : Variation in NICU Admissions across Insurance Type..... | 66 |
| 4.1 Introduction..... | 66 |
| 4.2 Background..... | 68 |
| 4.2.1 The Development of NICU..... | 68 |
| 4.2.2 The Current Status of NICU..... | 71 |
| 4.3 Methods..... | 73 |
| 4.3.1 Data..... | 73 |
| 4.3.2 Study Variables..... | 74 |
| 4.3.3 Statistical Analysis..... | 75 |
| 4.4 Results..... | 77 |
| 4.4.1 Descriptive Statistics..... | 77 |
| 4.4.2 Variation in NICU Admissions across Insurance Type..... | 80 |
| 4.5 Discussion..... | 83 |
| 4.6 Conclusion..... | 84 |
| Chapter 5 : Conclusion..... | 86 |
| Appendix A..... | 87 |
| A1. Coding of Comparison States..... | 87 |
| A2. Variable Definitions..... | 88 |
| A3. Pre-trends Test..... | 89 |
| A4. Effects by Year of Implementation..... | 91 |
| A5. Robustness to Alternative Comparison States..... | 91 |
| A6. Robustness to Alternative Implementation Timing..... | 91 |
| A7. Robustness to Alternative Model Specifications..... | 92 |
| A8. Robustness to Alternative Samples..... | 93 |
| A9. The Impact of GBR on Infant Health..... | 94 |
| A10. Measurement Error in NICU Admissions..... | 95 |
| A11. Synthetic Control Methods Estimates..... | 97 |
| A12. Dynamic Difference-in-Differences Estimates..... | 99 |
| A13. Models with State-specific Linear Trends..... | 99 |
| Appendix B..... | 119 |
| B1. Data Sources..... | 119 |
| B2. Model Fit Comparison..... | 123 |
| B3. Sensitivity Analysis using 2014 as the Implementation Year..... | 123 |

| | |
|--------------------|-----|
| Appendix C | 129 |
| Bibliography | 131 |

List of Tables

| | |
|---|-----|
| Table 1.1: Model Agreement | 4 |
| Table 2.1: Characteristics of the Study Population at Baseline, 2011-2014..... | 24 |
| Table 2.2: Effect of GBR on NICU Admissions, Overall and By Birthweight, 2011-2017..... | 30 |
| Table 2.3: Effect of GBR on NICU Admissions, By Gestational Age, 2011-2017 ... | 31 |
| Table 2.4: Effect of GBR on Infant/Neonatal Mortality Rates, 2011-2017..... | 32 |
| Table 3.1: Characteristics of Infants Before and After GBR, 2011-2016 | 46 |
| Table 3.2: Outcomes of Infants in Maryland at Baseline, 2011-2014..... | 48 |
| Table 3.3: Differential Pre-Trends Test of Infants, 2011-2014 | 54 |
| Table 3.4: Effect of GBR on Length of Stay, 2011-2016..... | 55 |
| Table 3.5: Effect of GBR on Total Cost of Birth, 2011-2016 | 57 |
| Table 3.6: Effect of GBR on Services Utilization, 2011-2016..... | 59 |
| Table 3.7: Aggregate Inpatient Hospitalization Costs Attributable to Each Department, 2011-2016 | 60 |
| Table 4.1: Descriptive Statistics by Insurance Type..... | 77 |
| Table 4.2: NICU Admission Rates by Insurance Type and by Birthweight..... | 80 |
| Table 4.3: The Association between Insurance Type and NICU Admissions..... | 81 |
| Table 4.4: The Association between Insurance Type and NICU Admissions, by Birthweight | 82 |
| Table A.1: Comparison States | 100 |
| Table A.2: Differential Pre-Trends Test, 2011-2014..... | 101 |
| Table A.3: Difference-in-Differences Estimates of the Effect of GBR on NICU Admission by Years of Implementation, 2011-2017 | 102 |
| Table A.4: Difference-in-Differences Estimates of the Effect of GBR on NICU Admission using Alternative Control States, 2011-2017 | 103 |
| Table A.5: Difference-in-Differences Estimates of the Effect of GBR on NICU Admission using Alternative Post-Implementation Time, 2011-2017 | 104 |
| Table A.6: Difference-in-Differences Estimates of the Effect of GBR on NICU Admission using Alternative Control States, 2011-2017 | 105 |
| Table A.7: Difference-in-Differences Estimates of the Effect of GBR on NICU Admission using Alternative Samples, 2011-2017..... | 106 |

| | |
|--|-----|
| Table A.8: Difference-in-Difference Estimates of the Effect of GBR on Infant Health Measures, 2011-2017 | 107 |
| Table A.9: State Weights for the Synthetic Maryland | 108 |
| Table A.10: Effect of GBR on NICU Admissions from SCM, 2011-2017 | 108 |
| Table A.11: Comparing Estimates from DID versus SCM | 109 |
| Table A.12: The Difference-in-Differences Estimates including State-specific Linear Trends, 2011-2017 | 109 |
| Table A.13: Characteristics of Mothers Before and After GBR, 2011-2017 | 110 |
| Table B.1: Sources of Covariates..... | 119 |
| Table B.2: Sources of Outcome Variables..... | 120 |
| Table B.3: Sources of Congenital Anomalies..... | 121 |
| Table B.4: Model Fit Comparison, Poisson Model versus Negative Binomial Model | 123 |
| Table B.5: Effect of GBR on Length of Stay (2014 as the Implementation Year), 2011-2016 | 124 |
| Table B.6: Effect of GBR on Length of Stay by Birthweight (2014 as the Implementation Year), 2011-2016..... | 125 |
| Table B.7: Effect of GBR on Total Cost of Birth (2014 as the Implementation Year), 2011-2016 | 126 |
| Table B.8: Effect of GBR on Total Cost of Birth by Birthweight (2014 as the Implementation Year), 2011-2016..... | 127 |
| Table B.9: Effect of GBR on Services Utilization (2014 as the Implementation Year), 2011-2016 | 128 |
| Table C.1: Percent of Births by State and Insurance Type, 2016-2017 | 129 |

List of Figures

| | |
|---|-----|
| Figure 2.1: Unadjusted Trends in NICU Admission: Maryland versus the Comparison States | 26 |
| Figure 3.1: Unadjusted Trends of Cost and LOS for All Infants: Maryland versus Comparison States | 49 |
| Figure 3.2: Unadjusted Trends of LOS for Infants by Gestational Age: Maryland versus Comparison States | 50 |
| Figure 3.3: Unadjusted Trends of Cost for Infants by Gestational Age: Maryland versus Comparison States | 51 |
| Figure 3.4: Unadjusted Trends of Services Utilization for Infants: Maryland versus Comparison States | 52 |
| Figure 3.5: Prevalence of Service Utilization by the Hospital Department, 2011-2016 | 62 |
| Figure 3.6: Distribution of Total Cost of Birth by Department for Infants, 2011-2016 | 63 |
| Figure A.1: Unadjusted Trends of Infant Mortality Rate, 2011-2017 | 112 |
| Figure A.2: Unadjusted Trends of Neonatal Mortality Rate, 2011-2017 | 113 |
| Figure A.3: Unadjusted Trends of Other Health Outcomes, 2011-2017 | 114 |
| Figure A.4: NICU Admission of Vital Statistics and Hospital Discharge Data in Maryland, 2011-2017 | 115 |
| Figure A.5: Trends in NICU Admission Rate: Maryland versus Synthetic Maryland | 116 |
| Figure A.6: Ratio of Post/Pre-GBR Mean Squared Prediction Error, Overall | 117 |
| Figure A.7: Dynamic Difference-in-Differences | 117 |

List of Abbreviations

| | |
|--|-----------------|
| AAP: American Academy of Pediatrics | passim |
| ACA: Affordable Care Act | 18, 25, 87, 111 |
| CDC: Centers for Disease Control and Prevention..... | passim |
| CMS: Centers for Medicare & Medicaid Services | passim |
| CT: computed tomography | passim |
| DC: District of Columbia..... | passim |
| ED: Emergency Department | 7, 9, 16, 34 |
| GBR: Global Budget Revenue..... | passim |
| GDP: Gross Domestic Product | 1 |
| HAC: hospital-acquired condition | 4 |
| HBW: high birthweight..... | passim |
| HCUP: Healthcare Cost and Utilization Project..... | passim |
| HSCRC: Health Services Cost Review Commission | passim |
| IPPS: Inpatient Prospective Payment System..... | 2, 5, 41 |
| IRB: Institutional Review Board | 23, 76 |
| LOS: length of stay | passim |
| MLBW: moderately-low birthweight | passim |
| MRI: magnetic resonance imaging | passim |
| MSPE: mean squared prediction error | 98, 99 |
| NBW: normal birthweight | passim |
| NICU: neonatal intensive care unit..... | passim |
| PPC: potentially preventable condition | 9 |
| SGA: small for gestational age | 75, 77, 94, 107 |
| SID: State Inpatient Databases..... | passim |
| TPR: Total Patient Revenue..... | 3, 7 |
| VLBW: very-low birthweight..... | passim |
| VS: Vital Statistics..... | 73, 96, 97 |

Chapter 1: Introduction

Rising health care spending is a major concern facing the United States. In 2016, health care spending accounted for 17.8% of GDP, far outpacing spending in other developed countries (Papanicolas, Woskie, and Jha 2018). This spending is driven by high prices that result from new medical technologies, non-competitive market structures, and the excess utilization of services resulting from fee-for-service payment systems (Dieleman et al. 2017; Schroeder and Frist 2013; Cutler and McClellan 2001). State and federal governments have been engaging in a number of health care system reform initiatives that could help contain costs by regulating both price and quantity (Center for Medicare & Medicaid Innovation 2018). Comprehensive evaluations of these initiatives are crucial for policymakers reshaping future large-scale payment reforms.

This dissertation evaluates the impact of Maryland's Global Budget Revenue (GBR) program, one of the most innovative statewide hospital payment reforms since 2014, focusing on birth-related hospital utilization that accounts for approximately 10% of overall U.S. inpatient admissions each year. The GBR program was designed to provide incentives for hospitals to reduce high-cost services and substitute them for lower-cost population health investments. This dissertation focuses on high-cost neonatal services and examines heterogeneous effects with respect to observable clinical needs and financial incentives.

In this introductory chapter, I provide background information on Maryland's hospital payment system, introduce detailed settings of the GBR program, and discuss

current findings on evaluating GBR. Next, I introduce the conceptual frameworks. I conclude this chapter with a brief preview of Chapters Two, Three, and Four.

1.1 Background

1.1.1 Maryland's All-Payer Rate-setting System

Maryland has operated a nationally unique “All-payer” hospital rate-setting system since 1977. The system was running under a Medicare Waiver (codified in Section 1814(b) of the Social Security Act), which allows Maryland to rate Medicare services separately from the Inpatient Prospective Payment System (IPPS) and Outpatient Prospective Payment System (OPPS) (Cohen 2005). The system allowed Maryland to pay higher Medicare fees, and then it set all payers to the same unit price (fees). Under this waiver, the price of a given service was equated across all payers, but not necessarily across all hospitals. Hospital rates were based on hospitals’ historical costs and set to equal among payers (Giuriceo et al. 2016). This rate-setting process was regulated by an independent agency: the Health Services Cost Review Commission (HSCRC).

The all-payer rate-setting system was expected to constrain hospital costs, guarantee access, improve equity and fairness of hospital financing, and keep the system accountable to the public (Kastor and Adashi 2011). As a result, the system led to lower hospital prices in Maryland compared with other states. From 1977 to 2009, Maryland hospital’s cost per admission changed from 26 percent above the national average to 2.5 percent below the national average (Kastor and Adashi 2011). However, the all-payer rate setting system was blamed for giving hospitals an incentive to increase the volume of admissions (Murray 2009). As a result, Maryland’s Medicare rates remained higher than

the national average (Pope 2019) and per capita Medicare total spending ranked among the highest in the US heading into 2010 (KFF 2014). Maryland hospital discharges per 1,000 Medicare enrollees were also higher than the national average (Dartmouth Atlas Project 2019). Based on this evidence, it appears that hospitals did compensate for lower average unit prices by increasing volumes.

1.1.2 Settings of Maryland’s Global Budget Revenue Program

To restraint total spending (rather than per capita cost only) and to further improve quality, Maryland initiated a new payment model in 2014 known as the "Global Budget Revenue (GBR) program" or “All-payer Model”, which capped the annual growth of the total hospital spending to 3.58%, the state's 10-year compound historic economic growth rate (CMS 2018). This was a bold step aiming at solving current problems while retaining its unique rate-setting system.

A pilot model was first tested by CMS in partnership with the State of Maryland before the formal implementation. This pilot model was known as the “Total Patient Revenue (TPR) system” and was implemented in 8 rural acute-care hospitals in Maryland from July 1, 2010 to June 30, 2013.

Full GBR implementation took effect on January 1, 2014 that expanded to all 46 Maryland acute-care hospitals. By July 2014, all of the hospitals in the state successfully transited to operate under this global budget system. The “Model Agreement” requires that key requirements must be met, as shown in Table 1.1 (HSCRC 2018b). The agreement contains four general requirements and three that are particular to Medicare.

Table 1.1: Model Agreement

| General Requirement | Medicare Requirements |
|---|--|
| All-payer per capita total hospital revenue growth must be limited to 3.58 percent per year | Five-year Medicare per beneficiary total hospital cost savings must equal or exceed \$330 million |
| The rate of hospital-acquired conditions (HACs) must be reduced by 30 percent | The aggregate Medicare 30-day all-cause readmission rate must be reduced to at or below the national average |
| Hospital payment must transition away from volume-based payments | Total Medicare spending per beneficiary growth must fall below certain national growth rates |
| Maryland must submit a plan at the end of 2016 to move beyond hospitals and limit the growth in total hospital and non-hospital Medicare spending | |

The program sets a fixed budget for each hospital on inpatient, outpatient, and emergency department services based on its historical volume and limits all-payer per capita total hospital revenue growth rate to be within 3.58 percent each year over a 5-year period. The budget was also based on patient mix and services and adjusted annually to take into account other uncertainties such as inflation, changes in the community, service levels, or shifting of services to other settings. Each hospital knows its total revenue of that year in advance.

The key feature of GBR is that hospital revenues are expected to conform closely to the global budgets. Penalties are applied if revenues vary from the allowed global budget beyond a narrow 0.5 percent corridor. Hospitals get penalties for the portion of over or under 0.5 percent of the budget. The charges that are under the budget within 0.5% will be credited into the following year's budget, and overage amounts will be debited from the following year's budget. Hospitals can also obtain additional payments by meeting quality benchmarks, which are based on clinical processes of care measures,

patient experience measures, and mortality. Considering that the actual utilization is unlikely to perfectly match the projected utilization on which the global budget is based, to compensate for some deviation, the program also gives some freedom for hospitals to adjust their prices within $\pm 5\%$ (price adjustments larger than ($\pm 5\%$) were subject to the review and approval of the HSCRC).

Given these features, GBR changed the incentive of hospitals to generate earnings created from the fee-for-service system. It moved hospital financing from fee-for-service, where hospitals get paid for more “heads in beds” and keeping them for a long time, to paying for value and outcomes while not bringing patients in unnecessarily. Before GBR, hospitals increased the volume of patients to earn more revenue since the prices were regulated. Under GBR, the expected revenue is stabilized. In order to maximize the revenue, hospitals need to keep their expenditure aligned with the budget by limiting volumes and managing high-cost services. The features also give hospitals a strong incentive to adjust their prices to reach their global budgets when they have lower-than-expected volumes, which may also increase operating margins per volume (Giuriceo et al. 2016, 2018).

In addition, Maryland’s all-payer system created incentives that were different from the IPPS. Specifically, Maryland’s system before GBR, which paid for each unit of service provided, incentivized hospitals to increase not only overall volume (admissions and readmissions, i.e., extensive margins) but also the number of services for each admission/readmission (i.e., intensive margins). Unlike Maryland’s system, other hospitals under the IPPS had incentives to increase overall volume (outside of readmissions penalties), while limited the incentive to increase within-case intensity (i.e.,

intensive margins) such as testing, procedures, LOS, and units of services. Although hospitals were paid under GBR by using previous rates from a universe of 51 revenue centers and the units of service provided (Giuriceo et al. 2016), by fixing the hospitals' total revenue, GBR created new incentives for hospitals to reduce both overall volume (which is not provided by the IPPS), i.e., admissions/readmission, and the number of services for each admission/readmission (similar to the IPPS), to fully receive the expected revenue and increase profit.

The global budget rate-setting applies to all the Maryland residents and most out-of-state residents who received services in Maryland hospitals, with an exception (before 2017) for out-of-state residents visiting Johns Hopkins Hospital and its affiliates in Maryland. Medicare beneficiaries who received out-of-state hospital care are also part of the hospitals' global budgets (Berenson 2015). More details on the GBR program can be found in RTI's annual reports (Giuriceo et al. 2016).

1.1.3 Previous Findings on the Effects of Maryland's Global Budget Revenue Program on Hospital Utilization

By July 2014, all 36 general acute-care urban hospitals shifted 95% (the remaining 5% excluded from the global budget was the revenue for out-of-state patients) of their revenue into global budgets (with the exception of Holy Cross Germantown in October 2014) (Giuriceo et al. 2016). Evaluations for the GBR program, including all Maryland hospitals, were reported annually by RTI International starting in 2015 (Giuriceo et al. 2016, 2017, 2018, 2019). These annual reports cover a wide range of content, including hospital service utilization, hospital service mix, market dynamics, quality of care, and spillover effects to services not subject to the global budget (e.g.,

whether service provided in hospital outpatient settings shifted to nonregulated settings). The first annual report was only conducted on Medicare beneficiaries, and later reports expanded to commercial plan members and Medicaid beneficiaries. Aside from the RTI's reports, there were a few pieces of literature evaluating the GBR for both pilot and full implementation periods on hospital service utilization such as readmission rate, ED visits, and hospital spending.

Previous evaluations for the pilot model (with the implementation period as of 2010-2013) using difference-in-differences designs and within-state control group found no significant effects on readmission rates or acute hospital stays among Medicare beneficiaries with an 18 and 36 month follow-up period (Mortensen, Perman, and Chen 2014; Roberts, Hatfield, et al. 2018). Two recent studies using all-payer claims and within-state controls found reductions in different hospital departments evaluating the whole implementation period (Done, Herring, and Xu 2019; Pines et al. 2019). Done et al. (2019) found a significant 8.9% reduction in outpatient visits using rural untreated Zip Code Tabulation Areas as the control group. Pines et al. (2019) found a 12% decline in ED admission, a 23% decline in non-ED admissions, a 45% decrease in ambulatory surgery center visits, and a 40% reduction in outpatient clinic visits and services using seven similar non-TPR hospitals in Maryland as the control group.

Evaluations of the formal GBR program using a difference-in-differences approach and across-state control group yielded mixed results on hospital service utilization (e.g., admission and readmission, ED visits, and outpatient department utilization) for the Medicare population (Giuriceo et al. 2017, 2018; Roberts, McWilliams, et al. 2018). Significant reductions were found in the RTI's third annual report (2018)

where all-cause acute inpatient admissions per 1,000 patients decreased nearly 5%, ambulatory care sensitive condition admissions per 1,000 decreased 9.4%, and per capita expenditures for inpatient, outpatient, ED visits, and observation stay all decreased (Giuriceo et al. 2018). Meanwhile, Roberts et al. (2018) showed evidence that evaluation results on primary care visits, hospital stays, ED visits, return hospital stays, hospital outpatient department use, and post-hospitalization primary care visits are not stable across different model specifications. For instance, there was a relative increase in primary care visits with no reduction in hospital stays in a parsimonious model that included only the difference-in-difference interaction terms and area and year fixed effects. However, adding linear trends that were specific to Maryland and the comparison group, there was a reduction in hospital stays with no increase in primary care (Roberts, McWilliams, et al. 2018). It is impossible to know with certainty which model is superior as it depends on which models' un-testable assumptions are more accurate. For commercial plan members, RTI found significant reductions in ED visits, potentially avoidable admissions, and unexpected reduction in admission severity by December 2017 (Giuriceo et al. 2019).

In summary, a certain level of inconsistency exists among these studies. The inconsistency may arise from differences in study population (RTI used all 46 hospitals while Roberts et al., used only 36 urban hospitals), choices on comparison group (RTI used matched hospitals while Roberts et al., used matched counties), and the evaluation time range (RTI's results were through 3 years implementation while Roberts et al., used data through the first two years). Additional uncertainty may also exist due to the

technical issues that come along with the difference-in-differences method, such as serial correlations or the violation of parallel trend assumption.

Studies using less robust study designs have found more consistently positive results. In 2014, the annual growth of per capita hospital costs for all payers increased by 2.11% and Medicare costs decreased by 1.08%; the inpatient admissions per 1,000 Medicare beneficiaries decreased by nearly 5%; the rate of 65 potentially preventable conditions (PPC) dropped by 26.3%; and the Medicare all-cause readmission rate dropped by 0.2% in Maryland compared to the national average (Patel et al. 2015). While such results are certainly compelling, it is unclear if the pre-period experience is the right counterfactual for what would have happened in the absence of the program. Nonetheless, from the HSCRC report in 2018, hospitals fulfilled all the requirements after four years of implementation of the GBR program. Specifically, as measured in 2017, all-payer hospital revenue growth was 3.54%; Medicare savings in hospital expenditure was 5.63% lower than the national average growth rate from the 2013 base year; Medicare savings in the total cost of care was 1.36% lower than the national average; all-payer quality improvement reductions in PPCs under Maryland Hospital-Acquired Conditions (MHAC) program met a 50% reduction; readmission reductions for Medicare were 0.19% below the national average, and 100% of hospitals transferred revenue to globally based (HSCRC 2018a). On the other hand, hospitals have made some changes to reduce admissions, such as working with nonprofit health services providers to visit patients at home for those who visit ED frequently or discharging patients into long-term care settings (Sharfstein, Kinzer, and Colmers 2015).

1.2 Conceptual Framework

I provide an illustrative conceptual framework to think about how the Maryland Global Budget Revenue Program might influence NICU admission and utilization (e.g. length of stay) decisions. With the rapid increase in numbers of NICUs and numbers of NICU beds, the availability of bed supply may directly lead to additional utilization (Roemer 1961). Evidence supports this theory as Freedman exploited short-run within-hospital-month deviations in the number of empty beds using data from California (1991-2001) and New York (1994-2003) and found a causal effect of higher NICU bed supply on admission for low birthweight (1500g-2500g) newborns rather than very low birthweight (<1500g) (Freedman 2016b). Harrison et al., (2018) conducted a descriptive study using US birth certificates (2013) and American Hospital Association (2012) data that found that newborns' admissions to NICUs were positively correlated with NICU bed supply (W. N. Harrison, Wasserman, and Goodman 2018). Although the evidence suggested overutilization of NICU service in areas with more bed supply, it's not practical to shut down these NICUs to reduce unnecessary utilization and cost. Maryland's all-payer reform, therefore, offers a potential solution to contain the utilization and cost within a reasonable range. This study is also motivated by the "financially sensitive" feature of the NICU where it's known as one of the major profit centers for hospitals. A 2010 Health Affairs article that profiled one academic medical center found that NICU admissions made up just for 4% of total hospital admissions but accounted for 69% of net profits (Lantos 2010). In the next section, I will further discuss how Maryland's GBR offers economic incentives that could drive the change in utilizations of NICU and other neonatal services in a theoretical framework.

Here, I introduce a weighted utility function model suggested by McGuire (2000) and Freedman (2016). The doctor acts as the key decision-maker (the hospital is assumed to face similar incentives as suggested by literature) and will maximize his/her utility by considering his/her own welfare as well as the patient's welfare within a weighted function in (1.1) where the weight $\alpha > 0$. That is, the doctor improves utility when he/she provides additional services to patients while the corresponding harm to patients (such as financial loss, mental stress, or adverse medical events) will decrease his/her utility. The utility gained from NICU admission varies with a series of infant risk factors, such as birthweight, gestational age, etc. For simplicity and following previous literature, I use a single index b that quantifies risk on a continuum. Therefore, the doctor must choose an optimal threshold on b that will lead to NICU admission when crossed. The value of the threshold, called b^* , is chosen to maximize his/her utility within a constrained choice set $(0, B(\text{capacity}))$. The upper limit of the choice set is bounded by capacity, which may include hospital bed supply and all other hospital resource constraints. Let $U(b)$ be the overall utility of a doctor admitting a newborn into the NICU, which equals a weighted sum of doctor's payoff $U_d(b)$ and patient's payoff $U_p(b)$:

$$U(b) = \alpha U_d(b) + (1 - \alpha)U_p(b) + \text{Constant} \quad (1.1)$$

$$U'(b) = \alpha U'_d(b) + (1 - \alpha)U'_p(b),$$

$$\text{Where } b \in (0, B(\text{capacity})), \frac{\partial U'_d(b)}{\partial b} \leq 0, \text{ and } \frac{\partial U'_p(b)}{\partial b} \leq 0.$$

The payoffs of admitting a newborn for both doctor and newborn are assumed to increase with higher risk factors. If $U'(b) > 0$, the doctor will admit the newborn into the NICU; if $U'(b) < 0$, then the newborn won't get admitted into the NICU; If $U'(b) = 0$,

then the doctor is indifferent between admitting the newborn or not. b^* is the threshold where the doctor is indifferent between admitting the newborn or not. If taken b as birthweight, then newborns weighing less than b^* will get admitted while those weighing more than b^* won't.

The theory of “supplier-induced demand” is supported by literature investigating the availability-effect or fee-effect (Mcguire 2000). While evidence of the fee-effect is mostly found in Medicare plans where the fee and reimbursement structure is observable, literature in NICU focuses on the availability-effect because of the dramatic capacity increase of NICU beds in the US since the 1980s when it's likely that demand is oversupplied. Freedman (2016) suggested three mechanisms that might lead b^* to change with NICU bed capacity. They are the *Income Effect*: As empty beds increase, physicians see this as a negative income shock and they are more willing to raise b^* to compensate their income loss; *Option Value*: As there are more empty beds available, the opportunity cost of admitting a heavier infant decrease so raising b^* may bring more benefit to physicians; and *Congestion Externalities*: As capacity increases, the spillover effects of “quality of care” is decreasing given physician and nurse resources are less congested, and this may allow physicians to be more likely to admit a marginal infant.

Similarly, I suggest some mechanisms that the GBR program might change the threshold of b^* (where $U'(b^*) = 0$). I assume that the utility functional form won't be affected by GBR. The capacity of NICUs (measured by NICU beds) stay relatively stable in Maryland during my study period (<3.1% change from 2010 to 2015) (Giuriceo et al. 2016).

Case I: The b^* is below the upper bound $B(\text{capacity})$ before GBR. Given that total spending growth is capped at 3.58%, it's possible that every hospital department is required to constrain their utilization to avoid over-spending, which gives practitioners incentives to admit fewer infants that would otherwise be admitted. Hence, this can be seen as the choice set shrinks due to GBR and I should observe a decrease or no change in b^* . On the other hand, if GBR does not affect the neonatal department (i.e., the change in the budget is not binding on the NICU) and practitioners are allowed to make their own admission decisions, then either the choice set is unchanged or expanded and the admission of infants won't be affected at all. In this case, I should observe no changes in b^* .

Case II: The b^* is bounded by $B(\text{capacity})$. In this case, the choice of b^* follows the change of $B(\text{capacity})$. GBR might lead to shrinking, expansion, or no change of the choice set through affecting the hospital's investment in infrastructure. Correspondingly, b^* will decrease, be unchanged, or increase.

I provided some evidence that Case I is probably the correct case, i.e., there will be an interior solution rather than a corner solution. I examined the occupancy rate of NICU in Maryland before GBR. Specifically, I calculated the occupancy rate by dividing NICU days by NICU bed days available (Halpern et al. 2016) using 2014 SID files from the HCUP discharge data (HCUP 2019a) and the number of NICU beds from the FY2015 MHCC's report (Maryland Health Care Commission 2015). I found an average occupancy rate of 39.9%. Similarly, Freedman also suggested that capacity constraints and congestion externalities are not the reasons that drove NICU admissions using data from California and New York (Freedman 2016a). In addition, GBR allows hospitals to

adjust prices within a certain range. Under GBR, hospitals might decrease NICU admissions and choose to increase the price of NICU (Giuriceo et al. 2016). The hospital doesn't have incentives to decrease prices in this circumstance. Therefore, the effect of a price change will fall into Case I. Moreover, it's possible that doctors are not employed by the hospital to do the admitting. In this case, GBR won't affect them, and the admission rate shouldn't be affected.

1.3 A Preview of Chapters Two, Three, and Four

In Chapter Two, I examine the impact of GBR on the Neonatal Intensive Care Unit (NICU) admissions and infant mortality. The NICU is a particular medical technology that is characterized by high prices and potentially inefficient utilization. I apply a difference-in-difference design comparing Maryland to 20 states (including DC) before and after the GBR. I use the restricted-use birth certificates data from the Vital Statistics, which collects demographic and clinical information of newborns and mothers from a near census of births in the US. I find that the GBR is associated with a 16.8% (1.26 percentage points, bootstrap p-value=0.03) decrease in the NICU admission rate. The decline is primarily driven by infants that were relatively low-risk, corresponding to birthweight above 1,500g or gestational age >32 weeks. There's no impact of GBR on neonatal or infant mortality rate. These findings suggest substantial potential savings in neonatal hospital services after capping hospital revenues, which can be achieved without decreasing measurable care quality. The lessons from Maryland could help those states that are planning to adopt a similar global budget model, such as Pennsylvania and Vermont.

Chapter Three expands the analysis to broader hospital services related to births. In this chapter, I examine the impact of GBR on lengths of stay (LOS), total cost of care, and specific services utilization of infants. I use the inpatient discharge data from the Healthcare Cost and Utilization Project (HCUP), comparing Maryland with New York, New Jersey, and Kentucky through a difference-in-differences design. I find that GBR is associated with a decline in total LOS and the utilization of a series of neonatal services for newborns. These findings are a supplement to current findings on the effect of GBR on aggregate outcomes and shed light on an essential population that is mainly financed by Medicaid and private insurance.

Chapter Four redirects to NICU service and further explores the variation in NICU utilization that is related to another source of financial incentive – insurance coverage. In this chapter, I used the nationwide birth certificate data, including all states and DC, to describe the overall variation in NICU admissions across insurance type, and then stratified by birthweight. I find a significant variation in NICU use between Medicaid and privately insured patients. However, the variation is gone after adjusting for infant risk among the very-low birthweight infants that need intensive services the most. Nevertheless, the variation persists for normal birthweight infants that were relatively healthy. Although these findings are descriptive, they suggest a great amount of variation that is not attributed to the demand-side when the infant risk factors and maternal characteristics are mostly controlled for.

Chapter 2 : Changes in NICU Admissions after Maryland's Global Budget Revenue Program

2.1 Introduction

The U.S. is engaged in a number of payment reform activities that are designed to constrain health care cost growth (Center for Medicare & Medicaid Innovation 2018). One of the most ambitious state-based programs is Maryland's Global Budget Revenue (GBR) program, which prospectively sets global budgets for every acute-care hospital in the state (CMS 2018). With a fixed budget encompassing revenues from inpatient, outpatient, emergency department, and hospital-based ambulatory surgery centers, and penalties for deviation of more than 0.5%, hospitals are now incentivized to limit volumes and substitute high-cost services for low-cost population health investments. A pilot version of the program was introduced in rural hospitals in 2010. The formal model, i.e., the GBR program, was then launched statewide in January 2014. With a short transition period, all 46 acute-care hospitals in Maryland were operating under a global budget setting from July 2014 (Giuriceo et al. 2016).

Comprehensive evaluations of both the intended and unintended consequences of GBR are crucial for policymakers refining the global budget model. Lessons from Maryland are also essential for other states that now considering the global budget, such as Pennsylvania and Vermont (CMS 2019b, 2019c). Prior findings on GBR are mixed for changes in aggregate hospital utilization, including inpatient admission and readmission, emergency department (ED) visits, and outpatient department utilization (Roberts,

Hatfield, et al. 2018; Mortensen, Perman, and Chen 2014; Done, Herring, and Xu 2019; Pines et al. 2019; Giuriceo et al. 2019, 2016, 2017, 2018). Moreover, the studies on GBR have predominantly focused on Medicare beneficiaries (Roberts, McWilliams, et al. 2018; Beil et al. 2019). Evidence from other populations and focusing specifically on distinct high-cost services is lacking.

This study considers a previously unexamined service: Neonatal Intensive Care Units (NICU). NICUs are highly effective for infants that need them but are costly (Scott A. Lorch et al. 2012; Phibbs et al. 2007). NICU services are also sensitive to financial incentives. Recent work has found positive relationships between NICU admissions with bed supply and low unit census (Freedman 2016a; W. N. Harrison, Wasserman, and Goodman 2018; Goodman et al. 2019; J. Schulman et al. 2018). These correlations are largest among infants who do not appear to possess clear clinical indicators of need, suggesting that not all infants treated in the NICU require it and that some infants might be adequately cared for outside of the NICU.

In this study, I estimate the association of the GBR program with NICU utilization and investigate how the associations vary across birthweight and gestational age categories. While the hope is that hospitals will substitute unnecessary high-cost care with lower-cost alternatives that produce equal or superior health outcomes, the program may inadvertently reduce services in ways that decrease health. To measure this dynamic, I also explore the association of GBR with infant and neonatal mortality rates.

I use a difference-in-differences design comparing Maryland with 20 states (including DC) before and after the implementation of GBR. Data comes from restricted-use Vital Statistics. I use the years of 2011 to 2017 and examine the effects after three full

years of implementation, a longer implementation period that has been observed by previous GBR studies.

2.2 Methods

2.2.1 Study Design

I used a difference-in-differences approach to compare NICU admission among newborns in Maryland versus newborns in other states, before (2011-2014) and after (2015-2017) the implementation of the GBR program (Dimick and Ryan 2014). To interpret these estimates as causal effects, one must assume that Maryland would have followed the same trend as the comparison states in the absence of the program. Comparison states included 19 states and the District of Columbia that collected NICU admission information on their birth certificate forms and adopted the ACA Medicaid Expansion as Maryland did. The full list of comparison states is provided in Table A.1 (Appendix A). While hospitals were subject to global budgets at the start of the Fiscal Year 2014 (July 2014) and applied penalties since Fiscal Year 2015, I specified that the post-intervention period started in 2015 which was the first full calendar year of the program implementation.

The American Academy of Pediatrics (AAP) recommends NICU admission for all very-low birthweight infants (weighing less than 1,500 grams) and very preterm infants (born before 32 weeks of gestation) (American Academy of Pediatrics 2012; Kilpatrick, Papile, and Macones 2017). Fewer infants above these thresholds are likely to need NICU care, and NICU utilization for such infants has been shown to be sensitive to financial incentives (Freedman 2016a; W. N. Harrison, Wasserman, and Goodman 2018;

American Academy of Pediatrics 2012; J. Schulman et al. 2018; Angert and Adam 2008; Cloherty et al. 2012). I hypothesized that GBR would only affect NICU admission rates among higher weight and longer gestation infants that are more likely to be safely treated outside of the NICU setting. I conducted subgroup analyses by birthweight and gestation using commonly used categories (defined below) (Freedman 2016a; W. Harrison and Goodman 2015; Kilpatrick, Papile, and Macones 2017).

In my secondary analyses, I assessed whether GBR led to changes in infant and neonatal mortality rates. Neonatal mortality (death in the first 28 days) is likely to be more sensitive to NICU care than infant mortality (death in the first year) (Goodman et al. 2002; WHO 2007). Mortality analyses were conducted for all infants, for moderately-low and normal birthweight infants, and for moderately preterm and term infants. Within each group, I considered all infants in the category and infants not admitted into a NICU. If GBR reduced NICU care among infants who needed it, then I expected that the largest increase in mortality would be for those not admitted to a NICU.

2.2.2 Data

I used data from restricted-use Vital Statistics which comprise a near census of live births from 2011 to 2017 (CDC 2019). I obtained state of birth, NICU admission status, and infant and maternal characteristics from the Birth Files. I used state-of-birth rather than state of mother's residence because, with some exceptions, GBR budgets are determined by the amount of care provided to both in-state and out-of-state residents. Mortality was obtained from the Linked Birth-Death Files. The sample included all singleton births weighing 500g and over. Both files were merged with state-by-year

characteristics obtained from the Area Health Resources Files (HRSA 2019), Kaiser Family Foundation (KFF 2017), and CDC Wonder (CDC 2018).

2.2.3 Study Variables

The primary outcome was an indicator of NICU admission status. NICU admission in the Birth Files is defined as “Admission into a facility or unit staffed and equipped to provide continuous mechanical ventilatory support for the newborn” (Center for Health Statistics 2003). Importantly, this indicator measures the admission of an infant to a certain clinical setting, but does not necessarily indicate the types of care that the infant received. I grouped infants into four birthweight categories: very-low (VLBW, 500-1,499g), moderately-low (MLBW, 1,500-2,499g), normal (NBW, 2,500-3,999g), and high (HBW, 4,000g and above) (W. Harrison and Goodman 2015). Separately, I also grouped the sample into four gestation categories: very preterm (<32 weeks), moderately preterm (32-36 weeks), term (37-41 weeks), and postterm (>41 weeks) (Kilpatrick, Papile, and Macones 2017). I measured infant and neonatal mortality rates as the number of deaths per 1,000 live births per state-year (M. H. Boudreaux, Dagher, and Lorch 2018).

The covariates in the individual-level NICU admission analysis included infant characteristics: birthweight, gestational age, gender, and an indicator of any congenital anomaly; and a set of maternal characteristics: race/ethnicity, age, education, insurance type, parity, an indicator of any maternal morbidity, an indicator of any maternal infection during pregnancy, and an indicator of any risk factor during pregnancy. The maternal characteristics were used to adjust for changes in health risks at the time of birth. I also included time-varying state-level characteristics, including poverty rates,

unemployment rates, birth rates, and the number of NICU beds per 1,000 residents to control for changes in the supply of NICU services. Further details about covariates are described in Appendix A.

In the aggregate-level mortality analyses, I controlled for the percent of infants that were moderately-low birthweight, preterm, and had any congenital anomalies; percent of mothers that were non-Hispanic White, aged less than 35 years, had less than high school education, first births, had any maternal morbidities, had any infections during pregnancy, had any risk factors during pregnancy, and the state-level characteristics as above.

2.2.4 Statistical Analysis

I implemented the difference-in-differences comparison in a linear regression framework. Individual-level NICU admissions were modeled using linear probability models and aggregate-level mortality rates with linear regressions. Linear models were chosen due to the ease of interpreting the coefficients. However, we come to similar results using logistic regressions (Table A.6). The models controlled for the covariates described above in addition to state and year fixed effects. State fixed effects accounted for unobserved state-specific factors that were stable over time and the year fixed effects controlled for year-specific changes that were common for all states. 95% confidence intervals were obtained from the clustered sandwich estimator to account for state clustering. I also report P-values that were obtained from a bootstrap method that better accounted for serial correlation in the presence of a single treated cluster (Ferman and Pinto 2019). My regression of interest took the following form:

$$Y_{ist} = \delta (Treatment_s * Post_t) + X_{ist} \beta_1 + Z_{st} \beta_2 + u_s + \gamma_t + e_{ist}$$

Where Y_{ist} is the indicator of NICU admission, $Post_t$ is a dummy variable for whether it's the post-intervention period, $Treatment_s$ is a dummy variable indicating whether the unit is treated in the post-intervention period, X_{ist} is a vector of individual-level covariates, Z_{st} is a vector of state-by-year level covariates, u_s is a vector of state fixed effects that control for time-invariant characteristics within a state, γ_t is a vector of year fixed effects that control for state-invariant changes across years, and e_{ist} is the error term. δ is the DID estimate of the policy effect that estimates the mean difference in outcomes between Maryland and comparison states, before and after the year 2015. A similar strategy was used to model state-year infant mortality rates.

I conducted additional analyses to determine if these results were sensitive to reasonable alterations in approach. I first examined if I came to similar conclusions using alternative comparison groups, different covariates, different post-period definitions, using state of residence rather than state of birth, restricting to births that occurred in urban hospitals, or adding state-specific linear trends. I also examined results after excluding Baltimore City and Baltimore County where a local initiative was launched in 2009 to reduce infant mortality (Baltimore City Health Department 2009). I considered alternative specifications such as logistic regression for binary outcomes. I also examined if GBR led to changes in health at birth to better understand if any changes I observed in NICU admission could have possibly resulted from changes to infant health. To measure health at birth, I used an indicator of preterm, an indicator of small for gestational age, detailed birthweight in grams, and weeks of gestation. Finally, I used a data-driven method, the synthetic control method, to examine if I came to a similar conclusion as

using the difference-in-differences. Details regarding these sensitivity analyses are provided in Appendix A. Analyses were conducted using Stata version 15 (StataCorp). Data analysis was approved by the University of Maryland Institutional Review Board (IRB).

2.3 Results

2.3.1 Newborn Characteristics

I observed a total of 11,965,997 newborns in Maryland and the comparison states. Baseline (2011-2014) characteristics are presented in Table 2.1. Most covariates were qualitatively similar between Maryland and the comparison states, although Maryland had a higher proportion of infants born to non-Hispanic Black mothers, a lower proportion born to Hispanic mothers, and a smaller Medicaid population. Hence, I did a balancing test for problematic compositional changes. Specifically, I replaced the outcome variable with the covariate and fitted the standard DID regression model, and then I examined whether the magnitude of delta (δ) is small and not significant. As shown in Table A.13, I found that the differences of most covariates are stable over time. The two exceptions were race and insurance type. We can see that the race distributions were quite stable in Maryland and the comparison group, but the Hispanic population increased by 3 percent in Maryland while almost unchanged in comparison states. Also, the Medicaid population increased more in Maryland after the intervention compared to the comparison states. While significant, the differences were relatively small in magnitude. And we will see later in the results section that these differences in covariates are much smaller compared to the size of the program's effects and they're not possible to drive the entire effect.

Table 2.1: Characteristics of the Study Population at Baseline, 2011-2014

| Characteristic, % | Maryland (n=254,832) | Comparison States (n=6,606,329) | P-Value |
|--|-------------------------|---------------------------------------|---------|
| Maternal Characteristics | | | |
| Maternal Race | | | <0.001 |
| Non-Hispanic White | 47.66 | 54.17 | |
| Non-Hispanic Black | 30.94 | 10.86 | |
| Non-Hispanic Other | 7.97 | 9.48 | |
| Hispanic | 13.43 | 25.49 | |
| Maternal Age, years | | | <0.001 |
| <20 | 5.74 | 6.89 | |
| 20-24 | 18.97 | 21.73 | |
| 25-34 | 57.4 | 55.27 | |
| 35-44 | 17.65 | 15.93 | |
| 45+ | 0.24 | 0.18 | |
| Maternal Education | | | <0.001 |
| Less than High School | 13.28 | 16.72 | |
| High School | 20.81 | 24.32 | |
| Some College | 27.98 | 28.58 | |
| University and Above | 37.93 | 30.38 | |
| Maternal Insurance | | | <0.001 |
| Medicaid | 33.44 | 43.38 | |
| Private Insurance | 57.72 | 49.53 | |
| Self-Pay | 3.31 | 3.05 | |
| Other | 5.53 | 4.04 | |
| Parity | | | <0.001 |
| First | 41.58 | 40.48 | |
| Second | 33.21 | 31.84 | |
| Third or Higher | 25.21 | 27.67 | |
| Any Indications of Maternal Infection | 2.35 | 2.14 | <0.001 |
| Any Indications of Maternal Morbidity | 2.43 | 1.6 | <0.001 |
| Any Indications of Maternal Pregnancy Risk | 30.56 | 27.25 | <0.001 |
| Infant Characteristics | | | |
| Any Indications of Congenital Anomaly | 0.32 | 0.3 | 0.068 |
| Male | 51.16 | 51.22 | 0.583 |
| Birthweight | | | <0.001 |
| Very-Low | 1.09 | 0.86 | |
| Moderately-Low | 5.41 | 4.83 | |
| Normal | 85.04 | 85.63 | |
| High | 8.47 | 8.69 | |

| Characteristic, % | Comparison | | P-Value |
|-------------------------------------|-------------------------|-------------------------|---------|
| | Maryland (n=254,832) | States (n=6,606,329) | |
| Gestational Age | | | <0.001 |
| Very Preterm | 1.26 | 1 | |
| Moderately Preterm | 6.65 | 6.13 | |
| Term | 91.77 | 92.38 | |
| Postterm | 0.32 | 0.49 | |
| <i>State Characteristics</i> | | | |
| Poverty Rate | 8 | 12.89 | <0.001 |
| Unemployment Rate | 6.71 | 8.41 | <0.001 |
| Birth Rate | 12.32 | 12.44 | <0.001 |
| NICU Bed, per 1,000 pop | 0.07 | 0.06 | <0.001 |

Note. The sample excludes those with birthweight less than 500g or unknown, non-singleton births, and those with missing values. The baseline period is from 2011 to 2014. Comparison states include 19 states and the District of Columbia that collected NICU admission information in our study period and adopted the ACA Medicaid Expansion as Maryland did. The full list of comparison states is provided in Appendix A (Table A1). The birthweight categories were defined as: very-low birthweight (VLBW, 500-1,499g), moderately-low birthweight (MLBW, 1,500-2,499g), normal birthweight (NBW, 2,500-3,999g), and high birthweight (HBW, 4,000g and above). The gestation categories were defined as: very preterm (<32 weeks), moderately preterm (32-36 weeks), term (37-41 weeks), and postterm (>41 weeks).

2.3.2 NICU Utilization

Before the implementation of GBR, the overall NICU admission rate in Maryland increased steadily from 7.28 per 100 births in 2011 to 7.86 in 2014 (Figure 2.1: Unadjusted Trends in NICU Admission: Maryland versus the Comparison States-A). Similar trends occurred in the comparison group. After the implementation of GBR, the overall NICU admission rate in Maryland decreased from 7.86 in 2014 to 6.79 in 2017, while NICU admissions continued to increase in the comparison states.

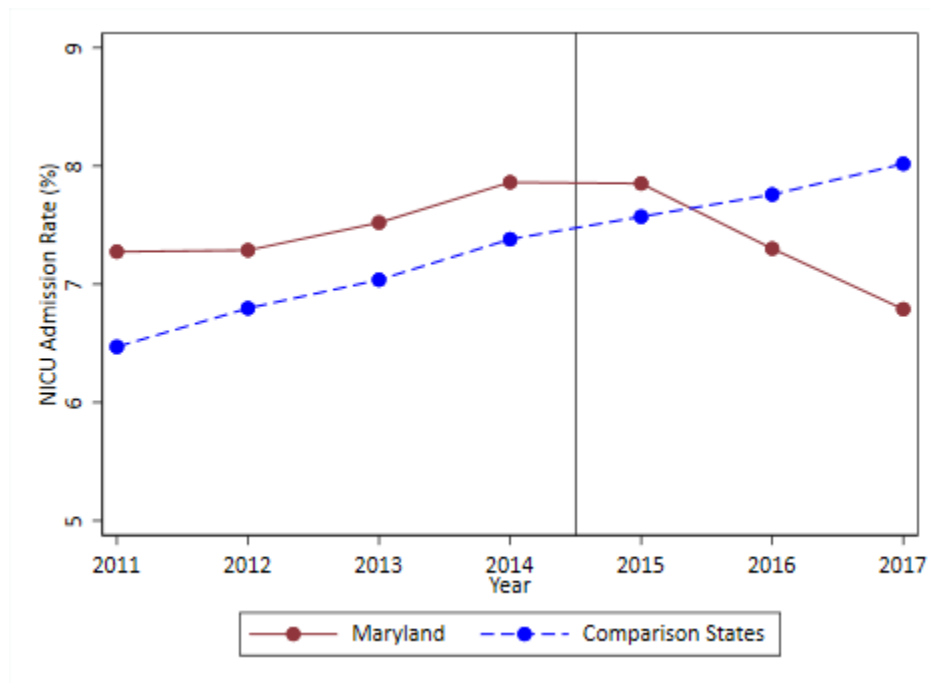
Figure 2.1-B and Figure 2.1-C suggested relatively large declines in NICU admission rates in Maryland among MLBW, NBW, and HBW infants, and among moderately preterm and term infants that were not observed in the comparison states. I

did not observe a similar pattern for VLBW or very preterm infants. I observed similar trends before GBR in both Maryland and the comparison states among most birthweight and gestation groups so that the parallel trends assumption, as required of difference-in-differences designs, was met. Additional analyses (Table A.2 in Appendix A) confirmed that trends in Maryland before GBR were statistically similar to trends in the comparison states. The exception was for the very-low birthweight group in which pre-period admissions in Maryland appeared to be increasing at a slightly faster rate than the comparison states. While this difference in pre-trends did not appear to be large enough to be of substantive concern, results for the VLWB group should be interpreted with caution.

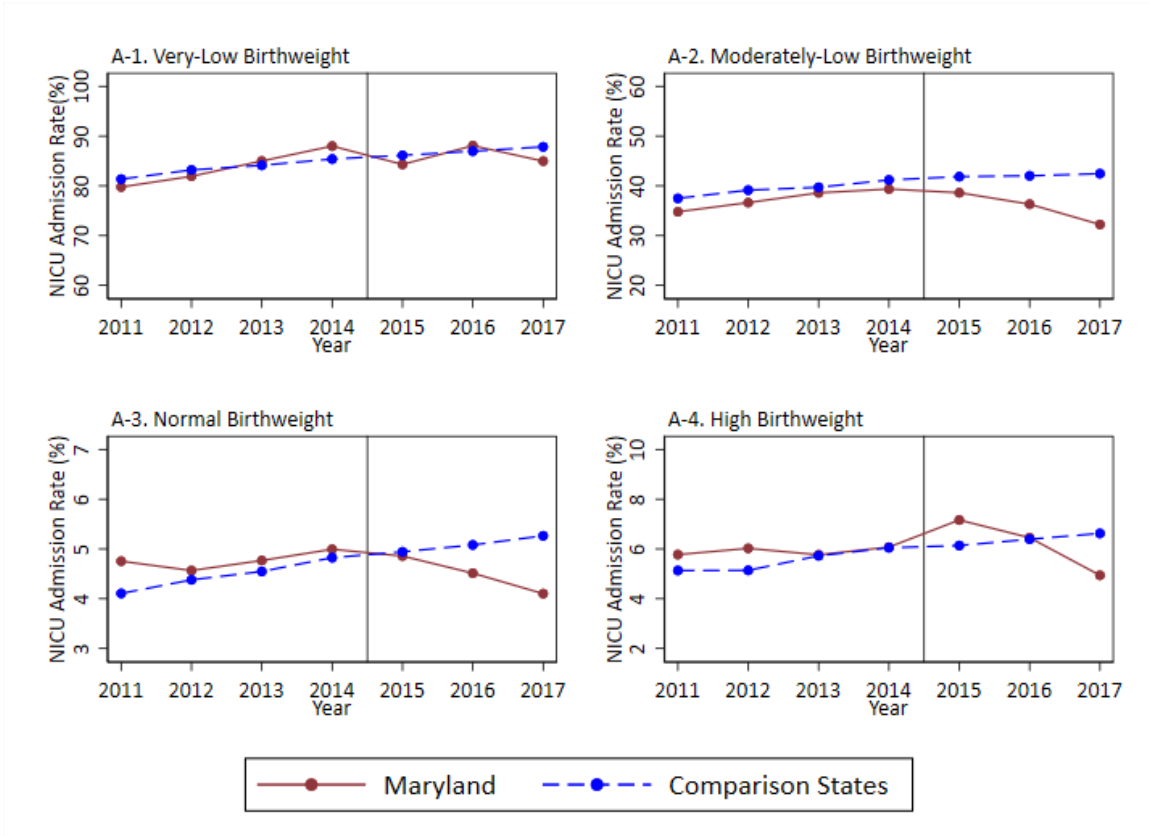
Figure 2.1: Unadjusted Trends in NICU Admission: Maryland versus the Comparison

States

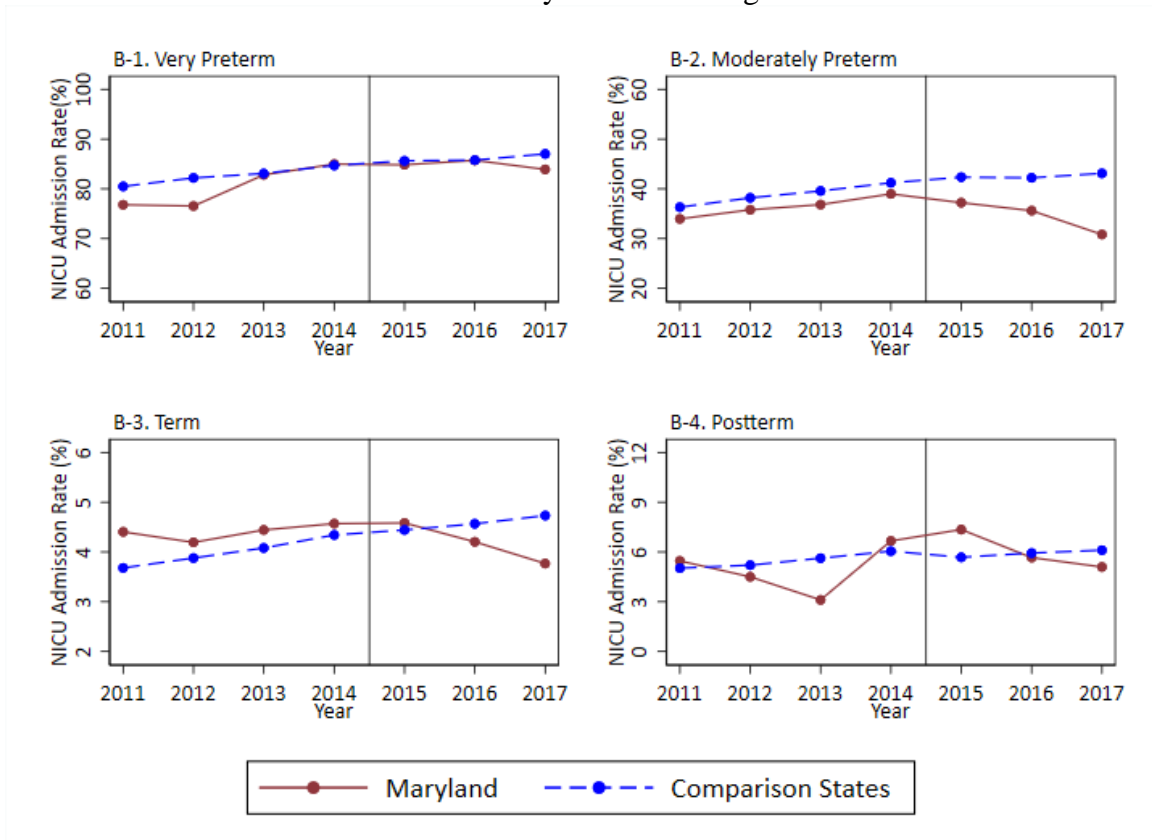
A. Overall



B. By Birthweight



C. By Gestational Age



Note. Comparison states included 19 states and DC. The birthweight categories were defined as: very-low birthweight (500-1,499g), moderately-low birthweight (1,500-2,499g), normal birthweight (2,500-3,999g), and high birthweight (4,000g and above). The gestation categories were defined as: very preterm (<32 weeks), moderately preterm (32-36 weeks), term (37-41 weeks), and postterm (>41 weeks).

Table 2.2 and Table 2.3 describe the adjusted difference-in-differences results for all births and for each birthweight and gestation category. The model suggests that GBR was associated with a 1.26 percentage points (-16.8%; 95% CI, -1.76 to -0.76; P=.03) decline in NICU admission (Table 2.2). The association among MLBW and NBW infants was -4.5 percentage points (-12.0%; 95% CI, -5.71 to -3.29; P=.003) and -1.1 percentage points (-23.1%; 95% CI, -1.58 to -0.62; P=.04), respectively. Similarly, in Table 2.3, I observed statistically significant associations of GBR with NICU admission for

moderately pre-term (-15.5%; 95% CI, -7.06 to -4.23; P=.01) and term (-22.3%; 95% CI, -1.45 to -0.52; P=.05) infants. Conversely, I observed small and non-significant associations for VLBW and very preterm infants. Changes among HBW and postterm infants were relatively large but not statistically significant (-14.2%/14.8%; 95% CI, -1.36 to -0.31/0.17 to 1.64; P=.13/.43).

2.3.3 Infant Mortality

As shown in Table 2.4, GBR did not have a statistically significant association with infant and neonatal mortality rates overall, among all MLBW and NBW infants, among moderately preterm and term infants, or among such infants that were not admitted into a NICU. Although not statistically significant, the point estimates suggested a reduction in mortality for all groups considered. In Appendix A, I present additional details for these analyses.

Table 2.2: Effect of GBR on NICU Admissions, Overall and By Birthweight, 2011-2017

| <i>Dependent Variable: NICU Admission Indicator</i> | Maryland | | | Comparison States | | | Adjust Difference-in-Differences Estimate (95% CIs) | | Bootstrap P-Value | Relative Effect from Baseline |
|---|--------------|-------------|-----------------------|-------------------|-------------|-----------------------|---|---------------|-------------------|-------------------------------|
| | % before GBR | % after GBR | Unadjusted Difference | % before GBR | % after GBR | Unadjusted Difference | | | | |
| All (N=11,965,997) | 7.5 | 7.3 | -0.2 | 6.9 | 7.8 | 0.9 | -1.26 | (-1.76,-0.76) | 0.03 | -16.80% |
| By Birthweight | | | | | | | | | | |
| Very-Low Birthweight <i>N=104,356</i> | 83.7 | 85.7 | 2 | 83.5 | 87 | 3.5 | -0.95 | (-2.58,0.69) | 0.59 | -1.10% |
| Moderately-Low Birthweight <i>N=591,120</i> | 37.4 | 35.7 | -1.7 | 39.4 | 42.1 | 2.7 | -4.5 | (-5.71,-3.29) | 0.003 | -12.00% |
| Normal Birthweight <i>N=10,238,271</i> | 4.8 | 4.5 | -0.3 | 4.5 | 5.1 | 0.6 | -1.1 | (-1.58,-0.62) | 0.04 | -23.10% |
| High Birthweight <i>N=1,032,250</i> | 5.9 | 6.2 | 0.3 | 5.5 | 6.4 | 0.9 | -0.84 | (-1.36,-0.31) | 0.13 | -14.20% |

Note. Estimates are from separate regressions for all births and for each birthweight group. The coefficients are in percentage points. The models control for birthweight (in full and gestation analyses), gestational age (in full and birthweight analyses), mother's age, race, education level, insurance type, parity, maternal morbidities, infections, and risks, infant's sex and congenital anomalies, state-level poverty rate, unemployment rate, birth rate, NICU bed per 1,000 population, and state and year fixed effects. N denotes the number of observations. The birthweight categories were defined as: very-low birthweight (500-1,499g), moderately-low birthweight (1,500-2,499g), normal birthweight (2,500-3,999g), and high birthweight (4,000g and above). The gestation categories were defined as: very preterm (< 32 weeks), moderately preterm (32-36 weeks), term (37-41 weeks), and postterm (>41 weeks). Comparison states included 19 states and DC. Confidence intervals are based on standard errors that are clustered at the state level. P-values are obtained using a bootstrap approach, developed by Ferman and Pinto (2019) that better accounts for serial correlation. The baseline rate refers to the average admission rate before GBR in Maryland, which is calculated using data from 2011 to 2014.

Table 2.3: Effect of GBR on NICU Admissions, By Gestational Age, 2011-2017

| | Maryland | | | Comparison States | | | Adjust Difference- in-Differences Estimate (95% CIs) | Bootstrap P-Value | Relative Effect from Baseline |
|---|-----------------|----------------|--------------------------|-------------------|----------------|--------------------------|--|----------------------|--|
| <i>Dependent Variable: NICU Admission Indicator</i> | % before GBR | % after GBR | Unadjusted Difference | % before GBR | % after GBR | Unadjusted Difference | | | |
| By Gestational Age | | | | | | | | | |
| Very Preterm <i>N=120,041</i> | 80.2 | 84.8 | 4.6 | 82.6 | 86.1 | 3.5 | 0.47 (-1.02,1.95) | 0.78 | 0.60% |
| Moderately Preterm <i>N= 744,512</i> | 36.4 | 34.5 | -1.9 | 38.8 | 42.6 | 3.7 | -5.65 (-7.06,-4.23) | 0.01 | -15.50% |
| Term <i>N=11,047,339</i> | 4.4 | 4.2 | -0.2 | 4 | 4.6 | 0.6 | -0.98 (-1.45,-0.52) | 0.05 | -22.30% |
| Post-term <i>N= 54,105</i> | 4.9 | 6.2 | 1.2 | 5.5 | 5.9 | 0.4 | 0.73 (-0.17,1.64) | 0.43 | 14.80% |

Note. Estimates are from separate regressions for each gestation group. The coefficients are in percentage points. The models control for birthweight (in full and gestation analyses), gestational age (in full and birthweight analyses), mother's age, race, education level, insurance type, parity, maternal morbidities, infections, and risks, infant's sex and congenital anomalies, state-level poverty rate, unemployment rate, birth rate, NICU bed per 1,000 population, and state and year fixed effects. N denotes the number of observations. The birthweight categories were defined as: very-low birthweight (500-1,499g), moderately-low birthweight (1,500-2,499g), normal birthweight (2,500-3,999g), and high birthweight (4,000g and above). The gestation categories were defined as: very preterm (< 32 weeks), moderately preterm (32-36 weeks), term (37-41 weeks), and postterm (>41 weeks). Comparison states included 19 states and DC. Confidence intervals are based on standard errors that are clustered at the state level. P-values are obtained using a bootstrap approach, developed by Ferman and Pinto (2019) that better accounts for serial correlation. The baseline rate refers to the average admission rate before GBR in Maryland, which is calculated using data from 2011 to 2014.

Table 2.4: Effect of GBR on Infant/Neonatal Mortality Rates, 2011-2017

| | Maryland | | | Comparison States | | | Adjust Difference- in-Differences Estimate (95% CIs) | | Bootstrap P-Value | Relative Effect from Baseline |
|---|---------------|--------------|--------------------------|-------------------|--------------|--------------------------|---|--------------|----------------------|--|
| <i>Dependent Variable: Infant/Neonatal Mortality Rate</i> | Before GBR | After GBR | Unadjusted Difference | Before GBR | After GBR | Unadjusted Difference | | | | |
| <i>Infant Mortality Rate</i> | | | | | | | | | | |
| All | 4.22 | 4.05 | -0.17 | 3.99 | 3.85 | -0.14 | -0.19 | (-0.47,0.09) | 0.28 | -5.50% |
| Among MLBW and NBW Infants | | | | | | | | | | |
| Overall | 2.62 | 2.71 | 0.09 | 2.77 | 2.74 | -0.03 | -0.06 | (-0.29,0.17) | 0.7 | -2.30% |
| Not NICU Admitted | 2.81 | 2.9 | 0.09 | 2.96 | 2.95 | -0.01 | -0.17 | (-0.40,0.06) | 0.2 | -6.10% |
| Among MPT and Term Infants | | | | | | | | | | |
| Overall | 2.45 | 2.57 | 0.12 | 2.61 | 2.58 | -0.03 | -0.01 | (-0.25,0.22) | 0.9 | -0.40% |
| Not NICU Admitted | 2.62 | 2.74 | 0.12 | 2.79 | 2.78 | -0.01 | -0.11 | (-0.35,0.13) | 0.25 | -4.20% |
| <i>Neonatal Mortality Rate</i> | | | | | | | | | | |
| All | 2.45 | 2.27 | -0.18 | 2.29 | 2.17 | -0.12 | -0.15 | (-0.41,0.11) | 0.23 | -6.10% |
| Among MLBW and NBW Infants | | | | | | | | | | |
| Overall | 1.11 | 1.14 | 0.03 | 1.23 | 1.21 | -0.02 | -0.03 | (-0.20,0.14) | 0.74 | -2.70% |
| Not NICU Admitted | 1.19 | 1.22 | 0.03 | 1.31 | 1.3 | -0.01 | -0.06 | (-0.26,0.15) | 0.58 | -5.00% |
| Among MPT and Term Infants | | | | | | | | | | |
| Overall | 0.96 | 1.05 | 0.09 | 1.14 | 1.11 | -0.03 | 0.01 | (-0.17,0.19) | 0.85 | 1.00% |
| Not NICU Admitted | 1.03 | 1.12 | 0.09 | 1.21 | 1.2 | -0.01 | -0.01 | (-0.25,0.23) | 0.85 | -1.00% |

Note. Infant mortality rate is the number of deaths per 1,000 live births at the state-year level. Neonatal mortality rate is the number of neonatal deaths (infant age<28 days) per 1,000 live births at the state-year level. There are 147 state-by-year cells in each model. All estimates are weighted by the number of total births in each cell. All models control for percents of infants that are moderately-low birthweight, with congenital anomalies, and preterm; percents of mothers that are non-Hispanic White, aged less than 35 years old, with less than high school education, with first birth order, have maternal morbidity, had infections during pregnancy, had risk factors during pregnancy, state-year level poverty rate, unemployment rate, birth rate, NICU bed per 1,000 population, and state and year fixed effects. Standard errors are clustered at the state level. P-values are obtained using a bootstrap approach, developed by Ferman and Pinto (2019). The baseline rate refers to the average admission rate before GBR in Maryland which is calculated using data from 2011 to 2014. MLBW refers to moderately-low birthweight (1,500g-2,499g); NBW refers to normal birthweight (2,500g-3,999g); MPT refers to moderately preterm (32-36 weeks).

2.3.4 Sensitivity Analyses

The results of the sensitivity analyses supported our main findings. My findings on NICU admission rates remained almost unchanged in magnitude and significance when using different comparison states (Table A.4), using January or July of 2014 as the start of the post-period (Table A.5), in models with different covariates (Table A.6), using logistic regression models (Table A.6), restricting to state residents or urban hospital births (Table A.7), or adding state-specific linear trends (Table A.12). The reduction in NICU admission became larger after excluding Baltimore City and County which also had a local ongoing initiative to reduce infant mortality (Table A.7). However, my main conclusions were unaltered. We failed to find evidence that GBR was associated with changes in health at birth (Table A.8), suggesting that changes in NICU admission were not mediated by changes in clinical need. I came to the same conclusion using the synthetic control methods (Section A11 in Appendix A). Further details about the sensitivity analysis can be found in Appendix A.

2.4 Discussion

In this study, I estimated the impact of Maryland's Global Budget Revenue program on NICU utilization and infant mortality. I found that the implementation of GBR was associated with a substantial decrease in NICU admissions in Maryland. The difference-in-difference estimate suggests that GBR was associated with approximately 2,527 fewer NICU admissions between 2015 and 2017, than would have occurred in the absence of GBR. Associations were largest for moderately-low and normal birthweight infants, and among moderately preterm and term infants. This aligns with previous

evidence that suggests that financial incentives are the most likely to affect NICU admissions for infants that do not have clear indicators of clinical need. The reduction in NICU care that I observed did not appear to result in worse birth outcomes as measured by infant or neonatal mortality rates.

This study measured changes in NICU admission where the NICU was defined as a unit that could provide continuous mechanical ventilatory support. This roughly aligns with AAP's level III-IV nurseries. However, this study did not measure changes in actual care delivered to patients and it is possible that while GBR altered the location of where care was delivered, it might not alter the content of that care. For that reason, I cannot come to any specific conclusions about the potential magnitudes of cost-savings. In Chapter Two, I will use hospital discharge data to extend the analysis to the changes in care and related costs.

Previous studies of GBR have come to mixed findings about hospital utilization among Medicare beneficiaries. Discrepancies in these results are due to differences in how services were defined, the choice of control groups, and the length of evaluation periods. One study found no changes in hospital or primary care use after 2 years (Roberts, McWilliams, et al. 2018); another found a 4.9% relative reduction in inpatient admissions and approximately \$554 million savings for hospital services after 3 years (Beil et al. 2019). In addition to these evaluations, two recent studies of the rural pilot program found reductions in outpatient visits of 9-40 percent but came to inconsistent conclusions regarding inpatient and ED care (Done, Herring, and Xu 2019; Pines et al. 2019).

This study focusing on NICU admissions suggests relatively large associations of GBR compared to previous studies. This could be attributed to several factors. First, unlike previous GBR studies that focused on the Medicare population, my comparison group was less affected by national Medicare payment reform activities (e.g., Accountable Care Organizations) because I focused on a service that is primarily financed by Medicaid and private insurance. This gives me a well-performed counterfactual that passes the pre-trend assumption and entitles a lower risk of biased estimates. Second, this study also benefits from a long follow-up period relative to previous studies. Another potential explanation for why I found relatively large associations is that NICU admission decisions are made by physicians who largely practice in the hospital and who as hospitals' employees may have been more sensitive to the hospital-based incentives of GBR compared to physicians who often practice outside of hospitals, and make admission decisions by themselves (M. Schulman 2003; Freedman 2016a). Lastly, the NICU is well known as one of the major profit centers for hospitals and is documented to be affected by financial incentives (Lantos 2010). This makes it a service that is more likely affected by payment reforms like GBR that target high-cost services.

In addition to informing policy discussions about the effects of GBR, my results also offer important insights about NICU care. They are consistent with previous studies showing that NICU admissions among moderate to low-risk infants are sensitive to financial incentives and appear more discretionary than admissions for high-risk infants (i.e., birthweight<1500g and gestation<32 weeks) (Freedman 2016a; W. N. Harrison, Wasserman, and Goodman 2018; Goodman et al. 2019). In the context of a nationwide

increase in NICU admissions, more than 20% since 2007 (W. Harrison and Goodman 2015), my findings indicate that NICU care is a place of potential health system savings and that reducing NICU utilization among some patient populations can be achieved without apparent harm.

This study had limitations. First, there might be measurement error in the NICU admissions indicator obtained from birth certificates. Such errors would be problematic if they varied within a state across time. In Figure A.4, I show that NICU admission rates in Maryland, as measured from vital statistics, followed a similar trend as admission rates estimated from hospital discharge data. While these data sources are not perfectly comparable, the similarity of the trends suggests that my estimates were unlikely to be solely driven by measurement error (See Section A10 in Appendix A for further details). Second, like all quasi-experimental studies, a potential limitation is that unobserved factors could be correlated with the timing of GBR implementation. One clear threat came from the adoption of Medicaid expansion under the Affordable Care Act in 2014, which recent research demonstrates led to improved birth outcomes among African-American infants (Brown et al. 2019). However, by restricting the comparison group to states that adopted the expansion, I limited this concern. Finally, I was unable to detect more detailed neonatal risks (e.g. complications of prematurity, neonatal abstinence syndrome, etc.) or identify harm other than death given the limited information provided in the birth certificate.

2.5 Conclusion

Maryland has enacted one of the largest and most innovative payment reform initiatives in the United States. My findings suggest that hospitals reacted to the policy by

altering NICU service patterns in directions consistent with the intentions of the program. This is important information for other states that are considering Maryland's model and for national policymakers that have suggested using a global budget model to finance a single-payer system (Keith K 2019; Woolhandler and Himmelstein 2019).

Chapter 3 : The Impact of Maryland's Global Budget Revenue Program on Birth-Related Hospitalization: Length of Stay, Cost, and Service Utilization

3.1 Introduction

In the previous chapter, I discussed the impact of Maryland's' global budget revenue (GBR) program on the neonatal intensive care unit (NICU) admissions at the population level, using the birth certificate data. My findings suggested a substantial decline in NICU admissions after GBR, which was mainly driven by infants with relatively higher birthweights or longer gestational ages. In addition, I failed to find evidence of changes to mortality rates, suggesting that changes in utilization did not have negative health consequences. While these findings add important information to the current literature, birth certificate data lack information on specific service intensity measures such as actual services received and the costs of care. Evidence on the impact of GBR on intensive margins, i.e. services provided during each admission, and on birth-related hospitalization, which accounted for approximately 10% of total inpatient admissions and the most frequent reason for hospital stays (Kowlessar, Jiang, and Steiner 2013), is still lacking.

In this paper, I studied changes to length of stay, cost, and neonatal services utilization in response to the GBR program which applied capitated annual budget restrictions to all 46 acute-care hospitals in Maryland since 2014.

This study makes several contributions to the literature. First, it provides the first empirical evidence on the impact of GBR on birth-related hospitalization which accounted for an essential, steady, and large portion of hospitals' annual admissions and revenues. Second, it considers a critically important patient population: newborns. Third, it studies the effect of GBR on a patient population that is half financed by private payers and half by Medicaid. GBR dynamics for this mixed payment population may not track previous analyses focusing on single payment populations such as Medicare. Finally, unlike previous studies that focused on extensive margin changes (i.e., admission), my focus on intensive margin effects offers a relatively unique perspective given that hospitals respond to financial incentives both on the admission decision and on the care provided conditional on admission.

3.2 Methods

3.2.1 Study Design

I used a difference-in-difference approach that compared outcomes in Maryland to those in comparison states over four pre-implementation years, 2011-2014, and two post-implementation years, 2015-2016. Due to the penalty for underage or overage the global budget began in Fiscal Year 2015 (July 2014), and as such, I began the post-period in 2015. Results using 2014 as the post-period starting time were also provided in Appendix B. The comparison states (New Jersey, New York, and Kentucky) were chosen from states that expanded Medicaid at the same time with Maryland, i.e., January 1st, 2014. A limited number of comparison states were chosen due to study financial constraints.

3.2.2 Data

The main data used in this study was from the State Inpatient Database (SID) of the Healthcare Cost and Utilization Project (HCUP) which collected all inpatient care records (HCUP 2019a). It encompassed more than 95 percent of all U.S. hospital discharges each year. For the purpose of this study, I used SID Files of Maryland, New Jersey, New York, and Kentucky from 2011 to 2016. Data from 2017 were not available at the time of the writing of this paper. The medical records for mothers and newborns were separate in the SID and could not be linked in those states that I used, except for New Jersey. This limited my ability to estimate the total cost of birth at the family level, which billed separately for mothers and infants. Therefore, in this paper, I focus my analysis on infants only. To select the newborn sample, live births of newborns were identified using diagnosis codes described in Table B.1 (Appendix B).

The annual hospital-level “cost-to-charge ratio” supplemental files from the HCUP were used to transform hospital charges into actual cost (HCUP 2019b). To account for inflation while comparing estimates of costs of inpatient services from different years, the medical care component of the Producer Price Index from the U.S. Bureau of Labor Statistics was used where costs were adjusted to 2016 US dollars (Dunn, Grosse, and Zuvekas 2018; BLS 2019). Given that the U.S. transitioned to the International Classification of Diseases, 10th Revision, Clinical Modification/Procedure Coding System (ICD-10-CM/PCS) coding scheme on October 1, 2015, I used ICD-9-CM/PCS for all the data before the third quarter of 2015 and ICD-10-CM/PCS thereafter (CMS 2019a). Other state-by-year characteristics used as covariates were collected from

the Kaiser Family Foundation, Area Health Resources Files, and CDC wonder (KFF 2017; HRSA 2019; CDC 2018; KFF 2019).

3.2.3 Study Variables

In this study, I explored a series of utilization measures, including length of stay, NICU care (level III&IV nursery), neonatal special care (level II nursery), and certain high-cost yet commonly used neonatal services, such as X-ray, ultrasound, CT scan, MRI, and respiratory services. I identified the NICU care using revenue codes that denoted level III or level IV nursery care (revenue code = 0173 or 0174) (Goodman et al. 2019). One thing to notice was that the assigned revenue code corresponds to the level of care determined during the clinical evaluation rather than the level of facility (Maryland Department of Health 2017; New York State Department of Health 2012) . As stated on the New York State’s Statewide Planning and Research Cooperative System (New York State Department of Health 2012), “The levels of care and resulting revenue codes may, and likely will, fluctuate during the infants stay in the facility.” The special care use was defined as the use of any level II nursery care (revenue code = 0172). I was also interested in examining the number of units used for these services, while given the difference in payment systems for Maryland (all-payer) versus other states (IPPS)¹, the units of some services may not be comparable. I did observe large differences in units of radiology service utilization. Therefore, indicators of services rather than the units were examined in the study. The revenue codes used to measure other services were listed in Table B.2 (Appendix B).

¹ Maryland’s all-payer rate-setting system used 51 revenue centers as the basis for payment (HSCRC 2012). In contrast, the IPPS categorized inpatient admissions into one of 746 Medicare Severity Diagnosis-Related Groups (CMS 2019d).

I also examined the total costs of care for each newborn and aggregate inpatient costs by the hospital department (i.e., cost center) (Table B.2, Appendix B) (Salemi et al. 2013). The definitions of terms charge, payment, and cost are different in hospital financial settings. The inpatient hospital charge was the price that the hospital billed for reimbursement purposes and it varied widely across the U.S. (Health Care Cost Institute 2019). The charges for neonatal and perinatal services also varied across states and hospitals (Hall et al. 2015). Also, the charges may not fully reflect the actual hospital services cost where the cost-to-charge ratios also varied across hospitals (HCUP 2019b). The actual payment that the hospital received from public and private payers were generally much lower than the listed charges/prices (Karaca and Moore 2013; Park, Kim, and Werner 2015). The HCUP SID files only contained the amount of charges, rather than payments or costs. In other words, I was unable to access the actual payment amount a hospital received from payers or the out-of-pocket expenditure of patients. But I was able to observe the amount of hospital charges and use the annual hospital-level cost-to-charge ratios to obtain the actual cost of hospital services. These costs reflect the actual hospital expenses including wages, supplies, and utilities, which could serve as a good proxy of healthcare spending (Riley 2009; Andrews 2015; HCUP 2019b).

The HCUP data only provide limited demographic information like other claims data. The covariates used in newborn analysis included infants' sex, race and ethnicity, birthweight, the primary payer, median household income as state quartile for patient zip code, and an urban-rural indicator (coded as rural if not metropolitan or micropolitan, urban otherwise). I also created the gestational age categories, an indicator of having respiratory distress syndrome, and an indicator of singleton birth using the diagnosis

codes (Table B.1, Appendix B). An indicator of having any congenital anomaly was also created using diagnosis codes from the Clinical Classification Software created by the HCUP, which collapsed diagnosis codes and procedure codes into a smaller number of clinically meaningful categories (Table B.3, Appendix B).

I also included time-varying state-level characteristics, including poverty rates, unemployment rates, birth rates, and the number of NICU beds per 1,000 residents to control for changes in the supply of NICU services.

3.2.4 Statistical Analysis

Linear regression models were primarily used for each outcome to facilitate complex adjustment for standard error. Non-OLS models were used as supplements according to the type of dependent variables. For count data (i.e. length of stay), Poisson models were used which had a better fit to the data compared to the negative binomial models. The model fit was estimated by comparing the mean differences of observed and predicted counts from these two models (Long and Freese 2014). Considering that the distribution of cost was skewed, OLS with a log transformation as well as the generalized linear model (GLM) with a log link function and the gamma distribution were used (Buntin and Zaslavsky 2004). Marginal effects were reported for non-OLS models.

I estimated the impact of GBR using the following equation (in OLS):

$$Y_{ist} = \delta(Treatment_s \cdot Post_t) + X_{ist}\beta + Z_{st}\beta_2 + \mathbf{u}_s + \boldsymbol{\gamma}_t + e_{ist}$$

Where Y_{ist} measured cost, length of stay, or service utilization for individual i in hospital s at year t . $Treatment_s$ was an indicator for Maryland patients. $Post_t$ was an indicator of post-implementation years. The coefficient of interest, $\hat{\delta}$, is the impact of GBR. \mathbf{u}_s was a vector of the hospital fixed effects and $\boldsymbol{\gamma}_t$ was a vector of the year fixed

effects. The hospital fixed effects controlled for the unobserved hospital-specific factors that were stable over time and the year fixed effects controlled for year-specific changes that were common for all hospitals. X_{ist} was a vector of individual-level and Z_{st} was a vector of state-by-year level covariates as described. Other non-OLS models shared the same control variables.

The standard errors were obtained from the clustered sandwich estimator to account for state clustering. However, that approach does not perform well when there was only one treated cluster (Ferman and Pinto 2019). To provide supporting evidence on the robustness of inference, I also reported P-values that were obtained from a bootstrap method that better accounted for serial correlation and heteroscedasticity in the presence of a single treated cluster (Ferman and Pinto 2019).

While the treatment effects may be heterogeneous across the conditional distribution of the outcomes which implied some potential of using a quantile regression analysis, I decided to conduct subgroup analysis instead. This is because the reason for high-cost or high utilization case related to birth and delivery were highly predictable by indicators such as gestational age or birthweight. That is, infants with high cost or high utilizations were more likely to be preterm or immature. The subgroup analysis, therefore, conveyed more information compared to quantile regression in this case. If I observed larger effect among those high-risk cases, I would expect to see larger effects in the tail when using a quantile regression. Therefore, I conducted subgroup analysis by infant gestational age.

I conducted several robustness checks. First, I conducted the DID estimates using 2014 as the treatment year, which was used in previous literature. Although I assume that

the effects of GBR were mainly driven by actual applied penalties which were started in 2015, the results using 2014 as the implementation year would facilitate a comparison to previous literature. Also, I conducted subgroup analysis by birthweight.

3.3 Results

3.3.1 Descriptive Statistics

Table 3.1 provides descriptive statistics for infants, with the bootstrap p-value from a balancing test. The balancing test was conducted by replacing the outcome variable with the covariate and fitted the standard linear DID regression model (Wing, Simon, and Bello-Gomez 2018). Then I examined whether the magnitude of delta (δ , the DID estimator) is small and not significant. There are some differences between Maryland and the comparison states at the baseline period. For example, Maryland has a higher percentage of non-Hispanic black infants, and a lower poverty rate. However, the strength of the DID design is that all observed and unobserved differences between Maryland and comparison hospitals are controlled by the hospital fixed effects in the model, as long as these differences are stable over time.

What matters for the validity of the DID is that the differences between the two groups are stable over time and that the changes in treatment exposure are not associated with changes in the distribution of covariates. If the covariates are differentially changing over time in Maryland hospitals versus the comparison state hospitals, this would not be controlled for by the hospital fixed effects. The final column of Table 3.1 presents the p-value from the balancing test which directly assesses if covariates differences between Maryland and the comparison states vary over time.

Table 3.1: Characteristics of Infants Before and After GBR, 2011-2016

| Characteristic, % | Maryland | | Comparison States | | Bootstrap P-value |
|--|---------------|---------------|-------------------|---------------|----------------------|
| | 2011- 2014 | 2015- 2016 | 2011- 2014 | 2015- 2016 | |
| Birthweight | | | | | |
| <1500g | 0.02 | 0.02 | 0.01 | 0.01 | 0.80 |
| 1500-2499g | 0.07 | 0.07 | 0.07 | 0.07 | 0.93 |
| >=2500g | 0.91 | 0.91 | 0.92 | 0.92 | 0.75 |
| Gestational Age | | | | | |
| <=32 weeks | 0.02 | 0.02 | 0.02 | 0.02 | 1.00 |
| 33-36 weeks | 0.07 | 0.07 | 0.06 | 0.06 | 0.60 |
| >=37 weeks | 0.91 | 0.91 | 0.92 | 0.92 | 0.67 |
| Female | 0.49 | 0.49 | 0.49 | 0.49 | 0.59 |
| Singleton Birth | 0.96 | 0.97 | 0.96 | 0.96 | 0.95 |
| Infant with Respiratory Distress Syndrome | 0.03 | 0.03 | 0.02 | 0.02 | 0.37 |
| Infant with Congenital Anomaly | 0.16 | 0.08 | 0.11 | 0.06 | 0.12 |
| Race | | | | | |
| Non-Hispanic White | 0.45 | 0.44 | 0.53 | 0.51 | 0.04 |
| Non-Hispanic Black | 0.31 | 0.31 | 0.13 | 0.13 | 0.44 |
| Non-Hispanic Other | 0.12 | 0.12 | 0.20 | 0.21 | 0.32 |
| Hispanic | 0.11 | 0.13 | 0.13 | 0.15 | 0.80 |
| Insurance Type | | | | | |
| Medicaid | 0.44 | 0.44 | 0.42 | 0.45 | 0.42 |
| Private Insurance | 0.52 | 0.51 | 0.52 | 0.48 | 0.04 |
| Self-Pay | 0.02 | 0.02 | 0.04 | 0.05 | 0.91 |
| Other | 0.02 | 0.03 | 0.02 | 0.02 | 0.36 |
| Urban | 0.99 | 0.99 | 0.96 | 0.96 | 0.59 |
| Median Household Income | | | | | |
| 1st Quartile | 0.27 | 0.28 | 0.31 | 0.31 | 0.56 |
| 2nd Quartile | 0.26 | 0.26 | 0.24 | 0.25 | 0.30 |
| 3rd Quartile | 0.27 | 0.26 | 0.24 | 0.23 | 0.96 |
| 4th Quartile | 0.20 | 0.20 | 0.21 | 0.21 | 0.65 |
| State Characteristics | | | | | |
| Poverty Rate | 8.00 | 7.50 | 12.71 | 11.98 | 0.40 |
| Unemployment Rate | 6.72 | 5.30 | 8.00 | 5.56 | 0.08 |
| Birth Rate | 12.32 | 12.12 | 12.11 | 11.82 | 0.37 |
| NICU Beds, per 1,000 Population | 0.07 | 0.07 | 0.04 | 0.04 | 0.67 |
| <i>N</i> | 258,977 | 132,148 | 1,426,420 | 721,040 | |

Note. The sample excludes those with missing values. Comparison states include New York, New Jersey, and Kentucky. P-values are obtained using a bootstrap approach, developed by Ferman and Pinto (2019).

I find that most of the covariates' differences are stable over time. The two exceptions are the proportions of infants being non-Hispanic White and covered by private insurance. For instance, Maryland and the comparison states both have 52% of births covered by private insurance before GBR but the decrease 1 percentage point after GBR in Maryland and 3 percentage points in the comparison states. Given that the magnitudes of these differences are very small and the significance level is on the margin, the effects of these compositional changes are assumed to be ignorable.

Table 3.2 presents the summary statistics for outcome variables for in Maryland at the baseline period. The average length of stay in Maryland before GBR is 3.71 days and the average cost of birth is \$3811.35 for infants. In addition, 7.66% of infants have used NICU care and 6.81% have used special care. The rates of radiology diagnostic and imaging service rates are 6.79% and 5.60%, respectively. The rates of using CT scan and MRI are relatively small, with 0.15% and 0.29%. Both the NICU care and special care may incorporate respiratory services where I observe a 13.35% utilization rate among all infants. The C-section rate is 32.69%.

The categories of gestation are following pediatric guidelines and literature, i.e., very-preterm (≤ 32 weeks), preterm (33-36 weeks), and term (≥ 37 weeks) (Kilpatrick, Papile, and Macones 2017; American Academy of Pediatrics 2012; W. N. Harrison, Wasserman, and Goodman 2018; W. Harrison and Goodman 2015). Only 2.19% of infants are very preterm (< 32 weeks) and the majority of infants are term births. As expected, both the LOS and total costs vary greatly across gestations, suggesting that the gestational age was a powerful indicator of the relative risk for infants. The average LOS is 36.74 days among preterm infants (< 32 weeks) while only 2.65 days among full-term

infants (≥ 37 weeks). The average total costs of births for the very-preterm infants are \$61,222.01 compared to only \$2,051.48 among term infants.

Table 3.2: Outcomes of Infants in Maryland at Baseline, 2011-2014

| Outcomes | Sample Size (%) | Mean |
|--------------------------------------|-----------------|-----------|
| Length of Stay, Days | | |
| <i>Overall</i> | 258,977 | 3.71 |
| <i>Infant Gestation</i> | | |
| <=32 weeks | 5,679 (2.19) | 36.74 |
| 33-36 weeks | 17,572 (6.79) | 7.27 |
| ≥ 37 weeks | 235,726 (91.02) | 2.65 |
| Total Cost, \$ | | |
| <i>Overall</i> | 258,977 | 3,811.35 |
| <i>Infant Gestation</i> | | |
| <=32 weeks | 5,679 (2.19) | 61,222.01 |
| 33-36 weeks | 17,572 (6.79) | 8,865.43 |
| ≥ 37 weeks | 235,726 (91.02) | 2,051.48 |
| Indicators of Services Use, % | | |
| NICU Care | 258,977 | 7.66 |
| Special Care | 258,977 | 6.81 |
| Radiology Diagnostic | 258,977 | 6.79 |
| CT Scan | 258,977 | 0.15 |
| Other Imaging Services | 258,977 | 5.60 |
| Magnetic Resonance Technology | 258,977 | 0.29 |
| Respiratory Services | 258,977 | 13.35 |
| Caesarean Section | 258,977 | 32.69 |

Note. The sample excludes those with missing values.

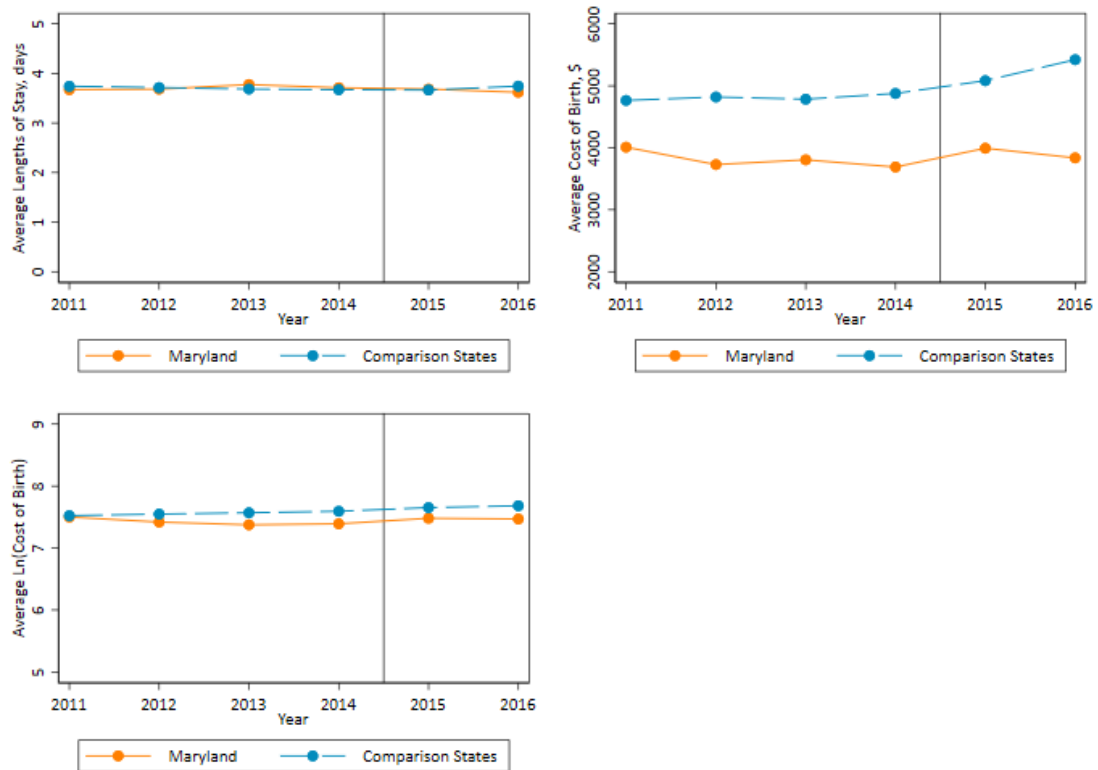
3.3.2 Pre-trends

The key assumption underlying the difference-in-difference method is that the outcomes of the treatment and control group would have followed the same trend had GBR never been implemented. Although this assumption cannot be tested directly, I followed the literature where I first examined the parallel trend visually and then conducted formal tests to assess differential trends prior to the GBR. Specifically, I used

the specification following the main model (without the state-specific linear trend) where I replaced the previous interaction term with an interaction term of an indicator of Maryland and a linear time trend, sub-setting to the pre-period.

Figure 3.1 plots the unadjusted means of LOS, cost, and rate of service utilization in Maryland and the comparison states, respectively. We can see from Figure 3.1 that trends in LOS and mean cost of birth are quite similar in Maryland and the comparison states. The log transfer of cost has slightly different trends.

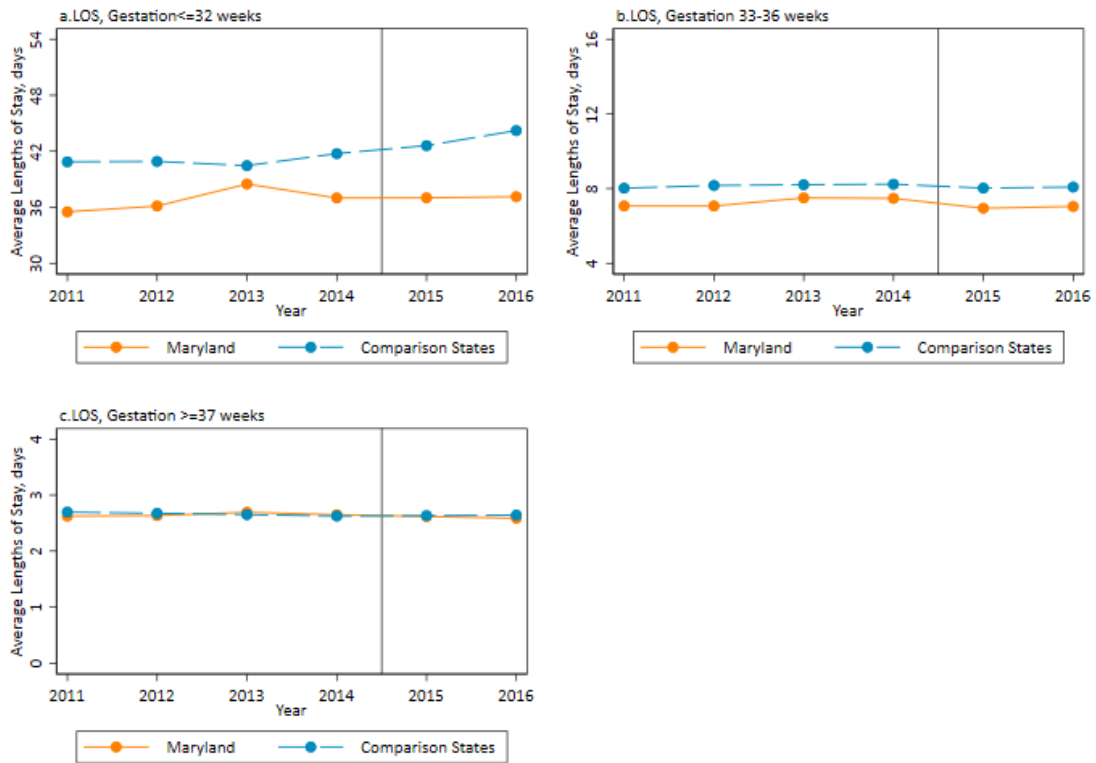
Figure 3.1: Unadjusted Trends of Cost and LOS for All Infants: Maryland versus Comparison States



Note. Comparison states include New York, New Jersey, and Kentucky.

Figure 3.2 depicts the mean LOS in Maryland and the comparison states by gestational age. The trends look similar in all three gestational age groups, except for an outlier in the year 2013 among very-preterm infants. The trends look unchanged after GBR among preterm and term infants, while increased faster in the comparison states than Maryland among the very-preterm infant group.

Figure 3.2: Unadjusted Trends of LOS for Infants by Gestational Age: Maryland versus Comparison States

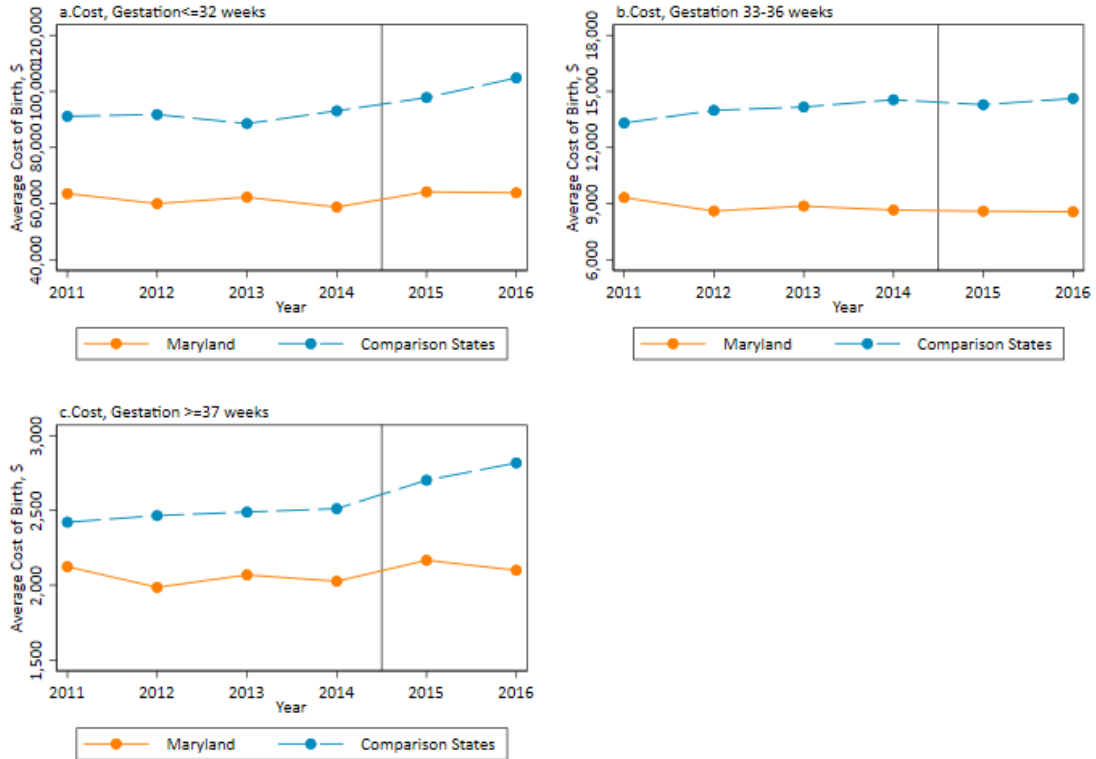


Note. Comparison states include New York, New Jersey, and Kentucky.

Figure 3.3 depicts the mean costs of births in Maryland and the comparison states by gestational age. Again, the trends look similar in all three gestational age groups. The

cost of birth seems unchanged after GBR for all Maryland infants, while it increased faster in the comparison states.

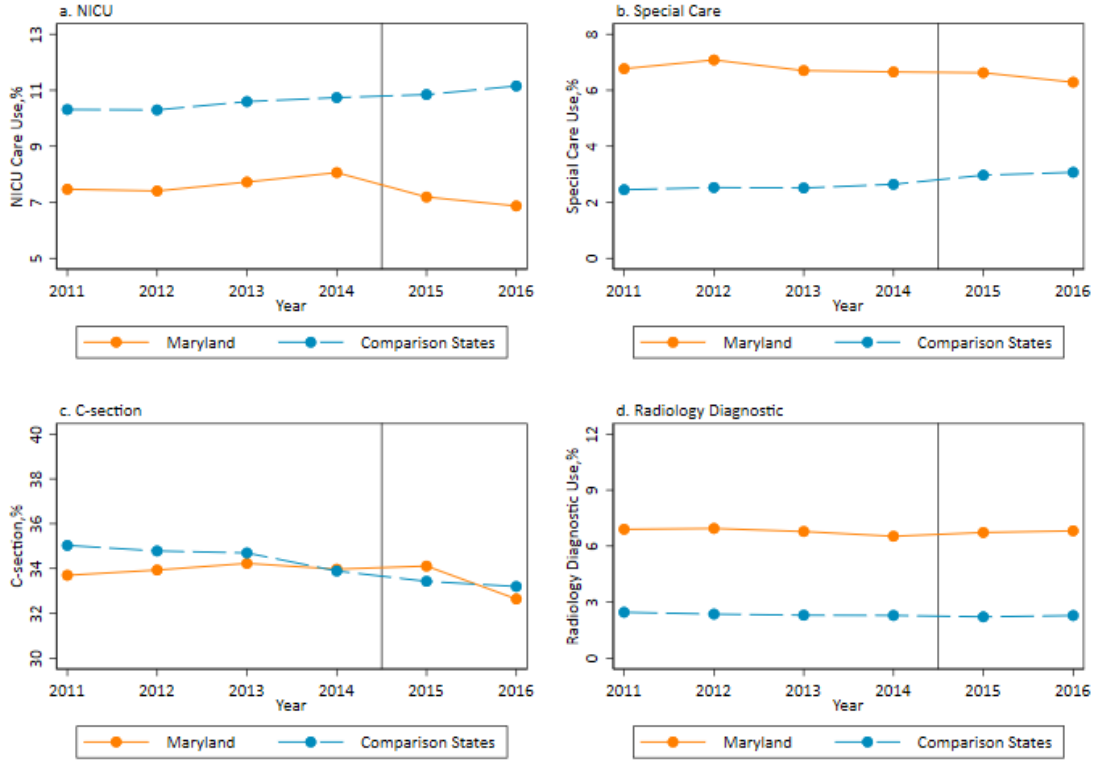
Figure 3.3: Unadjusted Trends of Cost for Infants by Gestational Age: Maryland versus Comparison States

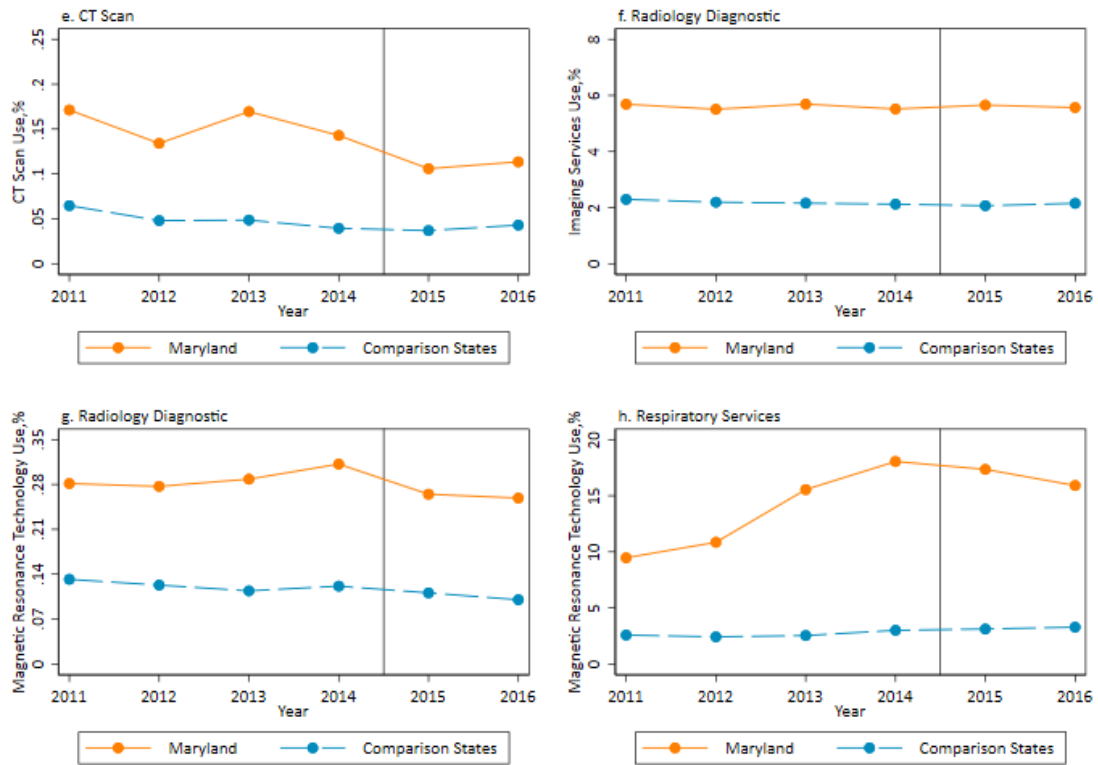


Note. Comparison states include New York, New Jersey, and Kentucky.

Figure 3.4 depicts the unadjusted trends of the utilization rate of each service in Maryland and the comparison states among all births. While visual inspections suggested that prior trends are mostly similar in Maryland and the comparison states, the parallel trends assumption may be violated among some cases, such as the C-section rate (Figure 3.4-c) or respiratory services (Figure 3.4-h). These potential violations are reassured by regression results.

Figure 3.4: Unadjusted Trends of Services Utilization for Infants: Maryland versus Comparison States





Note. Comparison states include New York, New Jersey, and Kentucky.

Table 3.3 presents the results of the pre-implementation trends test for all outcomes and subgroups. The tests for services utilization are mostly passed, except for respiratory services. Unfortunately, there are violations among overall group LOS and cost. The parallel trends assumption is also violated among infants with gestational age ≥ 37 weeks, which consists of the majority of all births. Therefore, I reported results from the models with state-specific linear trends for those groups who violated the parallel trends assumption.

Table 3.3: Differential Pre-Trends Test of Infants, 2011-2014

| Outcomes | Coef. | 95% CIs | P-value |
|--------------------------------------|-------|---------------|---------|
| Length of Stay, Days | | | |
| <i>Overall</i> | 0.07 | (0.05,0.08) | 0.001 |
| <i>Infant Gestation</i> | | | |
| <=32 weeks | 0.06 | (-1.42,1.54) | 0.9 |
| 33-36 weeks | 0.06 | (-0.19,0.31) | 0.5 |
| >=37 weeks | 0.03 | (0.01,0.05) | 0.015 |
| Log (Cost) | | | |
| <i>Overall</i> | -0.09 | (-0.14,-0.03) | 0.013 |
| <i>Infant Gestation</i> | | | |
| <=32 weeks | -0.06 | (-0.13,0.01) | 0.076 |
| 33-36 weeks | -0.09 | (-0.18,-0.01) | 0.038 |
| >=37 weeks | -0.09 | (-0.14,-0.04) | 0.01 |
| Indicators of Services Use, % | | | |
| NICU Care | -0.1 | (-0.56,0.36) | 0.526 |
| Special Care | 0.18 | (-0.26,0.61) | 0.285 |
| Radiology Diagnostic | -0.14 | (-0.33,0.06) | 0.115 |
| CT Scan | -0.01 | (-0.03,0.01) | 0.169 |
| Other Imaging Services | -0.26 | (-0.53,0.00) | 0.052 |
| Magnetic Resonance Imaging | 0.01 | (-0.00,0.02) | 0.155 |
| Respiratory Services | 2.36 | (1.30,3.42) | 0.006 |
| Caesarean Section | 0.19 | (-0.02, 0.40) | 0.064 |

Note. Standard errors in the parentheses are clustered at the state level. Models control for birthweight, gestational age, sex, race, insurance type, relative household income, an indicator of congenital anomalies, an indicator of respiratory distress syndrome, urban/rural; state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects.

3.3.3 Effect of GBR on LOS

I begin by considering the effect of GBR on birth-related length of stay. Regression estimates of OLS and Poisson models for all infants are displayed in Table 3.4. Marginal effects are reported for Poisson regressions. For those with violations of parallel trend assumption, I present the results with and without state-specific linear trends. Inconsistent results between these two groups should be interpreted with caution.

Table 3.4: Effect of GBR on Length of Stay, 2011-2016

| | OLS | | | | Poisson | | | |
|--|-----------------|-----------------|-----------------|------------------|-----------------|------------------|-----------------|-------------------|
| | Overall | <=32 weeks | 33-36 weeks | >=37 weeks | Overall | <=32 weeks | 33-36 weeks | >=37 weeks |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| <i>Without State-Specific Linear Trend</i> | | | | | | | | |
| DID Point Estimates | -0.01 (0.03) | -1.38 (0.51) | -0.07 (0.08) | -0.01 (0.01) | -0.03 (0.03) | -1.22* (0.54) | -0.04 (0.08) | -0.01 (0.01) |
| Ferman-Pinto P-value | 0.89 | <0.001 | <0.001 | <0.001 | | | | |
| <i>With State-Specific Linear Trend</i> | | | | | | | | |
| DID Point Estimates | -0.1 (0.06) | NA | NA | -0.06* (0.02) | -0.08 (0.05) | NA | NA | -0.06** (0.02) |
| Ferman-Pinto P-value | <0.001 | | | <0.001 | | | | |
| Baseline Mean in Maryland | 3.71 | 36.74 | 7.27 | 2.65 | 3.71 | 36.74 | 7.27 | 2.65 |
| N | 2,538,585 | 48,228 | 154,961 | 2,335,396 | 2,538,585 | 48,228 | 154,961 | 2,335,396 |

Note. Standard errors are in the parentheses, clustered at the state level. Models control for birthweight, gestational age, sex, race, insurance type, relative household income, an indicator of congenital anomalies, an indicator of respiratory distress syndrome, urban/rural; state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects. Marginal effects are reported for Poisson models. *** p<0.001, ** p<0.01, * p<0.05

During the two full years' implementation of GBR, I find that average LOS decreased by 0.1 days (bootstrap p-value<0.001) using the OLS model which corresponds up to a 2.7 percent decrease in Maryland relative to the comparison states, adjusting for the state-specific linear trends. But I do not find a similar decrease using the Poisson model. Among infants with gestational age ≤ 32 weeks, the decreases are consistent in both OLS and Poisson models by 1.38 (bootstrap p-value<0.001) or 1.22 days that correspond to a 3.8 or 3.3 percent decrease in Maryland compared to comparison states after GBR. Among infants with gestational age between 33 and 36 weeks, there is a small 0.07 days decrease (bootstrap p-value<0.001). The average LOS decreased by 0.06 days which corresponds to a 2.3 percent decrease in Maryland compared to the comparison states after GBR.

3.3.4 Effect of GBR on Cost of Birth

Next, I study the effect of GBR on costs of birth for infants. Marginal effects are reported for GLM regressions. Log transformation is used for the outcomes in OLS regressions.

Regression estimates of OLS and GLM models for all infants are displayed in Table 3.5. I begin by describing the results from models that do not include state-specific linear trends. My results OLS suggests a 4.0% (i.e. $\exp(-0.04)=0.96$) decrease in the cost of birth in Maryland compared to that in the comparison states after the implementation of GBR. The GLM results suggest a 4.7% decreases. The OLS and GLM models suggest similar treatment effect magnitudes as the decline is larger among very-preterm infants compared to preterm and term infants.

Table 3.5: Effect of GBR on Total Cost of Birth, 2011-2016

| | OLS | | | | GLM | | | |
|--|------------------|------------------|------------------|-------------------|----------------------|------------------------|------------------------|----------------------|
| | Overall | <=32 weeks | 33-36 weeks | >=37 weeks | Overall | <=32 weeks | 33-36 weeks | >=37 weeks |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| <i>Without State-Specific Linear Trend</i> | | | | | | | | |
| DID Point Estimates | -0.04 (0.02) | -0.06 (0.03) | -0.03 (0.02) | -0.04 (0.02) | -180.55* (86.58) | -4403.08* (1901.57) | -408.37* (203.95) | -92.31* (41.83) |
| Ferman-Pinto P-Value | <0.001 | <0.001 | <0.001 | <0.001 | | | | |
| <i>With State-Specific Linear Trend</i> | | | | | | | | |
| DID Point Estimates | 0.14** (0.01) | NA | 0.15* (0.04) | 0.14*** (0.01) | 684.48*** (74.74) | NA | 1691.53*** (491.32) | 334.10*** (25.08) |
| Ferman-Pinto P-Value | <0.001 | | <0.001 | <0.001 | | | | |
| Baseline Mean in Maryland | 3,811.35 | 61,222.01 | 8,865.43 | 2,051.48 | 3,811.35 | 61,222.01 | 8,865.43 | 2,051.48 |
| N | 2,538,585 | 48,228 | 154,961 | 2,335,396 | 2,538,585 | 48,228 | 154,961 | 2,335,396 |

Note. Standard errors are in the parentheses, clustered at state-level. Models control for birthweight, gestational age, sex, race, insurance type, relative household income, an indicator of congenital anomalies, an indicator of respiratory distress syndrome, urban/rural; state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects. Marginal effects are reported for GLM models. N denotes the number of observations.*** p<0.001, ** p<0.01, * p<0.05

The inclusion of state-specific trends leads to results with the opposite sign. The OLS outcome suggests a 15.0% (i.e. $\exp(0.14)=1.15$) increase in costs of birth, and similar results are obtained by GLM (17.9%). Differences between the model with and without linear trends can cast doubt on the validity of the study design. While such differences are concerning, they are consistent with Roberts et al., who studied the effects of GBR on hospital and primary care use and found that the sign of the treatment effect differed when including or excluding linear trends. Much like Roberts et al., the validity of either model depends on which unobserved assumption is actually at play. Given these differences, I cannot come to firm conclusions about the causal effects of GBR on hospital-based costs for infants.

3.3.5 Effect of GBR on Service Utilization

In Table 3.6, I study the effect of GBR on service utilization patterns. Given that the outcomes in columns 1-6 and 8 passed the pre-trend tests at the 5% significance level, I used a model without state-specific linear trends for these outcomes. I find a 0.6 percentage points decrease in NICU care, which corresponds to a 7.8 percent decrease after GBR in Maryland compared to the comparison states. There's a 0.76 percentage points (i.e., 11.2%) decrease in special care, which is a lower level neonatal care compared to NICU that provides respiratory services and mainly serves infants with relatively lower sickness. In addition, utilization for CT scans decreased significantly by 20.0%. The respiratory services use are not consistently estimated using models with or without state-specific linear trends since there's a 24.4% decline when adding the trends.

Table 3.6: Effect of GBR on Services Utilization, 2011-2016

| | NICU Care | Special Care | Radiology Diagnostic | CT Scan | Other Imaging Services | Magnetic Resonance Imaging | Respiratory Services | Caesarean Section |
|--|--------------|-----------------|-------------------------|------------------|------------------------------|----------------------------------|-------------------------|----------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| <i>Without State-Specific Linear Trend</i> | | | | | | | | |
| DID Point Estimates | -0.60* | -0.76** | -0.04 | -0.03* | -0.15 | -0.03 | 1.88* | 0.1 |
| | (0.14) | (0.10) | (0.14) | (0.01) | (0.20) | (0.02) | (0.39) | (0.14) |
| Ferman-Pinto P-Value | 0.02 | 0.02 | 0.80 | <0.001 | 0.65 | 0.45 | <0.001 | 0.49 |
| <i>With State-Specific Linear Trend</i> | | | | | | | | |
| DID Point Estimates | NA | NA | NA | NA | NA | NA | -2.86** | NA |
| | | | | | | | (0.41) | |
| Ferman-Pinto P-Value | | | | | | | <0.001 | |
| Baseline Mean in Maryland | 7.66 | 6.81 | 6.79 | 0.15 | 5.60 | 0.29 | 13.35 | 32.69 |

Note. N=2,538,585. Standard errors are in the parentheses, clustered at the state-level. Models control for birthweight, gestational age, sex, race, insurance type, relative household income, an indicator of congenital anomalies, an indicator of respiratory distress syndrome, urban/rural; state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects. N denotes the number of observations. *** p<0.001, ** p<0.01, * p<0.05

3.3.6 Cost Estimation by Hospital Department

The analyses of service utilization and cost in further detail are always favorable. Although the examination of all types of services might not be feasible in one paper, I show in Table 3.7 that the majority (73.7%) of inpatient costs of births in Maryland come from the nursery department. My analyses on the NICU and special unit care, therefore, captured the big picture of neonatal hospital care. Given different hospital payment settings, the distributions of services use and the cost by the department are different in Maryland, compared to the comparison states (Table 3.7), which impede my ability to conduct a DID analysis for each department separately. I show some descriptive results in this section instead, to help better understand the distributions and patterns of the birth-related service use and cost at the aggregate level, by comparing Maryland and the other comparison states.

Table 3.7: Aggregate Inpatient Hospitalization Costs Attributable to Each Department, 2011-2016

| Department, % | Maryland | Comparison State |
|---------------------|----------|------------------|
| Clinic | 0.00 | 0.05 |
| Special Care Units | 1.09 | 0.54 |
| Routine Bed Units | 1.18 | 0.37 |
| Nursery | 73.70 | 84.84 |
| All Other Ancillary | 0.56 | 0.58 |
| Operating Room | 0.77 | 0.93 |
| Therapies | 6.23 | 2.35 |
| Pharmacy | 5.58 | 3.34 |
| Laboratory | 10.89 | 7.02 |

Note. Comparison states include New York, New Jersey, and Kentucky

Figure 3.5 presents the prevalence of service utilization for infants by each department. For example, almost every infant has been billed on nursery care (i.e., the

prevalence is around 100%), while very few of them use special care units (i.e., the prevalence is below 10%). We can see that in Maryland, the prevalence of therapy services increases substantially before GBR, while it started to decrease after GBR. Service utilization in other departments is relatively stable over time. In the comparison states, there is a rapid increase in pharmacy department use and a slow increase in services within the therapies department. The distributions across the department are also different between Maryland and the comparison states where Maryland has a higher prevalence of billing into the therapies department.

Furthermore, I examined the distribution of the average cost of birth attribute to each department (Figure 3.6). The main cost comes from the department of nursery both in Maryland and the comparison states, which is also suggested in Table 3.7. The average costs attributed to the nursery department has a descending pattern in Maryland before GBR and starts to increase after GBR, while the nursery cost in other comparison states keeps increasing over time.

Figure 3.5: Prevalence of Service Utilization by the Hospital Department, 2011-2016

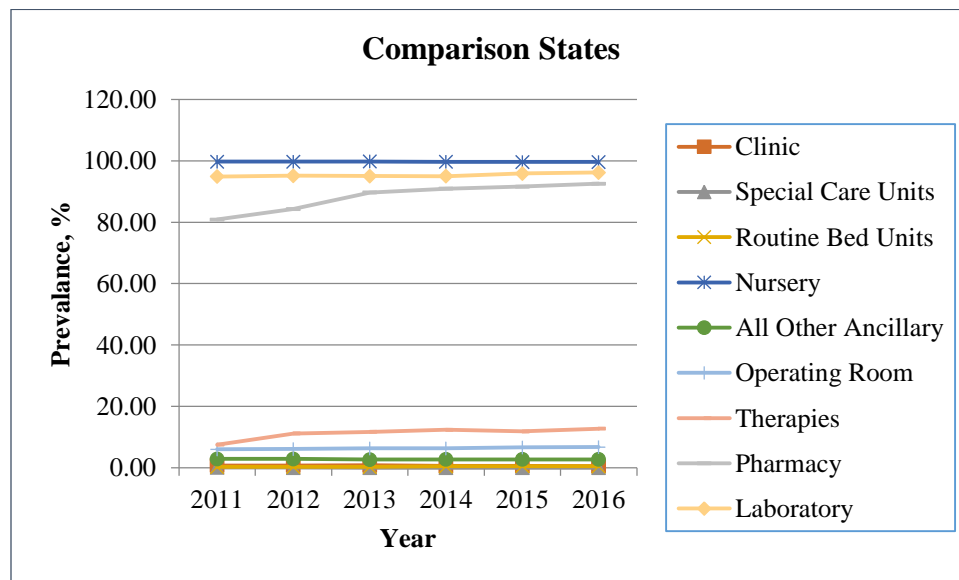
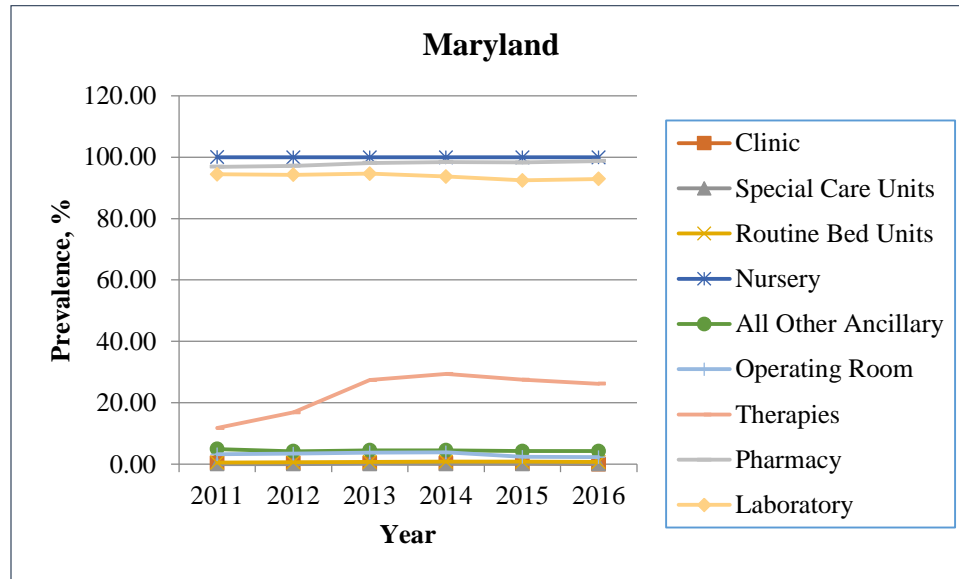
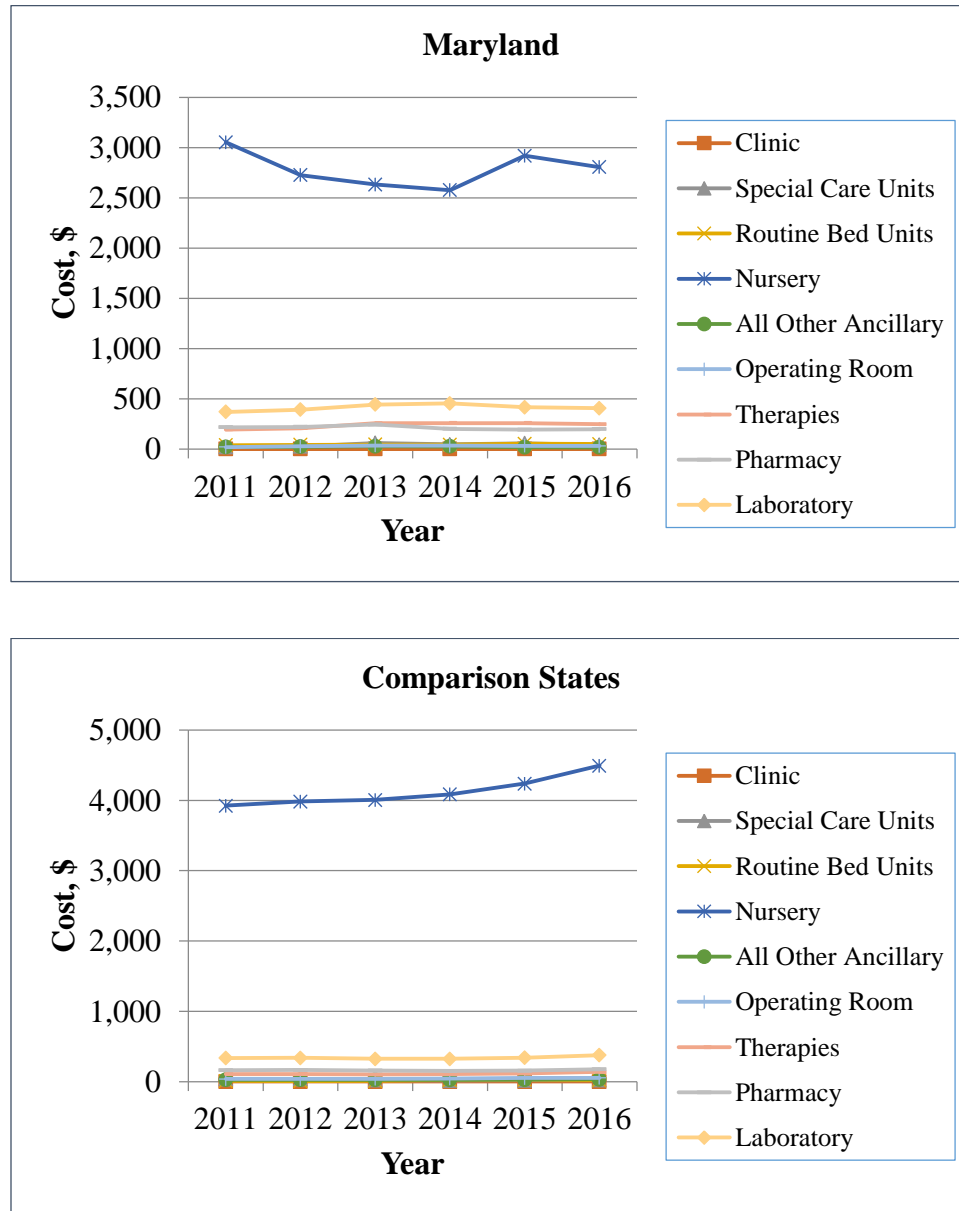


Figure 3.6: Distribution of Total Cost of Birth by Department for Infants, 2011-2016



3.3.7 Sensitivity Analyses

The results of the sensitivity analyses supported my main findings (Appendix B). The effects of GBR using 2014 as the implementation year are larger compared to using the year of 2015. Specifically, I find that GBR in Maryland led to a 7% decline in average LOS for infants, which are mainly driven by the decrease among preterm (<37

weeks) infants who used to have longer stays compared to term (≥ 37 weeks) infants. In addition, I find significant declines in the use of NICU care, special care, CT scan, MRI, and respiratory services. The results using birthweight also came to a similar conclusion using gestational age.

3.4 Discussion

In this paper, I study the hospitals' responses to a statewide global budget program in Maryland by focusing on birth-related services. I find that the GBR program in Maryland led to a 2.7% decline in average LOS for infants, which are mainly driven by the decrease among very-preterm (< 32 weeks) infants who used to have longer stays compared to other infants. In addition, I find significant declines in the use of NICU care (7.8%), special unit care (11.2%), and CT scan (20.0%). Assuming parallel trends, I find a significant decrease in costs of birth by 4% among all births and by 6% among very-preterm births; but assuming differential trends, I find a 14% increase in the total cost of birth for the overall sample.

Under GBR, all the acute-care hospitals in Maryland were assigned an annual revenue cap since 2014. Hospitals faced the risk of a reduced budget in the next year if its expenditures are not aligned with the budget in the current year starting July 2015, regardless of underage or overage. This program creates meaningful incentives for hospitals to manage their utilization, price, and expenditure. In particular, hospitals now have the incentive to reduce intensive margins where they can maintain the same revenue by raising the unit price, which in turn may rise the operating margins (Giuriceo et al. 2016). This is supported by my findings, where I find a significant decrease in LOS and NICU services use, but no significant change in the total cost of care.

3.5 Conclusion

I find that GBR was associated with a decline in total LOS and utilization of a series of neonatal services for newborns, while no consistently significant change in the cost of birth. These findings are a supplement to current findings of GBR on hospital utilization and expenditure, providing encouraging evidence on the effectiveness of a global hospital budget model on reducing unnecessary volumes, and also shed light on an essential population whose healthcare is mainly financed by Medicaid and private insurance.

Chapter 4 : Variation in NICU Admissions across Insurance Type

4.1 Introduction

The neonatal intensive care unit (NICU), defined by CDC's birth registration program as a facility that is "staffed and equipped to provide continuous mechanical ventilatory support", provides highly specialized care to newborns (Martin and Menacker 2004). NICU care is highly valuable and has been linked to substantial reductions in infant mortality for infants that need it, typically infants who are born preterm or very low birthweight. Along with the development of NICUs, the infant mortality rate in the US has declined from 26 per 1,000 live births in the 1960s to 5.9 per 1,000 live births in 2016.

NICUs are also costly and highly profitable facilities, serving as one of the major profit centers for hospitals. A NICU-stay costs approximately \$56,000 for commercial members and \$39,000 among Medicaid beneficiaries (David C. Goodman, George A. Little, Wade N. Harrison, Atle Moen, Meredith E. Mowitz, Cecilia Ganduglia-Cazaban, Kristen K. Bronner 2019). A 2010 Health Affairs article that profiled one academic medical center's efforts found that NICU admissions made up just for 4% of total hospital admissions but accounted for 69% of net profits (Lantos 2010). Not surprisingly, the marriage of high effectiveness and high profitability led to the rapid growth of the utilization of NICUs, attaining over a 20% increase from 2007-2012 (W. Harrison and

Goodman 2015). Hence, there are compelling reasons for researchers to investigate the financial incentives related to NICU care and the consequences of its fast expansion.

There are at least three strands of evidence that suggest that admissions among relatively low-risk (those with higher birthweight) infants are particularly sensitive to financial incentives. The first is the variation of the NICU admission rate across geography and hospitals, controlling for infant's birthweight (a key measurement of infant risk). For example, a report of the Dartmouth Atlas Project found that there was no regional variation of NICU admission among very-low birthweight infants (<1500g), while it varied 3 times for moderately-low birthweight infants (1500-2499g) and up to 5 times among those normal birthweight infants. Second, the composition of NICU admitted infants has changed over time. Specifically, from 2007 to 2017, among admitted infants, the percent of very-low birthweight infants decreased from 16.1% to 12.7%, while the percent of normal birthweight infants increased from 42.2% to 48.0%. Third, NICU admissions were found to be correlated with the unit census and bed supply; and these correlations were concentrated among infants who do not possess clear clinical indicators of need (W. N. Harrison, Wasserman, and Goodman 2018; Freedman 2016a; J. Schulman et al. 2018). For example, Freedman (2016) exploits exogenous capacity variation documenting a causal impact of empty beds on NICU admission among infants with higher birthweight (Freedman 2016a).

Insurance plays an important role in reimbursement and hospital financing and is a key lever in hospital financial incentives. Previous studies suggest that more neonatal services are used among newborns covered by private insurance, compared with Medicaid or the uninsured (Braveman et al. 1991; Currie and Gruber 2001). Given that

private payers have higher unit prices, this pattern is consistent with hospitals oversupplying NICU services to private pay patients, relative to patient need, or undersupplying it to publically funded patients. Medicaid expansion has also been shown to affect hospitals' decisions to adopt the NICU, which was attributed to the increase in the relative price of Medicaid compared to private (Freedman, Lin, and Simon 2015).

While these studies exploit sound identification strategies, the data used in these studies are from decades ago such that conclusions from these studies may not reflect the recent rapid change of NICU care. Little is known about the variation in NICU care across the type of payers in the contemporary setting.

This descriptive study uses birth certificate data to explore the variation of NICU use across insurance payers and to examine how much such variation could be explained by infant risk factors, maternal characteristics, and state characteristics.

4.2 Background

4.2.1 The Development of NICU

The infant mortality rate (death within the first year of life) in the United States has declined from 26.0 per 1,000 live births on average in the 1960s to 9.2 in the 1990s to 5.9 in 2016 (Xu et al. 2018). The decline that happened during the 1960s to the 1990s was mainly driven by the drop in neonatal infant mortality (under 28 days from birth) among preterm infants (Cutler and Meara 2000). From 1950 to 1990, mortality rates of very low birthweight infants (VLBW) (<1500g) declined by 42 percent, and mortality rates of low birthweight infants (LBW) (1500g-2500g) declined 7.5 percent in total

(Cutler and McClellan 2001)². These reductions of mortality were accomplished through the diffusion of technologies related to birth and the corresponding development of NICUs (Cutler and McClellan 2001).

In the first half of the 20th century, the incubator was the only "high tech" treatment for preterm infants (Jorgensen 2012). With the development of the ventilator for infants, the first modern NICU opened in the 1960s, with ancillary facilities to maintain temperature and nutrition and obstetric monitoring facilities. In the 1970s, the major innovation of ventilators happened, and the American Board of Pediatrics developed a Sub-Board on Neonatal-Perinatal Medicine (NPM) to further strengthen the specialty of neonatal care. In the 1980s-1990s, major innovations, such as antenatal corticosteroid treatment, tocolytics, high-speed, high-frequency ventilation, and the use of surfactant (approved by FDA which was believed to significantly help the prevention of respiratory distress syndrome (RDS) in premature infants with high risks of RDS), were adopted in neonatal intensive care (Cutler and Meara 2000).

In 2004, the American Academy of Pediatrics (AAP) designated 3 levels of neonatal care (American Academy of Pediatrics 2004). Level I provided basic care which is required of all inpatient maternity facilities, Level II provided specialty care for moderately ill newborns, and Level III provided subspecialty care for severely ill newborns with three subdivisions based on degrees of complexity and risk. However, there was great heterogeneity in applying this classification, and as a result, some VLBW infants were not sent to level III hospitals as expected. The classification lacked detailed

² The average birthweight among singleton term birth decreases from 1990-2005 (Donahue et al. 2010), which results in mortality for very-low birthweight infants (<1500g) increasing from 1983 to 2005 (Lau et al. 2013).

and standardized instructions, and until 2009, only 5 states had at least 90% of VLBW infants delivered to high-risk (level III) facilities (American Academy of Pediatrics 2012).

As a result, in 2009, the AAP updated the levels of neonatal care into 4 levels that consist of basic care (level I), specialty care (level II), and subspecialty intensive care (level III and level IV) with detailed information on functional criteria (e.g., birthweight and gestational age), capabilities (i.e., physical space, equipment, technology, and organization) and provider types (American Academy of Pediatrics 2012). Specifically, level I facilities, i.e., well newborn nurseries, provide the basic level of care to infants who are at low risk. They can also care for preterm infants at 35 to 37 weeks of gestation with certain conditions and late preterm infants at 34 to 36 weeks of gestation with further consideration. Level II facilities, i.e., special care nurseries, provide care for infants ≥ 32 weeks and ≥ 1500 g while they are stable or moderately ill with conditions, provide mechanical ventilation for less than 24 hours duration and/or continuous positive airway pressure, and provide care for those after intensive care. Level III, i.e., NICU, with extra capabilities based on level II, provide care for infants of < 32 weeks' gestation and weighing < 1500 g or with a critical illness. The designation of this level of care should be based on the clinical experience of certain specialists with expertise in neonatology. Level III NICUs are required to equip continuously available personnel (neonatologists, neonatal nurses, respiratory therapists) and equipment to provide sustained life support. It provides a full range of respiratory support, including conventional and/or high-frequency ventilation and inhaled nitric oxide. In addition, these units are required to have the capability to perform major surgery and advanced imaging such as CT scan, MRI, etc. Level IV facilities, i.e., regional NICUs, with extra

capabilities based on level III, provide care to the most complex and critically ill newborns as a regional center. These units are usually located within an institution, with additional capabilities to perform surgical repair of the complex condition and to equip a full range of pediatric subspecialists on-site, have easy transportation, and provide outreach education to keep pace with the latest knowledge.

4.2.2 The Current Status of NICU

Given the economic benefit of hospital constructing a NICU (regardless of size and capacity), a "deregionalization" trend to NICU occurred in the 1980s to 1990s, which expanded NICUs from originally large, regional hospital setting to smaller, community hospitals offering less sophisticated care (Schwartz, Kellogg, and Muri 2000; S A Lorch, Maheshwari, and Even-Shoshan 2012). While the number of births increased only by 17.6% from 1980 to 1995, the number of hospitals with a NICU increased by 98.9% and the number of NICU beds increased by 137.9% (Howell et al. 2002). In addition, 89% of new NICUs opened at that time were low-level NICUs, contrasted to 46% before 1980 (Baker and Phibbs 2002). There were also some geographic variations on NICU adoption. A perinatal survey in 2008 from AAP revealed that numbers of NICUs across states vary from 0 to 33, and numbers of NICU beds vary from 76 to 754 (S A Lorch, Maheshwari, and Even-Shoshan 2012). As of 2009, 37 states and District of Columbia with a certificate-of-need program (which sought to reduce expansions to healthcare infrastructure) had fewer hospitals with NICUs and fewer NICU beds in the hospitals (S A Lorch, Maheshwari, and Even-Shoshan 2012).

In 2012, 8.5 percent of newborns were admitted to a NICU. The units were extensively used for VLBW infants that were admitted at a rate of 844.1 per 1,000 live

births, compared with 43 per 1,000 of normal-birth-weight infants (2500g-4000g) (W. Harrison and Goodman 2015). Admission rates to NICUs increased from 64 to 77.9 per 1,000 live births during the six-year study period from 2007 to 2012 across the US (W. Harrison and Goodman 2015).

The increase of newborns admitted into NICUs could be solely due to medical reasons. For example, advancements in obstetric care might result in lowering the gestational age of viability. Thus, we might expect to see an increase in NICU services simply because there are more infants born alive that might benefit from NICU care. However, Harrison and Goodman (2015) showed that newborns who were admitted into a NICU were larger and less premature from year to year (W. Harrison and Goodman 2015), suggesting that the increasing admissions cannot be explained by an increase in the number of live but sick births. Schulman et al. (2018) found a 34-fold variation exists in admission rates across 130 NICUs in California, which cannot be fully explained by illness acuity. The NICU admission rate was negatively correlated with the percentage of admissions with a high degree of illness (J. Schulman et al. 2018).

About 4 million babies are born in the United States each year (CDC 2016a). As of 2011, births accounted for 10 percent of all inpatient discharges (Kowlessar, Jiang, and Steiner 2013). In 2009, the total cost of birth hospitalization and re-hospitalization was over \$13 billion (Barradas et al. 2016). The 9.1% of all births who are preterm and/or low birth weight accounted for 43.4% of the total cost (Barradas et al. 2016). Compared with a normal birth that costs up to \$3,200 for three-day stays, preterm newborns (gestational age < 37 weeks) cost \$21,500 and the average length of stay is 14.3 days; newborns who are low weight (1500g-2500g) and very low weight (< 1500g) cost on

average \$14,000 and \$76,700 respectively, and the average length of stay is 10.9 and 42.6 days, respectively. Newborns who have complications such as respiratory distress syndrome might cost \$54,900, and the length of stays extend to 31.3 days (Kowlessar, Jiang, and Steiner 2013). The cost can reach a maximum of nearly \$546,000 for very sick infants using the 2001 National Inpatient Sample from HCUP (Russell et al. 2007). Even though the prevalence of certain complications and conditions can be very rare, the chances of having a preterm and low birth weight baby are 8.5 percent and 6.1 percent, respectively (Kowlessar, Jiang, and Steiner 2013).

Costs for birth have grown from year to year. The cost per low birthweight (<2500g) infant was \$6,850 in the 1970s and increased to \$11,048 in 1988 (both measured in 1990 dollars) (Cutler and Meara 2000). There are also subsequent costs for caring for low birthweight babies with disabilities and other developmental difficulties. Cutler et al. (2000) estimated that the total spending for children with health problems, including costs of medical, benefit payments, and special education, was \$4,136 in 1960 and \$8,271 in 1990, as measured in 1990 dollars (Cutler and Meara 2000). And Kuo et al. (2018) found that most post-NICU spending occurred within the first year after discharge which is \$33,276 per person per year using newborn data from 2010 (Kuo et al. 2018).

4.3 Methods

4.3.1 Data

Data from the 2016-2017 restricted-use natality files of the Vital Statistics (VS) was used (CDC 2003). The natality files have information on nearly all births in the US. Birth certificates are collected annually through the U.S. Standard Certificate of Live

Birth, collecting mothers' and infants' information on demographic characteristics, medical and public program utilization, and health characteristics. (CDC 2003). NICU admission status was first available in seven states in 2004 using the 2003 revised version of US Birth Certificates. The adoption of this new version of the birth certificate expanded to all states, the District of Columbia, and territories as of January 1, 2016 (CDC 2016b). Given that NICU admission status was first available in 2016 for all states and 2017 was the latest year at the time of data application for this dissertation, only two years of data were used in this analysis. The birth file was merged with state-by-year characteristics, obtained from State Health Compare, Area Health Resources Files, Kaiser Family Foundation, and CDC Wonder (State Health Compare 2019; HRSA 2019; KFF 2017; CDC 2018).

The sample included all newborns in all 50 states and DC. Births in territories or with missing values in any covariates were excluded. Infants weighing less than 500g were also excluded from the final sample, given they were usually not considered as live births (W. N. Harrison, Wasserman, and Goodman 2018; W. Harrison and Goodman 2015).

I observed 7,809,667 births during the years 2016 and 2017 in 50 states and DC. Observations that were missing in any covariates were dropped, accounting for 3.36% (262,347) of total births. The final sample was 7,547,320 births.

4.3.2 Study Variables

The outcome variable was an indicator of NICU admission status. NICU admission in the Birth Files is defined as “admission into a facility or unit staffed and equipped to provide continuous mechanical ventilatory support for the newborn”.

The key independent variable was the principal source of payment for this delivery that was categorized into Medicaid, private insurance, self-pay, and other sources (including Indian Health Services, CHAMPUS/TRICARE, other government, and any other sources). The covariates of infant's characteristics included: birthweight, gestational age, gender, 5 min Apgar score, plurality, an indicator of small for gestational age (SGA), and an indicator of any congenital anomaly; maternal characteristics included race, age, education, parity, and indicators of maternal pregnancy risks, infections, and morbidity; state characteristics included poverty rate, unemployment rate, birth rate, NICU bed per 1,000 population, and prescription rates of Oxycodone and Hydrocodone. The birthweight was categorized into following groups: very-low birthweight (500g-1,499g), moderately-low birthweight (1,500g-2,499g), and normal birthweight ($\geq 2,500$ g). These categories were widely used in the pediatric literature and guideline which facilitates the comparison to previous studies (Kilpatrick, Papile, and Macones 2017; David C. Goodman, George A. Little, Wade N. Harrison, Atle Moen, Meredith E. Mowitz, Cecilia Ganduglia-Cazaban, Kristen K. Bronner 2019). Based on AAP's perinatal care guidelines, all infants who weighing less than 1,500g should be sent into NICUs while admissions of all other infants depend on infant's health condition and were decided by physicians. Infants weighing more than 2,500g were commonly considered as normal births and faced lower risks.

4.3.3 Statistical Analysis

Using single and multivariable linear probability models, I examined whether NICU admission rate is associated with insurance type with and without adjusting for covariates and state and year fixed effects. The adjusted models controlled for different

characteristics by each step: first adding state and year fixed effects, then adding infant risk factors, and finally adding maternal and state characteristics. The most saturated model takes the form:

$$NICU_{ist} = \beta_0 + \beta_1 Private_{ist} + \beta_2 Selfpay_{ist} + \beta_3 Other_{ist} \\ + X_{ist}\Gamma + Z_{st}\rho + \mathbf{u}_s + \boldsymbol{\gamma}_t + e_{ist}$$

where X_{ist} was a vector of infant and maternal covariates, Z_{st} was a vector of state characteristics, \mathbf{u}_s was a vector of state fixed effects, $\boldsymbol{\gamma}_t$ was a vector of year fixed effects, and e_{ist} is a random error term. I then performed subgroup analyses by birthweight. All standard errors are clustered at the state level to allow for unobserved determinants of NICU admission to be correlated within states but independent across states.

There might be some state-level factors that were correlated with maternal insurance type, such as the affordability and accessibility of the private insurance market, the complexity of enrollment process of Medicaid, or the take-up rate of Medicaid patients by physicians, etc. To reduce potential omitted variable bias, I added state fixed effects into models to control for both observed and unobserved the time-invariant state characteristics. Year fixed effects were also included to control for temporal variation over time that is constant across all states. Analyses were conducted using Stata version 15 (StataCorp). Data analysis was approved by the University of Maryland Institutional Review Board (IRB).

4.4 Results

4.4.1 Descriptive Statistics

Table 4.1 displays the descriptive statistics of the study population, demonstrating infant, maternal, and state characteristics by insurance types where 42.66% of births in the years 2016-2017 were covered by Medicaid, 49.29% were covered by private insurance, 4.18% were uninsured, and 3.87% were paid by other sources. I focus on the Medicaid and privately insured population since they accounted for more than 90% of all births in the US. The distributions of maternal characteristics differ greatly between Medicaid and privately insured population. I observe a much higher proportion of non-Hispanic Black and Hispanic, younger, less educated, with higher parity, and a higher prevalence of infections during pregnancy among Medicaid mothers. The infants covered by Medicaid were more likely to be low birthweight, have a shorter gestational age, and have a lower Apgar score, compared to privately insured infants. The Medicaid infants are also more likely to be small for gestational age, which may partly attribute to the racial differences in the prevalence of SGA, which is higher for African-Americans infants compare to Whites (Alexander et al. 1999). The state characteristics are very similar across insurance types.

Table 4.1: Descriptive Statistics by Insurance Type

| Characteristic, % | Overall | Medicaid | Private Insurance | Self-pay | Others |
|---------------------------------|---------|----------|-------------------|----------|--------|
| Mother's Characteristics | | | | | |
| Race | | | | | |
| Non-Hispanic White | 52.4 | 37.25 | 67.39 | 35.76 | 46.45 |
| Non-Hispanic Black | 14.36 | 22.12 | 8.08 | 10.74 | 12.71 |
| Non-Hispanic Other | 9.64 | 7.74 | 10.97 | 12.05 | 11.21 |
| Hispanic | 23.59 | 32.89 | 13.56 | 41.45 | 29.63 |
| Age | | | | | |
| <20 | 5.24 | 9.45 | 1.64 | 4.63 | 5.14 |

| | | | | | |
|--|-------|-------|-------|-------|-------|
| 20-24 | 20.16 | 30.03 | 11.46 | 17.72 | 24.85 |
| 25-34 | 57.44 | 49.39 | 64.49 | 57.58 | 56.21 |
| 35-44 | 16.94 | 11 | 22.09 | 19.78 | 13.63 |
| 45+ | 0.22 | 0.12 | 0.31 | 0.29 | 0.17 |
| Education | | | | | |
| Less than High School | 13.5 | 23.96 | 2.78 | 31.59 | 15.2 |
| High School | 25.29 | 38.71 | 13.94 | 22.05 | 25.47 |
| Some College | 28.92 | 29.77 | 28.64 | 18.27 | 34.55 |
| University and Above | 32.29 | 7.56 | 54.65 | 28.08 | 24.78 |
| Parity | | | | | |
| First | 38.09 | 33.02 | 42.93 | 32.28 | 38.55 |
| Second | 32.21 | 29.52 | 34.9 | 28.99 | 31 |
| Third or Higher | 29.71 | 37.46 | 22.17 | 38.73 | 30.46 |
| Indicator of Any Maternal Morbidity | 1.19 | 1.01 | 1.33 | 1.08 | 1.5 |
| Indicator of Any Risk Factors during Pregnancy | 29.68 | 30.01 | 30.08 | 23.31 | 27.86 |
| Indicator of Any Infections during Pregnancy | 2.67 | 4.59 | 1.05 | 2.28 | 2.44 |
| Infant's Characteristics | | | | | |
| Birthweight | | | | | |
| 500-1499g | 1.21 | 1.43 | 1.04 | 1.02 | 1.2 |
| 1500-2499g | 6.76 | 7.91 | 5.89 | 5.58 | 6.36 |
| >=2500g | 92.03 | 90.66 | 93.07 | 93.4 | 92.44 |
| Gestational Age | | | | | |
| <32 weeks | 1.38 | 1.61 | 1.19 | 1.21 | 1.39 |
| 32-36 weeks | 8.23 | 9 | 7.69 | 6.78 | 8.09 |
| >=37 weeks | 90.4 | 89.39 | 91.13 | 92.01 | 90.52 |
| 5 Min Apgar Score | | | | | |
| 0-3 | 0.45 | 0.53 | 0.38 | 0.46 | 0.5 |
| 4-6 | 1.38 | 1.51 | 1.27 | 1.28 | 1.53 |
| 7-8 | 12.08 | 12.37 | 11.78 | 11.21 | 13.55 |
| 9-10 | 86.09 | 85.59 | 86.57 | 87.06 | 84.42 |
| Plurality | | | | | |
| Single | 96.62 | 97.1 | 96.12 | 97.47 | 96.88 |
| Twin | 3.28 | 2.83 | 3.75 | 2.44 | 3.03 |
| Triplet or More | 0.1 | 0.06 | 0.13 | 0.09 | 0.09 |
| Male | 51.15 | 51.05 | 51.18 | 51.55 | 51.33 |
| Indicator of Any Congenital Anomaly | 0.34 | 0.36 | 0.33 | 0.35 | 0.32 |
| Indicator of Small for Gestational Age | 8.66 | 10.32 | 7.26 | 8.81 | 8.16 |
| State Characteristics | | | | | |

| | | | | | |
|--|-----------|-----------|-----------|---------|--------|
| Poverty Rate | 11.28 | 11.59 | 11.00 | 11.41 | 11.30 |
| Unemployment Rate | 5.23 | 5.30 | 5.17 | 5.18 | 5.22 |
| Birth Rate | 12.13 | 12.15 | 12.07 | 12.35 | 12.45 |
| NICU Bed, per 1,000 pop | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 |
| Prescription of Oxycodone, kilograms per 100,000 pop | 16.47 | 16.48 | 16.54 | 16.05 | 15.96 |
| Prescription of Hydrocodone, kilograms per 100,000 pop | 9.41 | 9.79 | 9.09 | 9.34 | 9.42 |
| | | | | | 292,27 |
| N | 7,547,320 | 3,219,485 | 3,719,746 | 315,814 | 5 |

Note. Data from the Vital Statistics 2016-2017. The sample includes 50 states and DC, infants with birthweight higher than 500g, and with no missing values in all covariates.

Table 4.2 presents the descriptive summary of outcome variables by insurance type, and for each birthweight subgroup. Among all births, I observe an 8.7% NICU admission rate. The utilization of all these services is much higher among infants with very-low birthweight, lower for moderately-low infants, and much lower among normal birth infants. Comparing across insurance types, the overall NICU admission rate was highest among Medicaid infants. However, for all very-low birthweight infants, the NICU was used the most among privately insured infants. This difference also applied to moderately-low birthweight infants. For normal birthweight infants, the NICU admission rate was the highest among Medicaid infants.

Table 4.2: NICU Admission Rates by Insurance Type and by Birthweight

| NICU Admission, % | Overall | Birthweigh | Birthweight | Birthweigh |
|--------------------------|---------|------------|-------------|------------|
| | | t | 1,500- | t |
| | | 500-1,499g | 2,499g | >=2,500g |
| Overall | 8.7 | 87.85 | 43.49 | 5.11 |
| <i>By Insurance Type</i> | | | | |
| Medicaid | 9.51 | 87.48 | 41.58 | 5.49 |
| Private Insurance | 8.19 | 88.72 | 45.9 | 4.9 |
| Self-pay | 6.91 | 84.08 | 41.87 | 3.98 |
| Other | 8.25 | 86.4 | 42.75 | 4.86 |

Note. Data from the Vital Statistics 2016-2017. The sample includes 50 states and DC, infants with birthweight higher than 500g, and with no missing values in all covariates.

4.4.2 Variation in NICU Admissions across Insurance Type

Table 4.3 presents the association of insurance type with the likelihood of admission into a NICU. In the unadjusted model, private NICU admissions were 1.32 percentage points lower than Medicaid. After adjusting for infant characteristics only, the association decreased to 0.34 percentage points difference. In the final model that adjusts for infant, maternal, state characteristics, and fixed effects, I observe a 0.61 percentage point's lower NICU admission rate among privately insured infants.

Table 4.3: The Association between Insurance Type and NICU Admissions

| | Unadjusted (1) | Adjusted (2) | Adjusted (3) | Adjusted (4) |
|--------------------------|--------------------|--------------------|--------------------|--------------------|
| Medicaid | Ref. | Ref. | Ref. | Ref. |
| Private Insurance | -1.32*** (0.16) | -1.39*** (0.16) | -0.34*** (0.09) | -0.61*** (0.07) |
| Infant Characteristics | N | N | Y | Y |
| Maternal Characteristics | N | N | N | Y |
| State Characteristics | N | N | N | Y |
| State Fixed Effects | N | Y | Y | Y |
| Year Fixed Effects | N | Y | Y | Y |

Note. N=7,547,320. Percentage points are reported. Standard errors are in the parentheses, clustered at the state level. Infant characteristics include birthweight, gestational age, sex, an indicator of congenital anomalies, Apgar score, an indicator of small for gestational age, and plurality; maternal characteristics included race, age, education, parity, and indicators of maternal pregnancy risks, infections, and morbidity; state characteristics included level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and prescription rates of Oxycodone and Hydrocodone. N denotes the number of observations. *** p<0.001, ** p<0.01, * p<0.05

Table 4.4 examined the association of insurance type and NICU admission by birthweight categories. It shows that among very-low birthweight infants that need NICU care, the unadjusted admission rate is 1.25 percentage points higher for infants covered by private insurance. The difference disappears both in magnitude and significance after controlling for infant characteristics. Including maternal and state characteristics do not change the results much. A similar pattern has been seen among moderately-low birthweight infants where the unadjusted NICU admission is 4.32 percentage points higher among private insured infants, and the adjusted admission becomes small in magnitude and much less significant. In contrast, the variation in NICU admissions

among normal birthweight infants persists after controlling for all the covariates and fixed effects.

Table 4.4: The Association between Insurance Type and NICU Admissions, by Birthweight

| | Unadjusted (1) | Adjusted (2) | Adjusted (3) | Adjusted (4) |
|---------------------------|--------------------|--------------------|--------------------|--------------------|
| Birthweight 500g-1,499g | | | | |
| Medicaid | | | Ref. | |
| Private Insurance | 1.25*** (0.35) | 1.21*** (0.34) | 0.6 (0.34) | -0.38 (0.41) |
| N | 91,506 | | | |
| Birthweight 1,500g-2,499g | | | | |
| | (1) | (2) | (3) | (4) |
| Medicaid | | | Ref. | |
| Private Insurance | 4.32*** (0.59) | 3.59*** (0.47) | 0.74* (0.28) | -0.48* (0.22) |
| N | 509,971 | | | |
| Birthweight >=2,500g | | | | |
| Medicaid | | | Ref. | |
| Private Insurance | -0.59*** (0.11) | -0.67*** (0.09) | -0.47*** (0.08) | -0.62*** (0.07) |
| N | 6,945,843 | | | |
| Infant Characteristics | N | N | Y | Y |
| Maternal Characteristics | N | N | N | Y |
| State Characteristics | N | N | N | Y |
| State Fixed Effects | N | Y | Y | Y |
| Year Fixed Effects | N | Y | Y | Y |

Note. Percentage points are reported. Standard errors are in the parentheses, clustered at the state level. Infant characteristics include birthweight, gestational age, sex, an indicator of congenital anomalies, Apgar score, an indicator of small for gestational age, and plurality; maternal characteristics included race, age, education, parity, and indicators of maternal pregnancy risks, infections, and morbidity; state characteristics included level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and prescription rates of Oxycodone and Hydrocodone. N denotes the number of observations.*** p<0.001, ** p<0.01, * p<0.05

4.5 Discussion

Most births are fully or partially covered by health insurance in the US. The two main sources of health insurance covering births are Medicaid and private insurance. As of 2011, 44.7 percent of all births were covered by Medicaid, and 48.7 percent were covered by private insurance, and the remaining 3.6 percent are uninsured, and 3 percent have other types of insurance coverage (Kowlessar, Jiang, and Steiner 2013). I find that Medicaid covered 42.66% of births and privately insurance covered 49.29% as of 2016-2017. Newborns' characteristics also differ between Medicaid and private payers such that newborns covered by Medicaid had higher incidence rates for preterm (8.9 percent vs. 8.1 percent) and low birth weight (6.8 percent and 5.5 percent) (Kowlessar, Jiang, and Steiner 2013). In this study, I also find the differences in maternal characteristics between Medicaid and privately insured births. For example, I observe a much higher proportion of non-Hispanic Black and Hispanic, younger, less educated, with higher parity, and a higher prevalence of infections during pregnancy among Medicaid mothers,

Health insurance is a crucial source of financial incentives where reimbursement rates for physicians and hospitals differed across payers. Higher reimbursement may lead to more utilization. There is plenty of literature discussing the relationship between health insurance and healthcare service utilization for different types of medical services (e.g., Anderson, Dobkin, & Gross, 2012; Finkelstein et al., 2012; Jackson, 2018; Meer & Rosen, 2004). Unfortunately, neonatal specific research was very limited. Given that private insurance usually reimburses medical services more generously compared with public insurance, Braveman et al. (1991) found that sick newborns that are either without health insurance or are covered by Medicaid receive less inpatient care and shorter

hospital stays compared with privately insured, using data from California (Braveman et al. 1991). Moreover, newborns that were covered by Medicaid received less “low-tech” neonatal care compared with privately covered newborns using US birth certificate data (1987-1992) (Currie and Gruber 2001). A recent descriptive study found no relationship between the number of special care days between commercial and Medicaid insured among very low birthweight or low-risk infants (David C. Goodman, George A. Little, Wade N. Harrison, Atle Moen, Meredith E. Mowitz, Cecilia Ganduglia-Cazaban, Kristen K. Bronner 2019).

This study adds suggestive evidence to previous findings on comparing “high-tech” NICU service utilization between Medicaid and privately insured. With stratification analysis by birthweight, I also show that the difference across payers can be explained by infant risk factors for very-low birthweight infants, but cannot be explained for lower-risk infants.

Within the population of normal weight infants, private coverage is associated with a reduction in NICU admission. While it’s possible that the higher reimbursement rates of private insurance did incentivize more NICU admissions among privately insured infants that were relatively healthy, my results suggest that the amount of these effects are likely to be quite small.

4.6 Conclusion

This study explores the variation in NICU utilization that is related to a critical source of financial incentive – insurance coverage. In spite of an existing significant variation in NICU admissions between Medicaid and privately insured infants, such

variation does not persist after adjusting for infant risk among very-low birthweight infants who need intensive service the most. The variations persist for normal birthweight infants that were relatively healthy. Although these findings are descriptive, they suggest that a great amount of variation is not attributed to demand-side factors given that the infant risk factors and maternal characteristics are mostly controlled for.

Chapter 5 : Conclusion

This dissertation provides the latest evidence of Maryland's Global Budget Revenue program on hospital-based neonatal care and the knowledge of the economic incentives related to the NICU service. This work contributes to the literature on three aspects.

First, lessons from the Maryland model are critical to national efforts at reforming the health care delivery system. This dissertation expands the evaluation of Maryland's model to a high value but expensive hospital service. I provide evidence on the direct or spillover effects of GBR on NICU services which is lacking in the previous evaluation.

Secondly, previous studies investigating economic incentives on NICU utilizations mainly focus on the relationship between capacity and admission. In this dissertation, I expand the literature by exploring the incentives generated from Maryland's payment reform. In addition, I provide evidence on more outcomes (i.e. admission into lower level neonatal nursery, length of stay, and cost) which are not captured by NICU admission rate in previous studies.

This work also contributes to the literature as it is the first time that NICU utilization is compared by insurance type from a population perspective. The findings help us to better understand the role that health insurance played in the NICU utilization and to extend the knowledge of variations in NICU service utilization.

Appendix A

A1. Coding of Comparison States

The primary comparison states used in the main analysis consist of states that had adopted the 2003 revision of the birth certificate by 2010 and adopted the ACA Medicaid expansion. The implementation of the 2003 revision of the U.S. Standard Certificate of Live Birth started in 2004. As of January 1, 2016, all states and the District of Columbia had implemented this revised certificate. Maryland implemented the revised birth certificate on January 1, 2010. I obtained the list of states that had implemented the 2003 revised birth certificate as of January 1, 2010 to ensure a comparable time series across states. Thirty-three states and District of Columbia met this criterion (CDC 2010). I didn't include 2010 in my study period as we considered it as a transition period in which hospitals were adapting to the new form.

I next obtained the ACA Medicaid expansion status from Boudreaux et al. (2019) which also provided the timing of the expansion (M. Boudreaux et al. 2019). The following table (Table A.1) summarized the coding of Medicaid expansion status. I included all states that expanded Medicaid regardless of the timing of expansion in my main analyses. The states in bold were used in the preferred specification of comparison states.

A2. Variable Definitions

The individual-level covariates included in the preferred specification were: birthweight category (500-1499g, 1500-2499g, 2500-3999g, 4000g+), gestational age category (<32 week (very preterm), 32-36 week (moderately preterm), 37-41 (term), >41week (postterm)), maternal race and ethnicity (non-Hispanic White, non-Hispanic Black, non-Hispanic others, Hispanic), maternal age (<20, 20-24, 25-34, 35-44, 45+), maternal education (less than high school, high school, some college, university and above), insurance type (Medicaid, private insurance, self-pay, other), parity (0, 1, 2, 3+), an indicator of any maternal morbidity (maternal transfusion, third or fourth degree perineal laceration, ruptured uterus, unplanned hysterectomy, admission to intensive care unit, unplanned operating room procedure), an indicator of any risk factors during pregnancy (prepregnancy diabetes, gestational diabetes, prepregnancy hypertension, gestational hypertension, eclampsia, previous preterm birth, other previous poor pregnancy outcomes, pregnancy resulted from infertility treatment, mother had a previous cesarean delivery), an indicator of any infections during pregnancy (gonorrhea, syphilis, chlamydia, hepatitis B, hepatitis C), infant's gender male (=1 if male, =0 if female), and an indicator of any congenital anomalies (anencephaly, meningomyelocele/spina bifida, cyanotic congenital heart disease, congenital diaphragmatic hernia, omphalocele, gastroschisis, limb reduction defect, cleft lip with or without cleft palate, cleft palate alone, down syndrome, suspected chromosomal disorder, hypospadias). These covariates were selected based on previous literature and the quality of variables (W. Harrison and Goodman 2015; W. N. Harrison, Wasserman, and Goodman 2018). For example, maternal marital status was missing in California in 2017 and smoking status during

pregnancy was missing in California, Georgia, and Michigan for several years. I excluded these two variables in my analysis.

In the aggregate-level infant mortality analysis, I controlled for the percent of moderately-low birthweight infants, preterm infants, and infants with congenital infections; percent of mothers that are non-Hispanic White, aged less than 35 years old, with less than high school education, with first birth order, have maternal morbidity, had risk factors during pregnancy, had maternal infections during pregnancy, and state-level characteristics as above.

A3. Pre-trends Test

The assumption of difference-in-differences, commonly referred to as the parallel trends assumption, is that the treatment group and the comparison group would have followed the same trend had the intervention never occurred. While the parallel trends assumption cannot be tested directly, following common practice I gauged how plausible the assumption was in the data by ascertaining whether the outcome trends in Maryland were parallel with trends in the comparison states in the years leading up to the start of the policy.

The outcome graphs shown in the main paper (Figure 2.1) suggested reasonably parallel pre-trends in the NICU outcome. Figure A.1 and Figure A.2 demonstrate the same thing for infant and neonatal mortality rates among all births, among moderately-low birthweight (MLBW) and normal birthweight (NBW) births, and among moderately preterm and term births; and stratified for overall and for infants who were not admitted to a NICU.

To obtain more formal statistical evidence I compared linear trends in Maryland and the comparison states using regressions based only on pre-period data (2011-2014). The first set of regressions, based on individual-level data, used the NICU admission indicator as the outcome. The predictors included the interaction of continuous year and the Maryland indicator and the full set of covariates and fixed effects. I estimated models in the full sample and in each birthweight category. Table A.2 reports the coefficients from the interaction term, its 95% confidence intervals, and the p-value (from the clustered robust method on the state level). Statistical inference was based on clustered robust standard errors which are known to over-reject the null hypotheses. Unfortunately, the bootstrap routine I used in the main analysis requires discrete time period dummies and I was unable to implement it in the pre-trends test because the time variable of interest was continuous. Furthermore, given the large sample, even small differences can be statistically significant. Thus, the statistical tests I employed in Table A.2 represent a conservative approach to identifying differential pre-trends. I found that the coefficients were quite small (compared to the baseline mean in Maryland) for all groups considered. The p-values were large or just crossed 0.05 for all groups, except for the very-low birthweight group. However, given that the point estimates were relatively small for all groups (relative to their baseline means) and that my approach to statistical inference in Table A.2 is known to over-reject the null, I interpreted Table A.2 as evidence in favor of the parallel trends assumption.

I conducted a similar test of infant and neonatal mortality rates. For all groups, I found small coefficients and large p-values, suggesting there were no pre-existing trends of infant and neonatal mortality rates in Maryland and the comparison states (Table A.2).

A4. Effects by Year of Implementation

Besides the average post-period effects shown in Table 2.2, I also estimated effects by the year of implementation. This alternative model allowed effects to flexibly evolve over each year of GBR. The analysis was conducted by changing the post*treatment interaction term with three interaction terms for each post-period year and the Maryland indicator. All other features of the model were the same.

Table A.3 reports the coefficients from these interaction terms for all infants, MLBW, and NBW infants. I found that the effects of GBR grew over the study period for all three groups. For example, NICU admission rates for NBW infants decreased by 13.2% in the first full year of the implementation and 35.6% in the third full year of the implementation.

A5. Robustness to Alternative Comparison States

I replicated the main model using alternative comparison states to test whether the results were sensitive to the choice of the control group. I first used all 33 states and DC that had NICU information regardless of their Medicaid expansion status. To further exclude the effect of the timing of Medicaid expansion, I next restricted the sample to the 2014 expansion states.

In Table A.4, I show that both the magnitude and significance of my results were unchanged with both sets of comparison states.

A6. Robustness to Alternative Implementation Timing

I chose to consider 2015 as the first year of implementation because hospitals became subject to the global budgets in July 2014 (budgets were set on a fiscal year base)

and I expected that there would be a transition period as hospitals made adjustments to their strategies under the global budget environment. This decision differed from previous studies of the GBR program which used 2014 as the implementation year. Below I report how sensitive my results are to alternative implementation dates. I either used January 2014 as other studies or July 2014.

In Table A.5, I report difference-in-differences estimates using these two alternative implementation times. I observed slightly smaller coefficients using 2014 and July 2014 as the implementation time compared to 2015, but I came to the same basic conclusions. Finding smaller effects in these models, compared to my preferred specification, is consistent with the idea that the policy did not start to lead to changes in outcomes until 2015.

A7. Robustness to Alternative Model Specifications

I conducted several tests to examine if my results were sensitive to alternative model specifications. First, I examined how sensitive my results were to the set of included covariates by estimating unadjusted models that included only the difference-in-differences interaction and the state and year fixed effects. I also examined if my main model results changed after including Apgar scores and an indicator of congenital anomalies. A linear probability model was used in our main analysis of NICU admissions for the ease in interpreting the coefficients and because it allowed us to generate p-values from a bootstrap method that properly accounted for auto-serial correlation in the presence of single treated cluster. However, I also investigated if using logistic regressions with state clustered standard errors suggested the same pattern of results as my preferred approach.

In Table A.6, I report the difference-in-differences coefficients from an unadjusted linear probability model (column 2) and the incremental effects implied by the logistic regression model based on my main set of covariates (column 3). I find consistent results using these alternative model specifications.

A8. Robustness to Alternative Samples

In this section, I examined whether my results are sensitive to certain changes of my study sample. I first tested whether the NICU admission results were robust to excluding Baltimore City and Baltimore County, which had an initiative (B'more for Healthy Babies) that led to improved infant health starting in 2009 (Baltimore City Health Department 2009). The coefficients became larger when removing Baltimore City and Baltimore County from our sample (Table A.7, column 1). The change in NICU admissions for all births was -1.77 percentage points (95% CI, -2.26 to -1.29), -7.45 percentage points (95% CI, -9.07 to -5.82) for VLBW infants, -7.75 percentage points (95% CI, -8.98 to -6.51) for MLBW infants, -1.41 percentage points (95% CI, -1.88 to -0.94) for NBW infants, and -1.36 percentage points (95% CI, -1.81 to -0.90) for HBW infants. However, my general conclusions were unchanged from the preferred model.

I also re-estimated the model by restricting the sample to state residents, as five hospitals (University of Maryland Medical Center, Johns Hopkins Hospital, Johns Hopkins Bayview, and Johns Hopkins Suburban, and University of Maryland Shock Trauma) were exempt for nonresident services from their budgets in 2014 (Giuriceo et al. 2016). The University of Maryland facilities' budgets stopped excluding nonresident revenues starting in FY 2015. The results are shown in Table A.7 (column 2). Even though the exempt hospitals have an incentive to increase the volume of services

provided to out-of-state patients as a way to increase their revenues, the population of non-residents was smaller and excluding them didn't change much to my main results.

The preferred sample included all births, including births at rural hospitals that had previously been under the pilot program. However, none of these hospitals have a NICU facility. I also included births delivered at birthing centers. While I believe that infants who were delivered in rural hospitals or birth centers should also be affected by GBR via transfers, I re-ran the model by restricting the sample to urban and hospital births only. The coefficients are quite similar such that our results were not sensitive to the inclusion or exclusion of those facilities (Table A.7, column 3).

A9. The Impact of GBR on Infant Health

I interpreted the results as suggesting that the GBR program reduced NICU admissions, holding health at birth constant. However, it is also possible that GBR could have reduced NICU services by improving health at birth. Such effects may have been controlled out of the preferred model by including infant and maternal health characteristics as covariates.

Nonetheless, to investigate if GBR had impacts on health at birth, I used our difference-in-differences model to examine indicators of infant health. The indicators I used were preterm (less than 37 weeks of gestational age), small for gestational age (SGA), average birthweight, and average gestational age. An infant was defined as SGA if her/his birthweight was lower than the tenth percentile of birthweight at a given gestational age (in weeks) (Fenton and Kim 2013). The SGA was calculated based on 2013 Fenton growth charts by infant gender (Fenton and Kim 2013).

Figure A.3 shows the trends for these four outcomes in Maryland and the comparison states. Compared to other states, Maryland had a higher rate of preterm births and the trends started to diverge in Maryland versus the comparison states prior to GBR. As such, results for the preterm indicator should be interpreted with caution. The trends in all outcomes appeared quite similar between Maryland and comparison states before GBR.

In Table A.8, I show the difference-in-differences estimates using these four measures as outcome variables. I did not find evidence that the GBR was associated with any health outcome considered. These findings provide additional evidence that GBR effects on NICU use arise from changes to practice patterns rather than changes to health at birth.

A10. Measurement Error in NICU Admissions

An important concern with our results is whether differential misclassification of NICU admission biased our findings. To better understand if that issue influenced our results, I compared NICU admission rates from hospital discharge data in Maryland, as obtained from the State Inpatient Database of the Healthcare Cost and Utilization Project (HCUP), to NICU rates observed in Vital Statistics.

In the Vital Statistics, NICU admission is defined as “admission into a facility or unit staffed and equipped to provide continuous mechanical ventilator support for the newborn”(National Center for Health Statistics 2012). The guidelines for completing the Birth Certificate through Facility Worksheets define the NICU admission as “include NICU admission at any time during the infant’s hospital stay following delivery. Do not

include units that do not provide continuous mechanical ventilation. Do not include well-baby nurseries or special care nurseries (i.e., Level II nursery). Do not include if the newborn was taken to the NICU for observation but is not admitted to the NICU". The definition from the Vital Statistics, therefore, attached the NICU admission to a facility or unit.

In the HCUP, there's no NICU admission indicator. A common approach to classify NICU admission using discharge data is following AAP's definition which defines a NICU as either a level III or level IV nursery and measuring the NICU admission using revenue codes that denoted level III or level IV nursery care (revenue code = 0173 or 0174) (Goodman et al. 2019). However, the assigned revenue code corresponds to the level of care determined during the clinical evaluation rather than the level of facility (Maryland Department of Health 2017; New York State Department of Health 2012). As stated on New York State's Statewide Planning and Research Cooperative System, "The levels of care and resulting revenue codes may, and likely will, fluctuate during the infants stay in the facility." (New York State Department of Health 2012)

Because the higher-level facilities includes the capabilities of previous levels (Kilpatrick, Papile, and Macones 2017), I expected the NICU admission rate from the Vital Statistics (VS) data to be higher compared to the HCUP data. I graphed the trends of the NICU admission rate for all births and by birthweight groups using each data set. The NICU admission in the HCUP data was measured using revenue codes that denoted level III or level IV nursery care (revenue code = 0173 or 0174).

Figure A.4 depicts NICU admission rates. While the levels were expectedly different, the trends of the NICU admission rate were generally similar among all groups except for the very-low birthweight group. The Vital Statistics suggested an upward trend in NICU admissions during the pre-period that was absent in the HCUP. The different trends in this group could reflect transfers that are captured with varying levels of accuracy in the Vital Statistics, but are expected to be more accurately measured in the HCUP. However, because the two data sources use different definitions of NICU services, I cannot confidently determine if differences in trend reflect measurement error or real changes in service use. While the differing trends in the VLBW group are somewhat concerning, the HCUP suggests very little change in its measure of NICU utilization among VBLW infants. That is consistent with our conclusions based on the analysis of the VS data featured in the main paper. Further, the HCUP does suggest declines in NICU admission for MLBW and NBW infants consistent with the implementation of GBR and the findings I present in the main paper.

A11. Synthetic Control Methods Estimates

Besides the preferred difference-in-differences design, I also explored an alternative approach, the synthetic control method (SCM), to estimate the effects of GBR. It's an increasingly popular method for policy evaluation and uses a data-driven approach for selecting a comparison group that is a weighted average of all states (Abadie, Diamond, and Hainmueller 2015). Non-negative weights are chosen to minimize the difference in outcomes during the pre-period, which ensures that pre-period trends are as similar as possible for the treatment unit and synthetic unit. I consider the same group of states (20 states including DC) as the donor pool to build the synthetic Maryland to deal

with the concerns from Medicaid expansion and facilitate comparison between models of DID and SCM. Treatment effects are then measured as the mean differences between Maryland and the synthetic Maryland in the post-period. The statistical inference for SCM is based on a permutation-based test where it assigned the treatment to each state in the donor pool and re-estimates the model, resulting in a series of placebo treatment effects. A p-value is then calculated as the probability that the original treatment effects surpass all other placebo ones by ranking the ratios of post/pre-intervention mean squared prediction error (MSPE). I used all the pre-treatment outcomes as my predictors following previous literature (Bilgel and Galle 2015; Kreif et al. 2016). There might be some concerns regarding including pre-treatment outcomes only as the predictors (Kaul et al. 2018) and there hasn't come to a consensus on the inclusion criteria of covariates (Botosaru and Ferman 2019), while I find that the results using the pre-treatment outcomes in SCM are quite similar to those using the DID, plus balanced covariates through conducting a balance test, suggesting the potential bias could be trivial.

Figure A.5 depicts Maryland and synthetic Maryland for the overall group. The pre-treatment trends between Maryland and synthetic Maryland are matched quite well before 2015. After 2015, the NICU admission rate decreased substantially in Maryland while it keeps increasing in synthetic Maryland. The average post-period NICU admission decreased 0.78 percentage points, corresponding to a 10.4% decrease from the baseline rate in Maryland. Table A.9 shows the list of states that contribute to the comparison group. Twenty states contributed in roughly equal proportions.

Table A.10 presents the results using the synthetic control method for overall and each birthweight group. The decrease is largest and significant in the MLBW group and

also significant in the overall group. Changes were small in other groups or insignificant. As mentioned above to calculate p-values, Figure A.6 depicts the ratio of post/pre-intervention MSPE ratios for Maryland and 20 comparison states where Maryland ranks the first. The exact p-value, i.e. the probability of obtaining a post/pre-intervention MSPE ratio as large as Maryland's for overall group, is $1/21=0.48$.

Table A.11 compares the size of effects from the DID method and the synthetic control method. The direction of the effects is consistent using these two methods, and the decreases are consistently significant for overall and for the MLBW group.

A12. Dynamic Difference-in-Differences Estimates

In addition to the average treatment effects, I explored the dynamic difference-in-differences results by replacing the interaction term of treatment dummy and the pre-post dummy with a series of interactions of the year and the treatment status. The estimated coefficients were standardized where the year before 2015 was taken as the baseline level with a coefficient as zero. Figure A.7 presents the results from the dynamic models for all births and by birthweight, where the dashed orange lines indicated the 95% confidence interval from standard errors clustered at the state level.

A13. Models with State-specific Linear Trends

There are some mild violations of pre-trend tests as shown in Table A.2. Following Wolfers (2006), I add the state-specific linear trend into the models and replace the interaction term with separate interactions of treatment status and post-period dummies for the years 2015, 2016, and 2017 (Wolfers 2006). This is to prevent the added state-specific trends from absorbing treatment effects.

Table A.13 presents the results while adding state-specific linear trends. Cluster robust standard errors at the state level are reported. I now find a larger impact among VLBW infants which reaches 11.4% decrease in the year 2017. I find declines among MLBW infants that are substantial and consistent right after the implementation of GBR. The changes among NBW and HBW infants are trivial in the first two years while becoming significant and substantial heading into the third year. Overall, the findings are consistent with those from my preferred difference-in-differences setting.

Table A.1: Comparison States

| State | Expansion Date | State | Expansion Date |
|-----------------------------|-----------------------|----------------------|-----------------------|
| California* | 2014 | Nevada | 2014 |
| Colorado | 2014 | New Hampshire | 2015 |
| Delaware | Early | New Mexico | 2014 |
| District of Columbia | Early | New York | Early |
| Florida | Never | North Dakota | 2014 |
| Georgia | Never | Ohio | 2014 |
| Idaho | Never | Oklahoma | Never |
| Illinois | 2014 | Oregon | Early |
| Indiana | 2015 | Pennsylvania | 2015 |
| Iowa | 2014 | South Carolina | Never |
| Kansas | Never | South Dakota | Never |
| Kentucky | 2014 | Tennessee | Never |
| Maryland | 2014 | Texas | Never |
| Michigan | 2014 | Utah | Never |
| Missouri | Never | Vermont | Early |
| Montana | 2016 | Washington | 2014 |
| Nebraska | Never | Wyoming | Never |

Note. The listed states are those that implemented the 2003 version of the U.S. Birth Certificate early or at the same time as Maryland. The states in bold are used in the preferred specification of comparison states.

Table A.2: Differential Pre-Trends Test, 2011-2014

| | Coef. | 95% CIs | Cluster Robust P-value |
|---------------------------------------|---------|---------------|------------------------|
| <i>NICU Admission</i> | | | |
| All | -0.12 | (-0.31,0.08) | 0.23 |
| Very-low Birthweight | 1.48 | (0.26,2.70) | 0.02 |
| Moderately-low Birthweight | 0.21 | (-0.61,1.02) | 0.6 |
| Normal Birthweight | -0.17 | (-0.33,-0.01) | 0.04 |
| High Birthweight | -0.23 | (-0.46,0.01) | 0.06 |
| Very Preterm | 1.17 | (0.06,2.29) | 0.04 |
| Moderately Preterm | -0.07 | (-0.77,0.63) | 0.84 |
| Term | -0.16 | (-0.32,0.01) | 0.06 |
| Postterm | -0.23 | (-0.67,0.21) | 0.29 |
| <i>Infant Mortality Rate</i> | | | |
| All | 0.21 | (-0.12,0.55) | 0.2 |
| MLBW and NBW | 0.19 | (-0.01,0.40) | 0.06 |
| MLBW and NBW & Not Admitted to NICU | 0.2 | (-0.03,0.43) | 0.09 |
| MPT and Term | 0.2 | (-0.01,0.41) | 0.06 |
| MPT and Term & Not Admitted to NICU | -0.002 | (-0.20,0.20) | 0.98 |
| <i>Neonatal Mortality Rate</i> | | | |
| All | -0.05 | (-0.33,0.23) | 0.71 |
| MLBW and NBW | -0.0001 | (-0.20,0.20) | 1 |
| MLBW and NBW & Not Admitted to NICU | -0.01 | (-0.20,0.18) | 0.91 |
| MPT and Term | -0.002 | (-0.20,0.20) | 0.98 |
| MPT and Term & Not Admitted to NICU | -0.004 | (-0.19,0.18) | 0.96 |

Note. The coefficients and 95% CIs for NICU admissions outcomes are in percentage point. Models control for the same covariates as the preferred models. Cluster robust p-values are obtained from the clustered sandwich estimator.

*Table A.3: Difference-in-Differences Estimates of the Effect of GBR on NICU Admission
by Years of Implementation, 2011-2017*

| Dependent Variable: NICU Admission Indicator | Average Admission Rate before GBR in Maryland | Adjusted Difference-in-Differences Estimate (95% CIs) | Cluster Robust P-Value | Relative Effect from Baseline, % |
|--|---|---|------------------------|----------------------------------|
| All (N=11,965,997) | | | | |
| Year 1 | 7.49 | -0.62 (-1.03,-0.22) | 0.005 | -8.28 |
| Year 2 | 7.49 | -1.07 (-1.52,-0.62) | <0.001 | -14.29 |
| Year 3 | 7.49 | -2.13 (-2.81,-1.45) | <0.001 | -28.44 |
| MLBW (N=591,120) | | | | |
| Year 1 | 37.37 | -1.88 (-3.32,-0.43) | 0.013 | -5.03 |
| Year 2 | 37.37 | -3.5 (-4.70,-2.31) | <0.001 | -9.37 |
| Year 3 | 37.37 | -8.17 (-9.53,-6.81) | <0.001 | -21.86 |
| NBW (N=10,238,271) | | | | |
| Year 1 | 4.77 | -0.63 (-1.00,-0.26) | 0.002 | -13.21 |
| Year 2 | 4.77 | -1 (-1.44,-0.57) | <0.001 | -20.96 |
| Year 3 | 4.77 | -1.7 (-2.36,-1.04) | <0.001 | -35.64 |

Note. Models control for birthweight (in full and gestation analyses), gestational age (in full and birthweight analyses), mother's age, race, education level, insurance type, parity, maternal morbidities, infections, and risks, infant's sex and congenital anomalies, state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects. The coefficients and 95% CIs are in percentage points. Standard errors are clustered at the state level. P-values are from clustered robust standard errors. The baseline rate refers to the average admission rate before GBR in Maryland which is calculated using data from 2011 to 2014.

Table A.4: Difference-in-Differences Estimates of the Effect of GBR on NICU Admission using Alternative Control States, 2011-2017

| | Main Model | | | All States | | | All 2014 Expansion States | | |
|------------------------------------|------------|---------------|-------------------|------------|---------------|-------------------|---------------------------|---------------|-------------------|
| | (1) | | | (2) | | | (3) | | |
| Dependent Variable: NICU Admission | Coef. | 95% CIs | Bootstrap P-Value | Coef. | 95% CIs | Bootstrap P-Value | Coef. | 95% CIs | Bootstrap P-Value |
| All | -1.26 | (-1.76,-0.76) | 0.03 | -1.26 | (-1.59,-0.93) | 0.03 | -1.06 | (-1.79,-0.32) | 0.03 |
| Very-low Birthweight | -0.95 | (-2.58,0.69) | 0.59 | -0.33 | (-1.43,0.77) | 0.91 | -1.94 | (-4.51,0.63) | 0.15 |
| Moderately-low Birthweight | -4.5 | (-5.71,-3.29) | 0.003 | -4.36 | (-5.25,-3.47) | 0.04 | -4.45 | (-7.02,-1.89) | 0.02 |
| Normal Birthweight | -1.1 | (-1.58,-0.62) | 0.04 | -1.11 | (-1.42,-0.80) | 0.02 | -0.89 | (-1.53,-0.25) | 0.04 |
| High Birthweight | -0.84 | (-1.36,-0.31) | 0.13 | -0.9 | (-1.25,-0.56) | 0.18 | -0.5 | (-1.29,0.29) | 0.31 |

Note. The baseline rate/mean for Maryland are calculated using data from 2011 to 2014. The coefficients and 95% CIs are in percentage points. Standard errors are clustered at the state level. P-values are obtained using a bootstrap approach, as developed by Ferman and Pinto (2019). The estimations are from separate regressions for all births and for each birthweight cohort. Models control for birthweight (in full and gestation analyses), gestational age (in full and birthweight analyses), mother's age, race, education level, insurance type, parity, maternal morbidities, infections, and risks, infant's sex and congenital anomalies, state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects.

Table A.5: Difference-in-Differences Estimates of the Effect of GBR on NICU Admission using Alternative Post-Implementation Time, 2011-2017

| Dependent Variable: NICU Admission | Main Model | | | Fiscal Year 2014 | | | Calendar Year 2014 | | |
|------------------------------------|------------|---------------|-------------------|------------------|---------------|-------------------|--------------------|---------------|-------------------|
| | (1) | | | (2) | | | (3) | | |
| | Coef. | 95% CIs | Bootstrap P-Value | Coef. | 95% CIs | Bootstrap P-Value | Coef. | 95% CIs | Bootstrap P-Value |
| All | -1.26 | (-1.76,-0.76) | 0.03 | -1.18 | (-1.67,-0.70) | 0.09 | -1.12 | (-1.62,-0.62) | 0.05 |
| Very-low Birthweight | -0.95 | (-2.58,0.69) | 0.59 | -0.25 | (-2.18,1.67) | 0.9 | 0.6 | (-1.77,2.97) | 0.76 |
| Moderately-low Birthweight | -4.5 | (-5.71,-3.29) | 0.003 | -4.2 | (-5.57,-2.82) | 0.06 | -3.8 | (-5.43,-2.17) | 0.06 |
| Normal Birthweight | -1.1 | (-1.58,-0.62) | 0.04 | -1.04 | (-1.50,-0.58) | 0.11 | -1.01 | (-1.47,-0.55) | 0.06 |
| High Birthweight | -0.84 | (-1.36,-0.31) | 0.13 | -0.82 | (-1.33,-0.32) | 0.26 | -0.93 | (-1.46,-0.41) | 0.15 |

Note. The baseline rate/mean for Maryland are calculated using data from 2011 to 2014. The coefficients and 95% CIs are in percentage points. Standard errors are clustered at the state level. P-values are obtained using a bootstrap approach, as developed by Ferman and Pinto (2019). The estimations are from separate regressions for all births and for each birthweight cohort. Models control for birthweight (in full and gestation analyses), gestational age (in full and birthweight analyses), mother's age, race, education level, insurance type, parity, maternal morbidities, infections, and risks, infant's sex and congenital anomalies, state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects.

Table A.6: Difference-in-Differences Estimates of the Effect of GBR on NICU Admission using Alternative Control States, 2011-2017

| Dependent Variable: NICU Admission | Main Model | | | Unadjusted Model | | | Logistic Regression | | |
|---------------------------------------|------------|---------------|-------------------|------------------|---------------|-------------------|---------------------|---------------|------------------------|
| | (1) | | | (2) | | | (3) | | |
| | Coef. | 95% CIs | Bootstrap P-Value | Coef. | 95% CIs | Bootstrap P-Value | Incremental Effects | 95% CIs | Cluster Robust P-Value |
| All | -1.26 | (-1.76,-0.76) | 0.03 | -1.01 | (-1.33,-0.69) | 0.12 | -1.18 | (-1.53,-0.83) | <0.001 |
| Very-low Birthweight | -0.95 | (-2.58,0.69) | 0.59 | -1.39 | (-3.24,0.47) | 0.48 | -1.62 | (-2.94,-0.29) | 0.017 |
| Moderately-low Birthweight | -4.5 | (-5.71,-3.29) | 0.003 | -4.38 | (-5.17,-3.58) | 0.06 | -4.53 | (-5.69,-3.38) | <0.001 |
| Normal Birthweight | -1.1 | (-1.58,-0.62) | 0.04 | -0.89 | (-1.22,-0.57) | 0.11 | -1.03 | (-1.36,-0.70) | <0.001 |
| High Birthweight | -0.84 | (-1.36,-0.31) | 0.13 | -0.56 | (-0.89,-0.24) | 0.33 | -0.76 | (-1.10,-0.42) | <0.001 |

Note. The baseline rate/mean for Maryland are calculated using data from 2011 to 2014. The coefficients and 95% CIs are in percentage points. Standard errors are clustered at the state level. The estimations are from separate regressions for all births and for each birthweight cohort. Models control for birthweight (in full and gestation analyses), gestational age (in full and birthweight analyses), mother's age, race, education level, insurance type, parity, maternal morbidities, infections, and risks, infant's sex and congenital anomalies, state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects. Bootstrap P-values are obtained using a bootstrap approach, as developed by Ferman and Pinto (2019). Cluster robust p-values are obtained from the clustered sandwich estimator.

Table A.7: Difference-in-Differences Estimates of the Effect of GBR on NICU Admission using Alternative Samples, 2011-2017

| Dependent Variable: NICU Admission | Excluding Baltimore City and County | | | Residents Only | | | Hospital & Urban County Births Only | | |
|---------------------------------------|--|---------------|----------------------|----------------|---------------|----------------------|--|---------------|----------------------|
| | (1) | | | (2) | | | (3) | | |
| | Coef. | 95% CIs | Bootstrap P-Value | Coef. | 95% CIs | Bootstrap P-Value | Coef. | 95% CIs | Bootstrap P-Value |
| All | -1.77 | (-2.26,-1.29) | 0.003 | -1.24 | (-1.74,-0.74) | 0.04 | -1.2 | (-1.70,-0.70) | 0.05 |
| Very-low Birthweight | -7.45 | (-9.07,-5.82) | 0.001 | -1.28 | (-2.95,0.40) | 0.49 | -0.67 | (-2.33,1.00) | 0.65 |
| Moderately-low Birthweight | -7.75 | (-8.98,-6.51) | 0.002 | -4.19 | (-5.40,-2.98) | 0.01 | -3.98 | (-5.19,-2.76) | 0.01 |
| Normal Birthweight | -1.41 | (-1.88,-0.94) | 0.02 | -1.11 | (-1.58,-0.63) | 0.04 | -1.08 | (-1.56,-0.60) | 0.05 |
| High Birthweight | -1.36 | (-1.81,-0.90) | 0.07 | -0.77 | (-1.31,-0.23) | 0.18 | -0.79 | (-1.33,-0.24) | 0.19 |

Note. The baseline rate/mean for Maryland are calculated using data from 2011 to 2014. The coefficients and 95% CIs are in percentage points. Standard errors are clustered at the state level. P-values are obtained using a bootstrap approach, as developed by Ferman and Pinto (2019). The estimations are from separate regressions for all births and for each birthweight cohort. Models control for birthweight (in full and gestation analyses), gestational age (in full and birthweight analyses), mother's age, race, education level, insurance type, parity, maternal morbidities, infections, and risks, infant's sex and congenital anomalies, state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects.

Table A.8: Difference-in-Difference Estimates of the Effect of GBR on Infant Health Measures, 2011-2017

| Outcome | Baseline Rate/Mean in Maryland | Coef. | 95% CIs | Bootstrap P-Value |
|--|--------------------------------|-------|---------------|-------------------|
| Preterm, % | 10.11 | -0.08 | (-0.32,0.17) | 0.73 |
| SGA, % | 8.55 | -0.34 | (-0.45,-0.23) | 0.73 |
| Average Birthweight, gram | 3306.67 | 5.67 | (1.60,9.73) | 0.4 |
| Average Gestational Age, week ^b | 38.66 | 0.02 | (-2.58,2.62) | 0.99 |

Note. The baseline rate/mean for Maryland are calculated using data from 2011 to 2014. The coefficients and 95% CIs of preterm and SGA models are in percentage points. Standard errors are clustered at the state level. P-values are obtained using a bootstrap approach, as developed by Ferman and Pinto (2019). Models control for mother's age, race, education level, insurance type, parity, maternal morbidity, infection, and risk, infant's sex, state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects. N=11,965,997. The gestational age is the obstetric estimates.

Table A.9: State Weights for the Synthetic Maryland

| | Overall | VLBW | MLBW | NBW | HBW |
|----------------------|---------|-------|-------|-------|-------|
| California | 0.025 | 0.036 | 0.049 | 0 | 0.016 |
| Colorado | 0.028 | 0.029 | 0.046 | 0 | 0.018 |
| Delaware | 0.005 | 0.144 | 0.02 | 0 | 0.011 |
| District Of Columbia | 0.016 | 0.021 | 0.049 | 0 | 0.06 |
| Illinois | 0.281 | 0.033 | 0.046 | 0.395 | 0.014 |
| Indiana | 0.034 | 0.022 | 0.043 | 0 | 0.016 |
| Iowa | 0.029 | 0.025 | 0.031 | 0 | 0.021 |
| Kentucky | 0.038 | 0.023 | 0.05 | 0 | 0.021 |
| Michigan | 0.02 | 0.027 | 0.071 | 0 | 0.305 |
| Montana | 0.019 | 0.016 | 0.047 | 0 | 0.018 |
| Nevada | 0.255 | 0.035 | 0.036 | 0.352 | 0.264 |
| New Hampshire | 0.027 | 0.03 | 0.046 | 0.093 | 0.013 |
| New Mexico | 0.013 | 0.242 | 0.133 | 0 | 0.014 |
| New York | 0.037 | 0.032 | 0.036 | 0 | 0.017 |
| North Dakota | 0.017 | 0.036 | 0.026 | 0 | 0.01 |
| Ohio | 0.018 | 0.028 | 0.045 | 0 | 0.016 |
| Oregon | 0.038 | 0.11 | 0.037 | 0.16 | 0.122 |
| Pennsylvania | 0.054 | 0.029 | 0.041 | 0 | 0.019 |
| Vermont | 0.029 | 0.021 | 0.112 | 0 | 0.008 |
| Washington | 0.017 | 0.062 | 0.037 | 0 | 0.018 |

Table A.10: Effect of GBR on NICU Admissions from SCM, 2011-2017

| NICU Admission, % | Average Baseline Admission Rate in Maryland | Average Post- Period Difference | Permutation Based P-Value | Relative Effect from SCM |
|----------------------|---|---------------------------------------|------------------------------|-----------------------------|
| Overall | 7.5 | -0.78 | 0.05 | -10.4%* |
| VLBW | 83.7 | -3.41 | 0.24 | -4.10% |
| MLBW | 37.4 | -6.2 | 0.05 | -16.6%* |
| NBW | 4.8 | -0.49 | 0.33 | -10.20% |
| HBW | 5.9 | -0.05 | 0.19 | -0.90% |

Note. The birthweight categories were defined as: very-low birthweight (VLBW, 500-1,499g), moderately-low birthweight (MLBW, 1,500-2,499g), normal birthweight (NBW, 2,500-3,999g), and high birthweight (HBW, 4,000g and above). Data comes from the 2011-2017 Vital Statistics. * denotes $p < 0.05$.

Table A.11: Comparing Estimates from DID versus SCM

| NICU Admission, % | Average Baseline Admission Rate in Maryland | Relative Effect from DID | Relative Effect from SCM |
|-------------------|---|--------------------------|--------------------------|
| Overall | 7.5 | -16.8%* | -10.4%* |
| VLBW | 83.7 | -1.1% | -4.10% |
| MLBW | 37.4 | -12.0%* | -16.6%* |
| NBW | 4.8 | -23.1%* | -10.20% |
| HBW | 5.9 | -14.20% | -0.90% |

Note. The birthweight categories were defined as: very-low birthweight (VLBW, 500-1,499g), moderately-low birthweight (MLBW, 1,500-2,499g), normal birthweight (NBW, 2,500-3,999g), and high birthweight (HBW, 4,000g and above). Data comes from the 2011-2017 Vital Statistics. * denotes $p < 0.05$.

Table A.12: The Difference-in-Differences Estimates including State-specific Linear Trends, 2011-2017

| | All | VLBW | MLBW | NBW | HBW |
|---------------------------|------------------------|-------------------------|-------------------------|------------------------|------------------------|
| Year 1 | -0.03 (-0.29,0.22) | -5.68 (-6.97,-4.39) | -2.19 (-2.91,-1.47) | -0.04 (-0.28,0.21) | 1.3 (1.04,1.56) |
| Year 2 | -0.67 (-1.19,-0.16) | -4.15 (-6.96,-1.34) | -5.01 (-6.29,-3.74) | -0.38 (-0.83,0.08) | 0.59 (0.13,1.06) |
| Year 3 | -1.34 (-2.04,-0.64) | -9.57 (-14.03,-5.11) | -9.91 (-11.90,-7.92) | -0.83 (-1.43,-0.23) | -0.88 (-1.56,-0.21) |
| Baseline Mean in Maryland | 7.5 | 83.7 | 37.4 | 4.8 | 5.9 |
| Relative Effect at Year 3 | -17.80% | -11.40% | -26.50% | -17.30% | -14.90% |
| N | 11,965,997 | 104,356 | 591,120 | 10,238,271 | 1,032,250 |

Note. The birthweight categories were defined as: very-low birthweight (VLBW, 500-1,499g), moderately-low birthweight (MLBW, 1,500-2,499g), normal birthweight (NBW, 2,500-3,999g), and high birthweight (HBW, 4,000g and above). Standard errors in the parentheses are clustered at the state level. Data comes from the 2011-2017 Vital Statistics.

* denotes $p < 0.05$.

Table A.13: Characteristics of Mothers Before and After GBR, 2011-2017

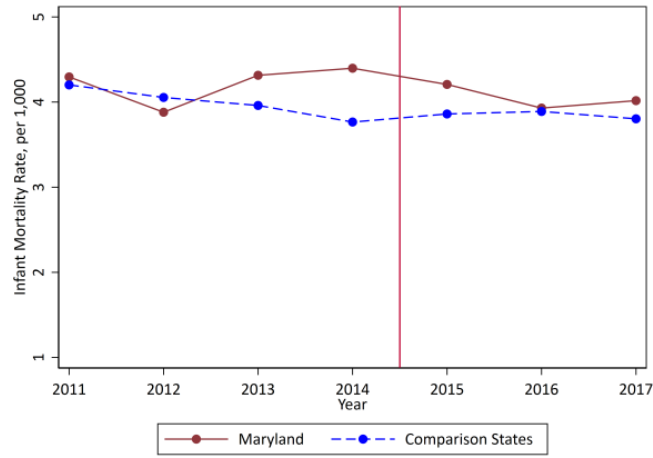
| Characteristic, % | Maryland | | Comparison States | | Bootstrap P-value |
|---|--------------------------------|---------------------------------|--------------------------------|---------------------------------|----------------------|
| | Pre Period 2011- 2014 | Post Period 2015- 2017 | Pre Period 2011- 2014 | Post Period 2015- 2017 | |
| Maternal Characteristics | | | | | |
| Maternal Race | | | | | |
| Non-Hispanic White | 47.66 | 44.55 | 54.17 | 52.83 | 0.06 |
| Non-Hispanic Black | 30.94 | 30.10 | 10.86 | 10.48 | 0.38 |
| Non-Hispanic Other | 7.97 | 9.70 | 9.48 | 11.51 | 0.61 |
| Hispanic | 13.43 | 15.65 | 25.49 | 25.18 | 0.00 |
| Maternal Age, years | | | | | |
| <20 | 5.74 | 4.21 | 6.89 | 4.96 | 0.21 |
| 20-24 | 18.97 | 16.37 | 21.73 | 19.37 | 0.57 |
| 25-34 | 57.40 | 59.13 | 55.27 | 57.63 | 0.38 |
| 35-44 | 17.65 | 20.01 | 15.93 | 17.83 | 0.39 |
| 45+ | 0.24 | 0.29 | 0.18 | 0.22 | 0.92 |
| Maternal Education | | | | | |
| Less than High School | 13.28 | 12.37 | 16.72 | 13.84 | 0.14 |
| High School | 20.81 | 20.70 | 24.32 | 24.52 | 0.80 |
| Some College | 27.98 | 27.04 | 28.58 | 28.62 | 0.36 |
| University and Above | 37.93 | 39.88 | 30.38 | 33.02 | 0.49 |
| Maternal Insurance | | | | | |
| Medicaid | 33.44 | 39.29 | 43.38 | 42.63 | 0.03 |
| Private Insurance | 57.72 | 53.34 | 49.53 | 50.80 | 0.04 |
| Self-Pay | 3.31 | 3.29 | 3.05 | 3.15 | 0.82 |
| Other | 5.53 | 4.08 | 4.04 | 3.42 | 0.50 |
| Parity | | | | | |
| First | 41.58 | 39.50 | 40.48 | 39.12 | 0.36 |
| Second | 33.21 | 33.64 | 31.84 | 32.33 | 0.92 |
| Third or Higher | 25.21 | 26.87 | 27.67 | 28.55 | 0.29 |
| Any Indications of Maternal Infection | 2.35 | 2.49 | 2.14 | 2.39 | 0.78 |
| Any Indications of Maternal Morbidity | 2.43 | 1.34 | 1.60 | 1.47 | 0.05 |
| Any Indications of Maternal Pregnancy Risk | 30.56 | 33.64 | 27.25 | 28.70 | 0.34 |
| Infant Characteristics | | | | | |
| Infant with Any Indications of Congenital Anomaly | 0.32 | 0.41 | 0.30 | 0.36 | 0.73 |
| Infant Male | 51.16 | 50.99 | 51.22 | 51.23 | 0.25 |
| Infant Birthweight | | | | | |
| Very-Low | 1.09 | 1.14 | 0.86 | 0.87 | 0.23 |

| | | | | | |
|-------------------------------------|---------|---------|---------|---------|------|
| Moderately-Low | 5.41 | 5.60 | 4.83 | 5.04 | 0.88 |
| Normal | 85.04 | 84.72 | 85.63 | 85.53 | 0.26 |
| High | 8.47 | 8.53 | 8.69 | 8.56 | 0.36 |
| Infant Gestational Age | | | | | |
| Very Preterm | 1.26 | 1.29 | 1.00 | 0.99 | 0.35 |
| Moderately Preterm | 6.65 | 6.87 | 6.13 | 6.29 | 0.86 |
| Term | 91.77 | 91.58 | 92.38 | 92.30 | 0.64 |
| Postterm | 0.32 | 0.27 | 0.49 | 0.42 | 0.76 |
| <i>State Characteristics</i> | | | | | |
| Poverty Rate | 800.00 | 733.39 | 1289.48 | 1126.99 | 0.11 |
| Unemployment Rate | 670.77 | 506.94 | 840.60 | 574.98 | 0.24 |
| Birth Rate | 1231.59 | 1202.90 | 1244.37 | 1197.71 | 0.49 |
| NICU Bed, per 1,000 pop | 6.58 | 6.47 | 5.80 | 5.64 | 0.94 |

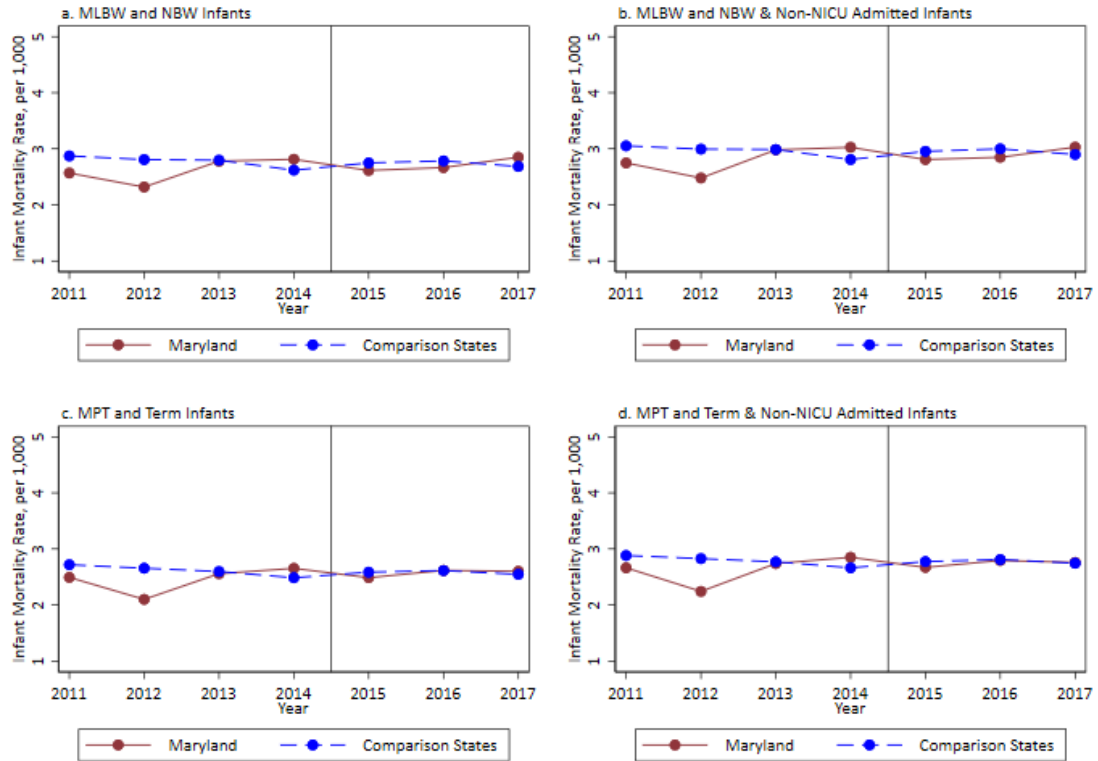
Note. The sample excludes those with birthweight less than 500g or unknown, non-singleton births, and those with missing values. The baseline period is from 2011 to 2014. Comparison states include 19 states and the District of Columbia that collected NICU admission information in our study period and adopted the ACA Medicaid Expansion as Maryland did. The full list of comparison states is provided in Table A.1. The birthweight categories were defined as: very-low birthweight (VLBW, 500-1,499g), moderately-low birthweight (MLBW, 1,500-2,499g), normal birthweight (NBW, 2,500-3,999g), and high birthweight (HBW, 4,000g and above). The gestation categories were defined as: very preterm (<32 weeks), moderately preterm (32-36 weeks), term (37-41 weeks), and postterm (>41 weeks).

Figure A.1: Unadjusted Trends of Infant Mortality Rate, 2011-2017

A. All Births



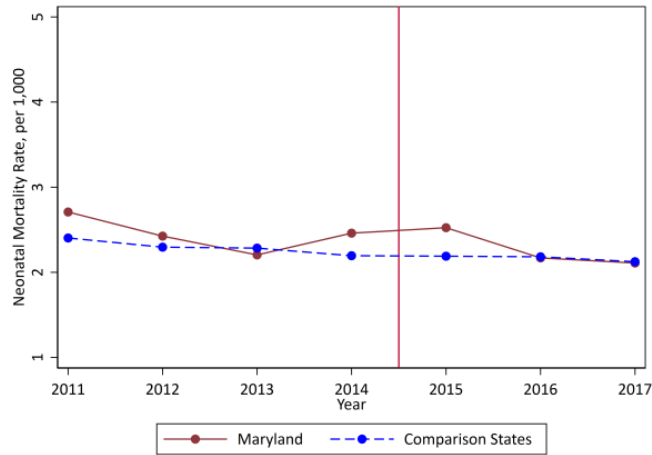
B. Subgroups



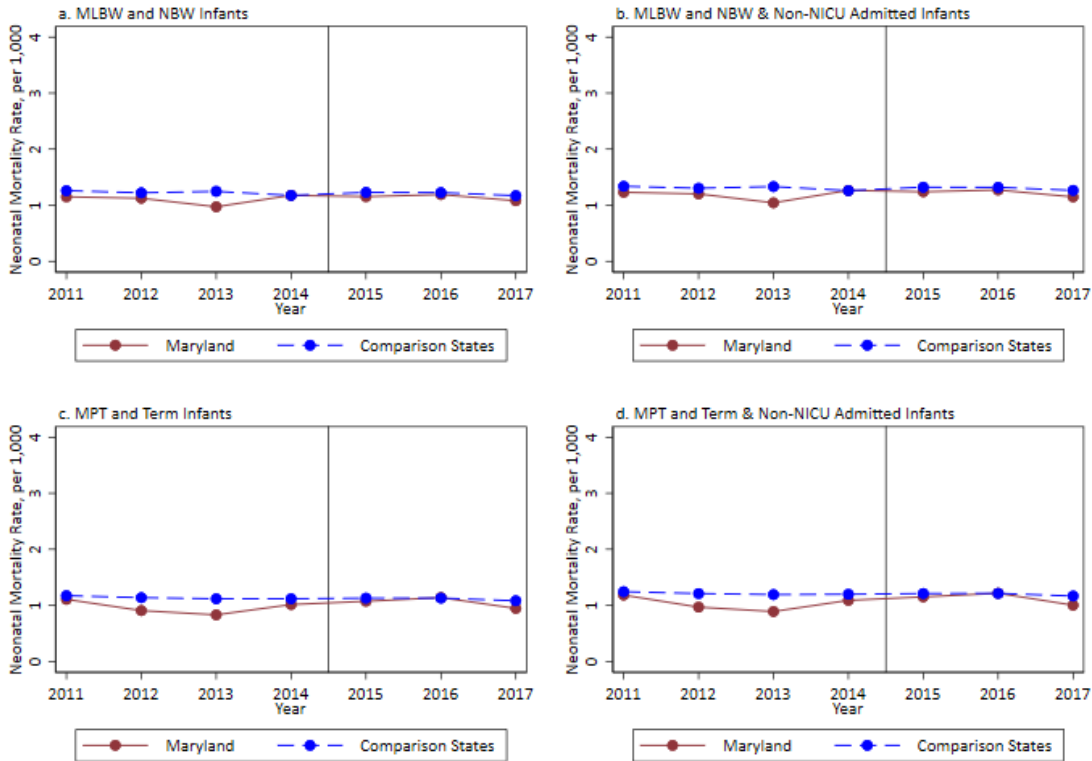
Source: 2011-2017 Natality Files and Period Linked Birth-Infant Death Data Files. Note. Comparison states included 19 states and DC. MLBW refers to moderately-low birthweight (1,500-2,499g); NBW refers to normal birthweight (2,500-3,999g); MPT refers to moderately preterm (32-36 weeks); term refers to 37-41 weeks.

Figure A.2: Unadjusted Trends of Neonatal Mortality Rate, 2011-2017

A. All Infants

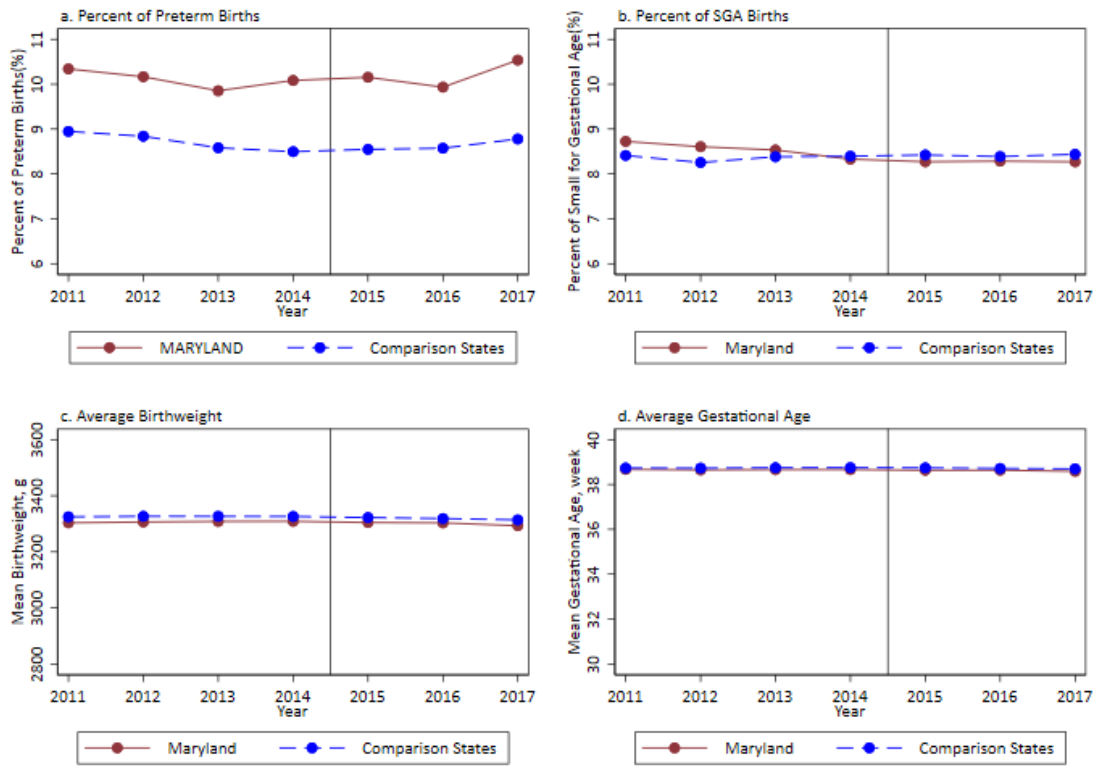


B. Subgroups



Source: 2011-2017 Natality Files and Period Linked Birth-Infant Death Data Files. Note. Comparison states included 19 states and DC. MLBW refers to moderately-low birthweight (1,500-2,499g); NBW refers to normal birthweight (2,500-3,999g); MPT refers to moderately preterm (32-36 weeks); term refers to 37-41 weeks.

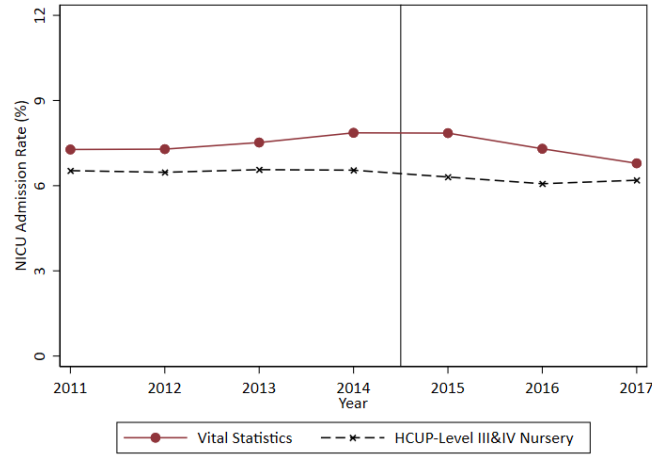
Figure A.3: Unadjusted Trends of Other Health Outcomes, 2011-2017



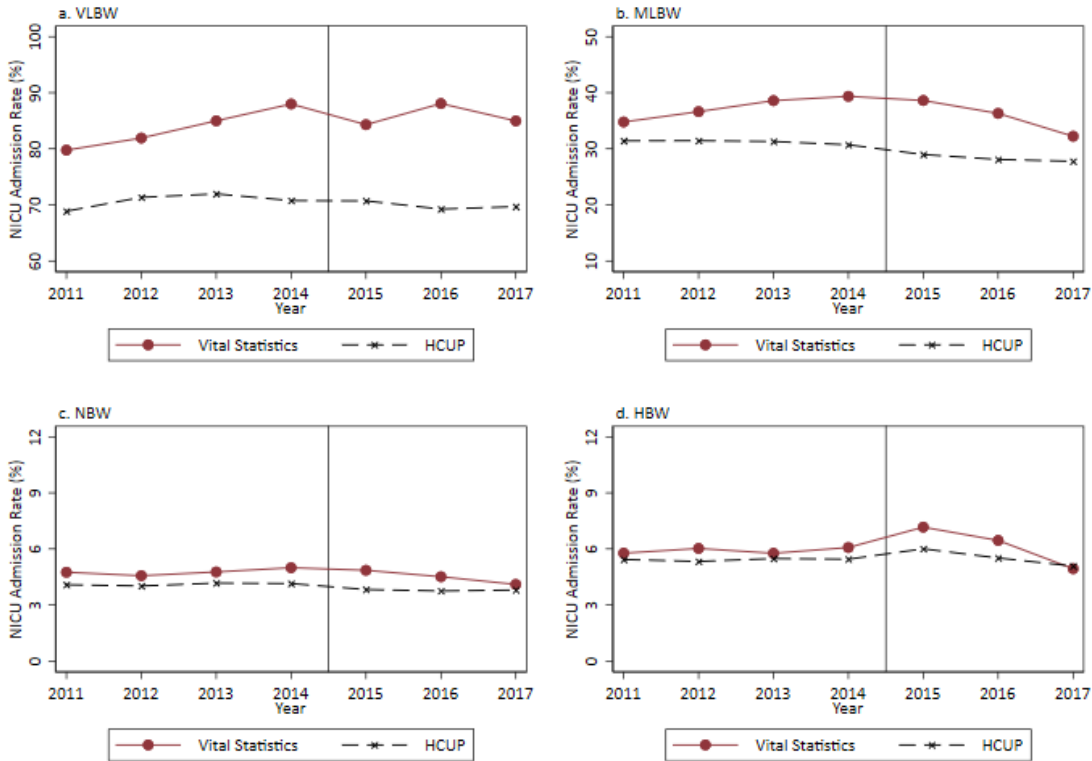
Source: 2011-2017 Natality Files. Note. Comparison states included 19 states and DC.

Figure A.4: NICU Admission of Vital Statistics and Hospital Discharge Data in Maryland, 2011-2017

A. Total



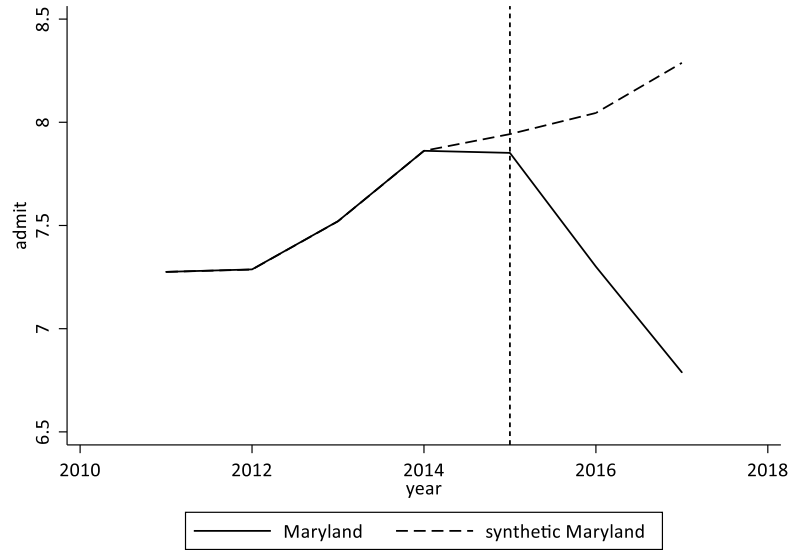
B. By Birthweight



Source: 2011-2017 Maryland’s State Inpatient Database of the Healthcare Cost and Utilization Project (HCUP). Note. The NICU admission in the HCUP data was defined as the use of level III or level IV nursery care. The birthweight categories were defined as: very-low birthweight (VLBW, 500-1,499g), moderately-low birthweight (MLBW, 1,500-2,499g), normal birthweight (NBW, 2,500-3,999g), and high birthweight (HBW, 4,000g and above).

Figure A.5: Trends in NICU Admission Rate: Maryland versus Synthetic Maryland

A. Overall



B. By Birthweight

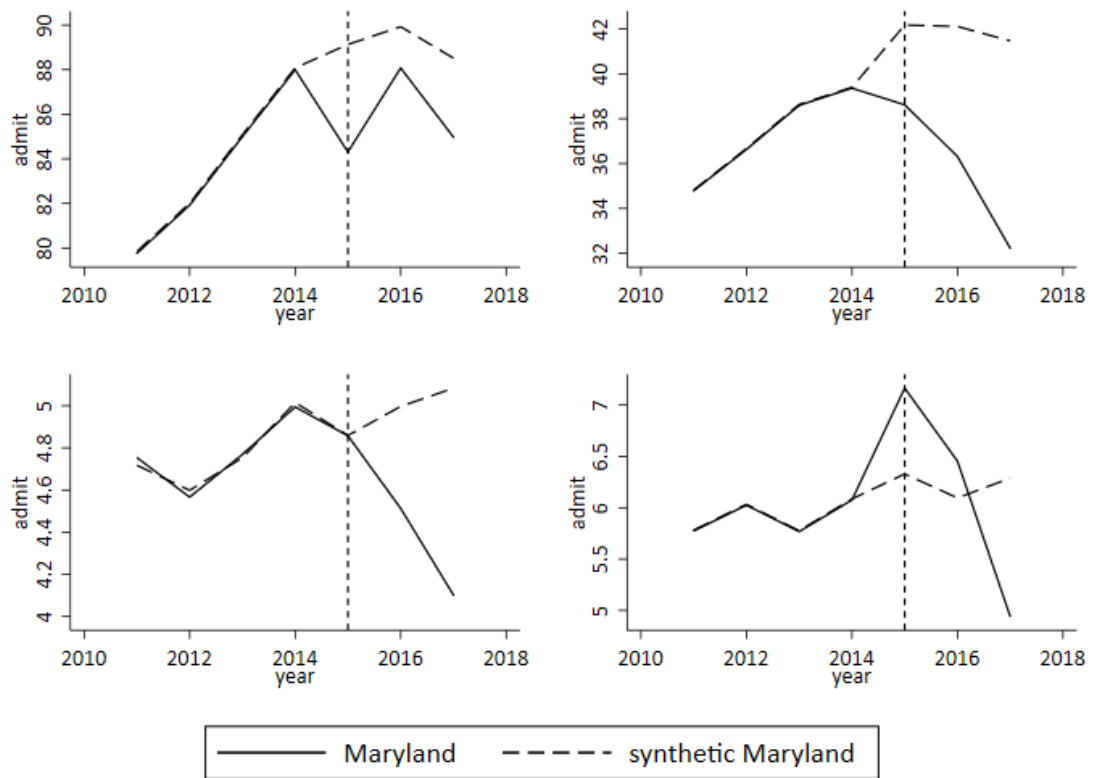


Figure A.6: Ratio of Post/Pre-GBR Mean Squared Prediction Error, Overall

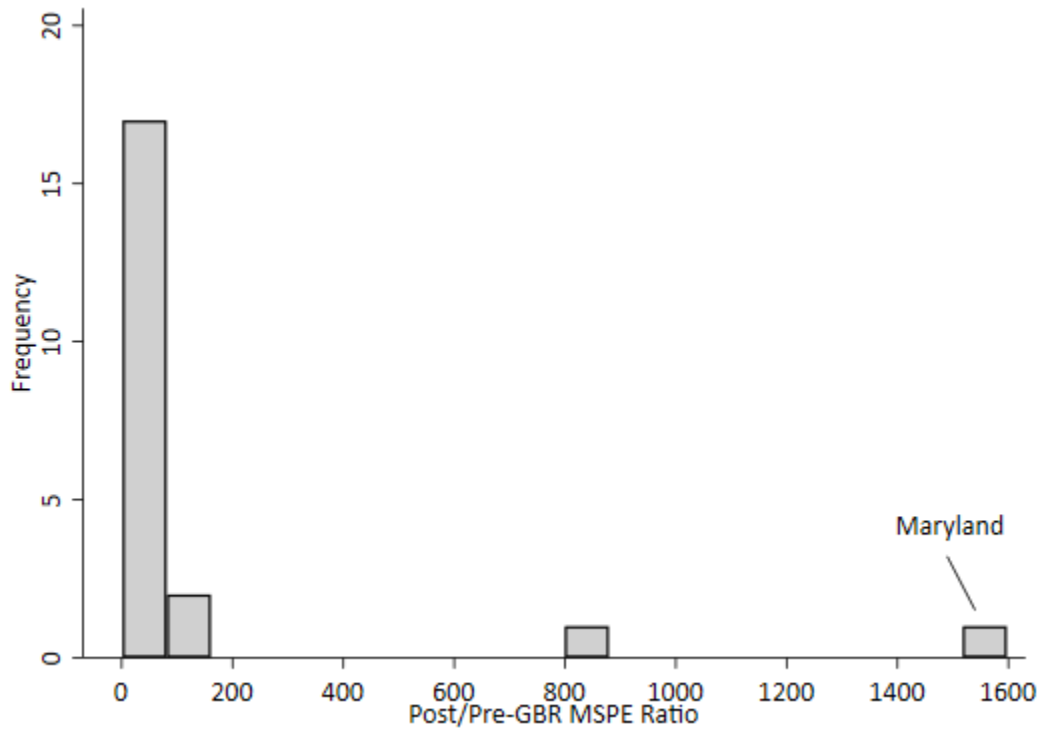
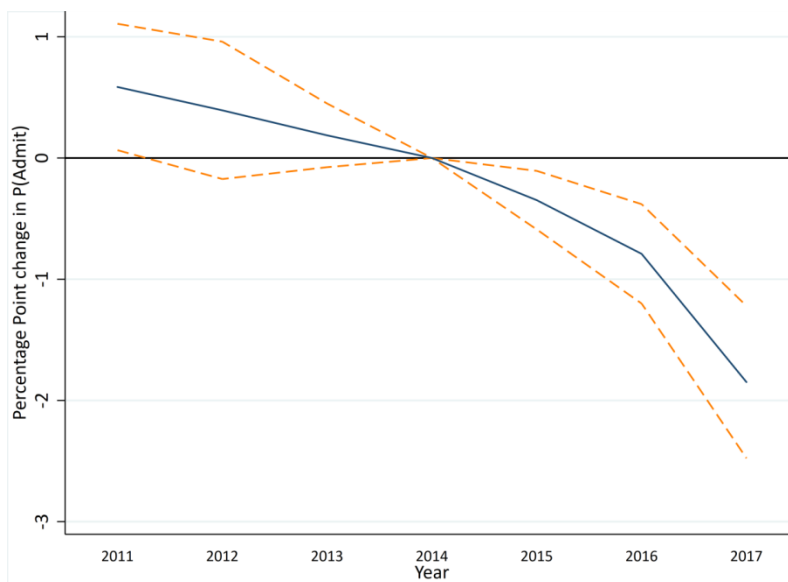
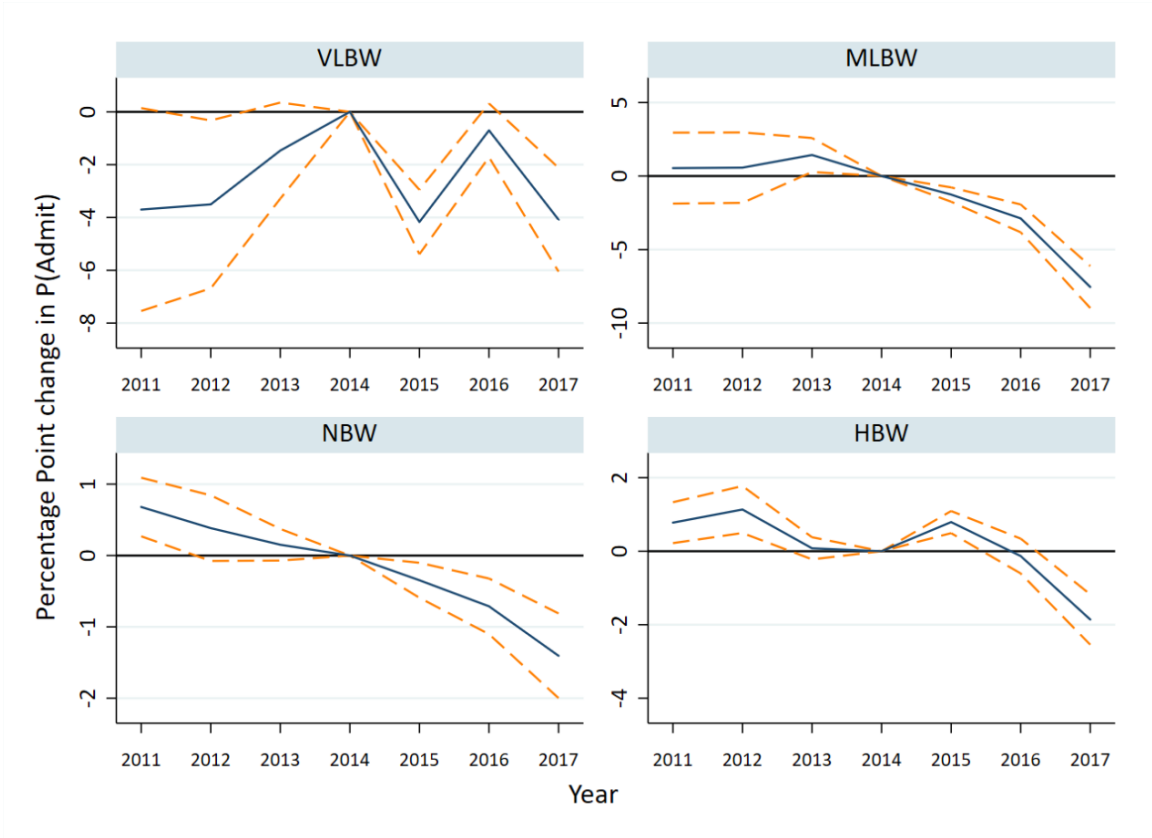


Figure A.7: Dynamic Difference-in-Differences

A. Overall



B. By Birthweight



Appendix B

B1. Data Sources

Table B.1: Sources of Covariates

| Variables | ICD-9 | ICD-10 |
|--------------------------------------|--|---|
| Hospital Born Infant | V3000, V3001, V3100, V3101, V3200, V3201, V3300, V3301, V3400, V3401, V3500, V3501, V3600, V3601, V3700, V3701, V3900, V3901 | Z3800, Z3801, Z3820, Z3830, Z3831, Z3850, Z3880, Z386 |
| Gestational Age | | |
| <=32 weeks | 76521 76522 76523 76524 76525 76526 | P0721-P0726 P0731-P0735 |
| 33-36 weeks | 76527 76528 | P0736-P0739 |
| >=37 weeks | 76621 76622 or not indicated | P0821 P0822 or not indicated |
| Birth Method | | |
| C-section | V3001 | Z3801 |
| Vaginal Birth | V3000 | Z3800 |
| Having Respiratory Distress Syndrome | 769 | P220 |

Note. The CMS' ICD-9-CM to and from ICD-10-CM and ICD-10-PCS Crosswalk are from: <https://data.nber.org/data/icd9-icd-10-cm-and-pcs-crosswalk-general-equivalence-mapping.html>.

Table B.2: Sources of Outcome Variables

| Variables | Revenue Codes |
|------------------------------------|--|
| <u>Nursery</u> | |
| NICU Care | 0173, 0174 |
| Special Care | 0172 |
| <u>Imaging</u> | |
| Chest X-ray | 0324 |
| Head CT Scan | 0351 |
| Body CT Scan | 0352 |
| Ultrasound | 0402 |
| Brain MRI | 0611 |
| Head/Neck MRI | 0615 |
| <u>Respiratory Services</u> | |
| General | 0410 |
| Inhalation services | 0412 |
| Hyperbaric oxygen therapy | 0413 |
| Other | 0419 |
| <u>By Cost Center</u> | |
| Clinics | 076*, 090*, 091*, 100* |
| Special Care Units | 020*, 021* |
| Routine Bed Units | 011*, 012*, 013*, 014*, 015*, 016* |
| Nursery, Labor/Delivery | 017*, 072* |
| All Other Ancillary | 028*, 048*, 068* |
| Operating Room | 036*, 037*, 045*, 071* |
| Therapies | 041*, 042*, 043*, 044*, 046* |
| Pharmacy | 025*, 027*, 062*, 063* |
| Laboratory | 030*, 031*, 032*, 033*, 034*, 035*, 040*, 061* |

Note. The revenue code for cost center is from Salemi et al. (2013).

Table B.3: Sources of Congenital Anomalies

| Category | CCS-Diagnosis | ICD-9 | ICD-10 |
|--|---------------|---|--|
| Cardiac and circulatory congenital anomalies | 213 | 7450 74510 74511 74512 74519 7452 7453 7454 7455 74560 74561 74569 7457 7458 7459 74600 74601 74602 74609 7461 7462 7463 7464 7465 7466 7467 74681 74682 74683 74684 74685 74686 74687 74689 7469 7470 74710 74711 74720 74721 74722 74729 7473 74731 74732 74739 74740 74741 74742 74749 7475 7476 74760 74761 74762 74763 74764 74769 74781 74782 74783 74789 7479 V1365 | Q256 Q212 Q219 Q204 Q244 Q205 Q262 Q249 Q234 Q265 Q2572 Q263 Q269 Q252 Q241 Q2579 Q254 Q240 Q250 Q231 Q255 Q245 Q251 Q218 Q233 Q232 Q230 Q229 Q210 Q238 Q213 Q203 Q2732 Q246 Q270 Q289 Q248 Q208 Q272 Q225 Q209 Q268 Q220 Q288 Q222 Z8774 Q282 Q200 Q243 Q278 Q283 Q260 Q2731 Q201 Q261 Q253 Q211 Q279 Q242 Q2571 Q221 P293 Q223 |
| Digestive congenital anomalies | 214 | 7500 75010 75011 75012 75013 75015 75016 75019 75021 75022 75023 75024 75025 75026 75027 75029 7503 7504 7505 7506 7507 7508 7509 7510 7511 7512 7513 7514 7515 75160 75161 75162 75169 7517 7518 7519 V1367 | Q450 Q383 Q408 Q443 Q400 Q459 Q395 Q392 Q419 Q444 Q453 Q451 Q441 Q429 Q390 Q438 Q437 Q388 Q435 Q458 Q446 Q396 Q445 Q442 Z87738 Q401 Q384 Q434 Q394 Q385 Q430 Q433 Q381 Q402 Q398 Q387 Q393 Q391 Q447 Q409 Q431 Q386 Q380 Q382 |
| Genitourinary congenital anomalies | 215 | 7520 75210 75211 75219 7522 7523 75231 75232 75233 75234 75235 75236 75239 75240 75241 75242 75243 75244 75245 75246 75247 75249 7525 75251 75252 7526 75261 75262 75263 75264 75265 75269 7527 7528 75281 75289 7529 7530 7531 75310 75311 75312 75313 75314 75315 75316 75317 75319 7532 75320 75321 75322 75323 75329 7533 7534 7535 7536 7537 7538 7539 V1361 V1362 | Q631 Q51821 Q605 Q624 Q6471 Q618 Q5210 Q633 Q51810 Q5521 Q644 Q6410 Q638 Q5279 Q6210 Q5212 Q6261 Q553 Q6119 Q5002 Q520 Q6474 Q551 Q514 Q513 Q512 Q559 Q630 Q516 Q515 Q6101 Q5031 Q5522 Q6473 Q524 Q602 Q563 Q628 Q619 Q6231 Q5523 Q51820 Q6419 Q6433 Q649 Q5032 Q526 Q504 Q6263 Q51811 Q6439 Q510 Q5529 Q544 Z87718 Q615 Q645 Q51828 Q558 Q5001 Q614 Q5562 Q5270 Q5110 Z87710 Q625 Q522 Q506 Q6239 Q550 Q640 Q6212 Q632 Q523 Q612 Q6432 Q529 Q5039 Q6431 Q528 Q5211 Q646 Q564 |

| | | | |
|-------------------------------------|-----|---|--|
| | | | Q6475 Q6479 Q6262 Q6100 Q5271 Q6211 Q613 Q5564 Q505 Q6102 Q5563 Q51818 Q539 Q642 Q5569 Q525 Q549 |
| Nervous system congenital anomalies | 216 | 7400 7401 7402 74100 74101 74102 74103 74190 74191 74192 74193 7420 7421 7422 7423 7424 74251 74253 74259 7428 7429 V1363 | Q001 Q042 Q062 Z87728 Q046 Q064 Q078 Q0703 Q052 Q038 Q079 Q041 Q031 Q019 Q054 Q058 Q0702 Q02 Q055 Q043 Q000 Q051 G901 Q068 Q002 Q057 Q063 Q0701 Q045 Q050 Q061 Q048 Q030 Q056 Q060 |
| Other congenital anomalies | 217 | 74300 74303 74306 74310 74311 74312 74320 74321 74322 74330 74331 74332 74333 74334 74335 74336 74337 74339 74341 74342 74343 74344 74345 74346 74347 74348 74349 74351 74352 74353 74354 74355 74356 74357 74358 74359 74361 74362 74363 74364 74365 74366 74369 7438 7439 74400 74401 74402 74403 74404 74405 74409 7441 74421 74422 74423 74424 74429 7443 74441 74442 74443 74446 74447 74449 7445 74481 74482 74483 74484 74489 7449 7480 7481 7482 7483 7484 7485 74860 74861 74869 7488 7489 74900 74901 74902 74903 74904 74910 74911 74912 74913 74914 74920 74921 74922 74923 74924 74925 7540 7541 7542 75430 75431 75432 75433 75435 75440 75441 75442 75443 75444 75450 75451 75452 75453 75459 75460 75461 75462 75469 75470 75471 75479 75481 75482 75489 75500 75501 75502 75510 75511 75512 75513 75514 75520 75521 75522 75523 75524 75525 75526 75527 75528 75529 75530 75531 75532 75533 75534 75535 75536 75537 | Q7951 Q654 Z87721 Q7100 Q840 Q9381 Q780 Q799 Q998 Q750 Q784 Q72899 Q7240 Q759 Q171 Q738 Q731 Q937 Q875 Q6502 Q872 Q7190 Q1381 Q681 Q129 Q673 Q793 Q772 Q992 Q690 Q359 Q142 Q128 Q833 Q308 Q9389 Q972 Q978 Q798 Q123 Q831 Q102 Q743 Q164 Q6680 Q667 Q844 Q175 Z87720 Q749 Q180 Q76425 Q9388 Q824 Q792 Q917 Q301 Q950 Q6589 Q870 Q691 Q101 Q7000 Q321 Q107 Q672 Q110 Q688 Q310 Q6581 Q8901 Q742 Q162 Q767 Q165 Q794 Q333 Q663 Q7010 Q892 Q7200 Q820 Q159 Q161 Z87730 Q8740 Q185 Q369 Q692 Q971 Q302 Q1389 Q828 Q76428 Q313 Q6689 Q173 Q680 Q970 Q984 Q134 Q894 Q933 Q318 Q913 Q132 Q909 E7871 Q848 Q704 Q766 Q338 Q778 Q763 Q348 Q765 Z87790 Q809 Q782 Q140 Q928 Q819 Q969 Q676 Q845 Q113 Q181 Q341 Q178 Q360 Q7160 Q841 Q781 Q7260 Q7250 Q7210 Q851 Q822 Q899 Q7120 Q783 Q163 Q158 Q183 Q121 Q846 Z8775 Q379 Q858 Q661 Q662 Q660 Q774 Q796 Q6501 Q7150 Q821 Q120 Z8776 Q776 Q160 Q999 Q684 Q135 Q76427 Q985 Q7649 Q871 Q8789 Q849 Q300 Q791 Q987 Q169 Q133 |

| | | |
|--|--|---|
| | 75538 75539 7554 75550 75551 75552 75553 75554 75555 75556 75557 75558 75559 75560 75561 75562 75563 75564 75565 75566 75567 75569 7558 7559 7560 75610 75611 75612 75613 75614 75615 75616 75617 75619 7562 7563 7564 75650 75651 75652 75653 75654 75655 75656 75659 7566 7567 75670 75671 75672 75673 75679 75681 75682 75683 75689 7569 7570 7571 7572 75731 75732 75733 75739 7574 7575 7576 7578 7579 7580 7581 7582 7583 75831 75832 75833 75839 7584 7585 7586 7587 7588 75881 75889 7589 7590 7591 7592 7593 7594 7595 7596 7597 7598 75981 75982 75983 75989 7599 7952 V136 V1364 V1366 V1368 V1369 | Q188 Q762 Q788 Q934 Q76426 Q106 Q7959 Q678 Q893 Q6650 Q709 Q682 Q130 Q789 Q71899 Q7030 Q349 E7872 Q843 Q6530 Q76419 Q324 Q357 Q651 Q334 Q730 Q111 Q7270 Q336 Q803 Q830 Q184 Q7230 Q825 Q664 Q804 Q988 Q771 Q891 Q130 Q182 Q335 Q186 Q331 Q332 Q674 Q170 Q7140 Q675 Q897 Q873 Q832 Q378 Q6500 Q671 Q683 Q8781 Q7110 Q189 Q842 Q666 Q187 Q670 Q677 Q311 Q6582 Q172 Q699 Q685 R898 Q150 Q748 Q330 Q104 Q790 Q100 Q141 Q143 Q148 Q838 Q179 Q131 Q340 Q740 Q124 Q103 Q112 Q8909 Q898 Q339 Q761 Q7020 Q7220 Q6531 Q760 Q174 Q741 Z87798 Q6532 Q752 |
|--|--|---|

Note. The CCS categories from AHRQ: <https://www.hcup-us.ahrq.gov/toolssoftware/ccs/AppendixASingleDX.txt>

B2. Model Fit Comparison

Table B.4: Model Fit Comparison, Poisson Model versus Negative Binomial Model

| Model | Maximum Difference | At Value | Mean Diff |
|-------------------|--------------------|----------|-----------|
| Poisson | 0.321 | 2 | 0.072 |
| Negative Binomial | 0.34 | 2 | 0.081 |

Note. The Stata user-written command *countfit* was used to generate this result.

B3. Sensitivity Analysis using 2014 as the Implementation Year

Table B.5: Effect of GBR on Length of Stay (2014 as the Implementation Year), 2011-2016

| | OLS | | | | Poisson | | | |
|--|--------------------|-------------------|------------------|-------------------|--------------------|---------------------|--------------------|--------------------|
| | Overall | <=32 weeks | 33-36 weeks | >=37 weeks | Overall | <=32 weeks | 33-36 weeks | >=37 weeks |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| <i>Without State-Specific Linear Trend</i> | | | | | | | | |
| DID Point Estimates | -0.08 (0.05) | -2.69** (0.36) | -0.23 (0.11) | -0.05 (0.03) | -0.11** (0.04) | -2.53*** (0.46) | -0.18 (0.12) | -0.05 (0.03) |
| Ferman-Pinto P-Value | 0.016 | <0.001 | 0.001 | 0.177 | | | | |
| <i>With State-Specific Linear Trend</i> | | | | | | | | |
| DID Point Estimates | -0.26*** (0.02) | NA | -0.33* (0.08) | -0.11** (0.01) | -0.24*** (0.02) | -3.98 *** (1.03) | -0.33*** (0.07) | -0.11*** (0.01) |
| Ferman-Pinto P-Value | <0.001 | | <0.001 | <0.001 | | | | |
| N | 2,538,585 | 48,228 | 154,961 | 2,335,396 | 2,538,585 | 48,228 | 154,961 | 2,335,396 |

Note. Standard errors are in the parentheses, clustered at the state level. Models control for birthweight, gestational age, sex, race, insurance type, relative household income, an indicator of congenital anomalies, an indicator of respiratory distress syndrome, urban; state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects. Marginal effects are reported for Poisson models. N denotes the number of observations.*** p<0.001, ** p<0.01, * p<0.05

Table B.6: Effect of GBR on Length of Stay by Birthweight (2014 as the Implementation Year), 2011-2016

| | OLS | | | Poisson | | |
|--|--------|------------|-----------|----------|------------|-----------|
| | <1500g | 1500-2499g | >=2500g | <1500g | 1500-2499g | >=2500g |
| <i>By Infant Birthweight</i> | (1) | (2) | (3) | (4) | (5) | (6) |
| <i>Without State-specific linear trend</i> | | | | | | |
| DID point estimates | -2.43* | -0.45 | -0.04 | -2.85*** | -0.62* | -0.04 |
| | (0.49) | (0.20) | (0.03) | (0.52) | (0.29) | (0.03) |
| Ferman-Pinto P-value | <0.001 | <0.001 | 0.156 | | | |
| N | 34,717 | 169,761 | 2,334,107 | 34,717 | 169,761 | 2,334,107 |
| <i>With State-specific linear trend</i> | | | | | | |
| DID point estimates | | -0.95** | -0.11** | -3.18*** | -0.98*** | -0.11*** |
| | | (0.08) | (0.01) | (0.31) | (0.07) | (0.01) |
| Ferman-Pinto P-value | | <0.001 | <0.001 | | | |
| N | | 169,761 | 2,334,107 | 34,717 | 169,761 | 2,334,107 |

Note. Standard errors in the parentheses, clustered at state-level. Models control for birthweight, gestational age, sex, race, insurance type, relative household income, an indicator of congenital anomalies, an indicator of respiratory distress syndrome, urban; state-level poverty rate, unemployment rate, birth rate, NICU bed per 1,000 population, and state and year fixed effects. Marginal effects are reported for Poisson models. N denotes as the number of observations. *** p<0.001, ** p<0.01, * p<0.05

Table B.7: Effect of GBR on Total Cost of Birth (2014 as the Implementation Year), 2011-2016

| | OLS | | | | Poisson | | | |
|--|-----------|------------|-------------|------------|------------|--------------|-------------|------------|
| | Overall | <=32 weeks | 33-36 weeks | >=37 weeks | Overall | <=32 weeks | 33-36 weeks | >=37 weeks |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| <i>Without State-Specific Linear Trend</i> | | | | | | | | |
| DID Point Estimates | -0.08* | -0.17** | -0.08** | -0.08* | -434.12*** | -11993.91*** | -1165.33*** | -209.50*** |
| | (0.02) | (0.02) | (0.01) | (0.02) | (67.56) | (1767.47) | (129.96) | (32.95) |
| Ferman-Pinto P-Value | <0.001 | <0.001 | <0.001 | <0.001 | | | | |
| <i>With State-Specific Linear Trend</i> | | | | | | | | |
| DID Point Estimates | 0.10** | -0.19* | 0.11* | 0.11** | 362.93*** | -8327.91 | 933.46** | 188.89*** |
| | (0.02) | (0.05) | (0.04) | (0.02) | (51.50) | (4360.25) | (306.11) | (29.69) |
| Ferman-Pinto P-Value | <0.001 | <0.001 | <0.001 | <0.001 | | | | |
| N | 2,538,585 | 48,228 | 154,961 | 2,335,396 | 2,538,585 | 48,228 | 154,961 | 2,335,396 |

Note. Standard errors are in the parentheses, clustered at the state level. Models control for birthweight, gestational age, sex, race, insurance type, relative household income, an indicator of congenital anomalies, an indicator of respiratory distress syndrome, urban; state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects. Marginal effects are reported for Poisson models. N denotes the number of observations. *** p<0.001, ** p<0.01, * p<0.05

Table B.8: Effect of GBR on Total Cost of Birth by Birthweight (2014 as the Implementation Year), 2011-2016

| | OLS | | | GLM | | |
|--|--------|------------|-----------|--------------|-------------|------------|
| | <1500g | 1500-2499g | >=2500g | <1500g | 1500-2499g | >=2500g |
| <i>By Infant Birthweight</i> | (1) | (2) | (3) | (4) | (5) | (6) |
| <i>Without State-specific linear trend</i> | | | | | | |
| DID point estimates | -0.15* | -0.14*** | -0.08* | -17095.17*** | -2593.06*** | -201.16*** |
| | (0.03) | (0.01) | (0.02) | (2392.64) | (95.00) | (32.23) |
| Ferman-Pinto P-value | <0.001 | <0.001 | <0.001 | | | |
| N | 34,717 | 169,761 | 2,334,107 | 34,717 | 169,761 | 2,334,107 |
| <i>With State-specific linear trend</i> | | | | | | |
| DID point estimates | | 0.04 | 0.12** | -10488.03*** | -254.75* | 201.95*** |
| | | (0.03) | (0.02) | (1374.44) | (129.43) | (30.72) |
| Ferman-Pinto P-value | | <0.001 | <0.001 | | | |
| N | | 169,761 | 2,334,107 | 34,717 | 169,761 | 2,334,107 |

Note. Standard errors in the parentheses, clustered at state-level. Models control for birthweight, gestational age, sex, race, insurance type, relative household income, an indicator of congenital anomalies, an indicator of respiratory distress syndrome, urban; state-level poverty rate, unemployment rate, birth rate, NICU bed per 1,000 population, and state and year fixed effects. Marginal effects are reported for GLM models. N denotes as the number of observations. *** p<0.001, ** p<0.01, * p<0.05

Table B.9: Effect of GBR on Services Utilization (2014 as the Implementation Year), 2011-2016

| | NICU Care | Special Care | Radiology Diagnostic | CT Scan | Other Imaging Services | Magnetic Resonance Technology | Respiratory Services |
|--|-----------|--------------|----------------------|---------|------------------------|-------------------------------|----------------------|
| <i>All Births</i> | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| <i>Without State-specific linear trend</i> | | | | | | | |
| DID estimates | -0.86* | -0.97* | -0.36 | -0.05* | -0.61 | -0.05 | 2.01* |
| | (0.19) | (0.27) | (0.14) | (0.01) | (0.26) | (0.02) | (0.44) |
| Ferman-Pinto P-value | <0.001 | <0.001 | <0.001 | <0.001 | 0.071 | <0.001 | <0.001 |
| <i>With State-specific linear trend</i> | | | | | | | |
| DID estimates | NA | NA | NA | NA | NA | -0.03* | -3.56** |
| | | | | | | (0.01) | (0.47) |
| Ferman-Pinto P-value | | | | | | <0.001 | <0.001 |

Note. N=2,538,585. Percentage points are reported. Standard errors in the parentheses, clustered at the state level. Models control for birthweight, gestational age, sex, race, insurance type, relative household income, an indicator of congenital anomalies, an indicator of respiratory distress syndrome, urban; state-level poverty rate, unemployment rate, birth rate, NICU beds per 1,000 population, and state and year fixed effects. N denotes the number of observations.*** p<0.001, ** p<0.01, * p<0.05

Appendix C

Table C.1: Percent of Births by State and Insurance Type, 2016-2017

| State | Medicaid | Private | Self-pay | Other |
|----------------------|----------|---------|----------|-------|
| Alabama | 50.42 | 45.12 | 2.08 | 2.39 |
| Alaska | 39.33 | 37.40 | 3.18 | 20.09 |
| Arizona | 52.54 | 40.30 | 4.39 | 2.76 |
| Arkansas | 46.98 | 49.06 | 2.94 | 1.03 |
| California | 42.96 | 48.50 | 4.27 | 4.27 |
| Colorado | 39.18 | 51.64 | 2.60 | 6.58 |
| Connecticut | 35.88 | 57.00 | 5.05 | 2.07 |
| Delaware | 43.26 | 51.36 | 1.91 | 3.47 |
| District Of Columbia | 37.22 | 57.08 | 0.90 | 4.80 |
| Florida | 48.98 | 41.72 | 6.21 | 3.08 |
| Georgia | 45.43 | 39.03 | 6.69 | 8.86 |
| Hawaii | 31.38 | 42.17 | 3.32 | 23.12 |
| Idaho | 37.20 | 53.38 | 5.83 | 3.59 |
| Illinois | 41.18 | 56.36 | 1.55 | 0.91 |
| Indiana | 40.82 | 52.54 | 4.64 | 1.99 |
| Iowa | 40.66 | 55.46 | 2.95 | 0.93 |
| Kansas | 30.56 | 57.25 | 6.77 | 5.42 |
| Kentucky | 49.80 | 43.57 | 3.53 | 3.09 |
| Louisiana | 62.48 | 34.06 | 0.84 | 2.62 |
| Maine | 39.51 | 54.33 | 4.04 | 2.12 |
| Maryland | 40.64 | 52.44 | 3.10 | 3.81 |
| Massachusetts | 27.50 | 67.97 | 0.80 | 3.74 |
| Michigan | 42.52 | 55.62 | 1.40 | 0.45 |
| Minnesota | 32.24 | 63.32 | 2.38 | 2.06 |
| Mississippi | 64.48 | 31.41 | 3.10 | 1.02 |
| Missouri | 39.26 | 55.27 | 3.47 | 2.00 |
| Montana | 41.27 | 47.84 | 5.23 | 5.66 |
| Nebraska | 33.94 | 59.14 | 3.88 | 3.05 |
| Nevada | 47.48 | 43.73 | 5.03 | 3.76 |
| New Hampshire | 27.84 | 67.02 | 1.70 | 3.44 |
| New Jersey | 31.12 | 59.61 | 8.26 | 1.00 |
| New Mexico | 59.49 | 27.18 | 6.67 | 6.66 |
| New York | 48.32 | 47.11 | 1.19 | 3.38 |
| North Carolina | 42.70 | 45.65 | 6.86 | 4.79 |
| North Dakota | 24.57 | 57.22 | 2.35 | 15.85 |
| Ohio | 41.53 | 51.03 | 4.67 | 2.77 |

| | | | | |
|----------------|-------|-------|------|------|
| Oklahoma | 51.55 | 39.97 | 2.06 | 6.42 |
| Oregon | 45.00 | 51.72 | 2.04 | 1.24 |
| Pennsylvania | 34.54 | 58.53 | 4.68 | 2.25 |
| Rhode Island | 46.88 | 50.62 | 0.71 | 1.80 |
| South Carolina | 51.09 | 41.61 | 2.95 | 4.36 |
| South Dakota | 30.82 | 60.81 | 2.96 | 5.41 |
| Tennessee | 49.73 | 43.22 | 2.03 | 5.02 |
| Texas | 46.65 | 39.15 | 8.21 | 5.98 |
| Utah | 25.95 | 64.33 | 5.40 | 4.32 |
| Vermont | 41.55 | 52.04 | 1.89 | 4.53 |
| Virginia | 30.09 | 62.51 | 5.33 | 2.08 |
| Washington | 39.72 | 52.58 | 1.11 | 6.59 |
| West Virginia | 51.21 | 45.85 | 2.07 | 0.88 |
| Wisconsin | 36.81 | 58.21 | 2.92 | 2.06 |
| Wyoming | 33.03 | 54.54 | 7.95 | 4.48 |

Note. The state is where the birth occurred. Data from the Vital Statistics 2016-2017.

Bibliography

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science* 59 (2): 495–510. <https://doi.org/10.1111/ajps.12116>.
- Alexander, Greg R., Michael D. Kogan, John H. Himes, Joanne M. Mor, and Robert Goldenberg. 1999. "Racial Differences in Birthweight for Gestational Age and Infant Mortality in Extremely-Low-Risk US Populations." *Paediatric and Perinatal Epidemiology* 13 (2): 205–17. <https://doi.org/10.1046/j.1365-3016.1999.00174.x>.
- American Academy of Pediatrics. 2004. "Levels of Neonatal Care." <https://doi.org/10.1542/peds.2004-1697>.
- . 2012. "Levels of Neonatal Care." <https://doi.org/10.1542/peds.2012-1999>.
- Anderson, Michael, Carlos Dobkin, and Tal Gross. 2012. "The Effect of Health Insurance Coverage on the Use of Medical Services." *American Economic Journal: Economic Policy* 4 (1): 1–27. <https://doi.org/10.1257/pol.4.1.1>.
- Andrews, Roxanne M. 2015. "Statewide Hospital Discharge Data: Collection, Use, Limitations, and Improvements." *Health Services Research* 50 (S1): 1273–99. <https://doi.org/10.1111/1475-6773.12343>.
- Angert, Robert, and Henry M Adam. 2008. "Care of the Very Low-Birthweight Infant." *Pediatrics in Review* 30 (1): 32–35. <https://doi.org/10.1542/pir.30-1-32>.
- Baker, Laurence C, and Ciaran S Phibbs. 2002. "Managed Care, Technology Adoption, and Health Care: The Adoption of Neonatal Intensive." *Source: The RAND Journal of Economics*. Vol. 33. <https://www.jstor.org/stable/pdf/3087471.pdf?refreqid=excelsior%3Af3325d7c54d53ba240954d493e6d1fae>.
- Baltimore City Health Department. 2009. "A Strategic Approach Baltimore City Infant Mortality Rate (Deaths/1,000 Live Births)." http://www.healthybabiesbaltimore.com/uploads/files/Bmore_5_Year_Brochure_for_viewing_lr.pdf.
- Barradas, Danielle T, Martin P Wasserman, Lekisha Daniel-Robinson, Marino A Bruce, Katherine Isselmann Disantis, Frederick H Navarro, Warren A Jones, et al. 2016. "Hospital Utilization and Costs Among Preterm Infants by Payer: Nationwide Inpatient Sample, 2009 HHS Public Access." *Matern Child Health J* 20 (4): 808–18. <https://doi.org/10.1007/s10995-015-1911-y>.
- Beil, Heather, Susan G. Haber, Katherine Giuriceo, Peter Amico, Marisa Morrison, Christopher Beadles, Olivia Berzin, et al. 2019. "Maryland's Global Hospital Budgets: Impacts on Medicare Cost and Utilization for the First 3 Years." *Medical Care* 57 (6): 417–24. <https://doi.org/10.1097/MLR.0000000000001118>.

- Berenson, Robert A. 2015. "Maryland's New All-Payer Hospital Demonstration." <https://www.urban.org/sites/default/files/publication/73836/2000517-Maryland%27s-New-All-Payer-Hospital-Demonstration.pdf>.
- Bilgel, Firat, and Brian Galle. 2015. "Financial Incentives for Kidney Donation: A Comparative Case Study Using Synthetic Controls." *Journal of Health Economics* 43 (September): 103–17. <https://doi.org/10.1016/j.jhealeco.2015.06.007>.
- BLS. 2019. "Producer Price Index (PPI)." 2019. <https://www.bls.gov/ppi/data.htm>.
- Botosaru, Irene, and Bruno Ferman. 2019. "On the Role of Covariates in the Synthetic Control Method." *Econometrics Journal* 22 (2): 117–30. <https://doi.org/10.1093/ectj/utz001>.
- Boudreaux, Michel, Yoon Sun Choi, Liyang Xie, and Daniel Marthey. 2019. "Medicaid Expansion at Title X Clinics." *Medical Care* 57 (6): 437–43. <https://doi.org/10.1097/mlr.0000000000001120>.
- Boudreaux, Michel H., Rada K. Dagher, and Scott A. Lorch. 2018. "The Association of Health Reform and Infant Health: Evidence from Massachusetts." *Health Services Research* 53 (4): 2406–25. <https://doi.org/10.1111/1475-6773.12779>.
- Braveman, Paula A., Susan Egerter, Trude Bennett, and Jonathan Showstack. 1991. "Differences in Hospital Resource Allocation Among Sick Newborns According to Insurance Coverage." *JAMA: The Journal of the American Medical Association* 266 (23): 3300. <https://doi.org/10.1001/jama.1991.03470230058031>.
- Brown, Clare C., Jennifer E. Moore, Holly C. Felix, M. Kathryn Stewart, T. Mac Bird, Curtis L. Lowery, and J. Mick Tilford. 2019. "Association of State Medicaid Expansion Status with Low Birth Weight and Preterm Birth." *JAMA - Journal of the American Medical Association* 321 (16): 1598–1609. <https://doi.org/10.1001/jama.2019.3678>.
- Buntin, Melinda Beeuwkes, and Alan M. Zaslavsky. 2004. "Too Much Ado about Two-Part Models and Transformation?: Comparing Methods of Modeling Medicare Expenditures." *Journal of Health Economics* 23 (3): 525–42.
- CDC. Centers for Disease Control and Prevention. 2003. "U.S. Standard Certificate of Live Birth." http://www.cdc.gov/nchs/vital_certs_rev.htm.
- . 2010. "User Guide to the 2010 Natality Public Use File." ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2010.pdf.
- . 2016a. "FastStats - Births and Natality." *Centers for Disease Control*. <https://www.cdc.gov/nchs/fastats/births.htm>.
- . 2016b. "User Guide to the 2016 Natality Public Use File." ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2016.pdf.
- . 2018. "CDC Wonder." 2018. <https://wonder.cdc.gov/>.

- . 2019. “NVSS - National Vital Statistics System.” 2019. <https://www.cdc.gov/nchs/nvss/index.htm>.
- Center for Health Statistics, National. 2003. “Birth Edit Specifications for the 2003 Proposed Revision of the U.S. Standard Certificate of Birth.” https://www.cdc.gov/nchs/data/dvs/birth_edit_specifications.pdf.
- Center for Medicare & Medicaid Innovation. 2018. “Innovation Models.” 2018. <https://innovation.cms.gov/initiatives/index.html#views=models>.
- Cloherly, John P, Eric C Eichenwald, Anne R Hansen, and Ann R Stark. 2012. *Manual of Neonatal Care*. 7th ed. LIPPINCOTT WILLIAMS & WILKINS. <https://doi.org/10.1097/01.pec.0000136078.50143.a7>.
- CMS. Centers for Medicare & Medicaid Services. 2018. “Maryland All-Payer Model.” 2018. <https://innovation.cms.gov/initiatives/Maryland-All-Payer-Model/>.
- . 2019a. “ICD-10.” 2019. <https://www.cms.gov/Medicare/Coding/ICD10/index?redirect=/ICD10>.
- . 2019b. “Pennsylvania Rural Health Model | Center for Medicare & Medicaid Innovation.” 2019. <https://innovation.cms.gov/initiatives/pa-rural-health-model/>.
- . 2019c. “Vermont All-Payer ACO Model | Center for Medicare & Medicaid Innovation.” 2019. <https://innovation.cms.gov/initiatives/vermont-all-payer-aco-model/>.
- . 2019d. “Acute Inpatient PPS.” 2019. <https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS>.
- Cohen, A Harold. 2005. “Maryland’s All-Payor Hospital Payment System.” <http://www.hsrc.state.md.us/Documents/pdr/GeneralInformation/MarylandAll-PayorHospitalSystem.pdf>.
- Currie, Janet, and Jonathan Gruber. 2001. “Public Health Insurance and Medical Treatment: The Equalizing Impact of the Medicaid Expansions.” *Journal of Public Economics* 82 (1): 63–89. [https://doi.org/10.1016/S0047-2727\(00\)00140-7](https://doi.org/10.1016/S0047-2727(00)00140-7).
- Cutler, David M., and Mark McClellan. 2001. “Is Technological Change in Medicine Worth It?” *Health Affairs* 20 (5): 11–29. <https://doi.org/10.1377/hlthaff.20.5.11>.
- Cutler, David M, and Ellen Meara. 2000. *The Technology of Birth: Is It Worth It? NBER/Frontiers in Health Policy Research*. Vol. 3. <https://doi.org/10.1162/109623100300091069>.
- Dartmouth Atlas Project. 2019. “Dartmouth Atlas of Health Care.” 2019. <https://www.dartmouthatlas.org/>.
- David C. Goodman, George A. Little, Wade N. Harrison, Atle Moen, Meredith E. Mowitz, Cecilia Ganduglia-Cazaban, Kristen K. Bronner, Julie R. Doherty. 2019. “The Dartmouth Atlas of Neonatal Intensive Care.”

- Dieleman, Joseph L., Ellen Squires, Anthony L. Bui, Madeline Campbell, Abigail Chapin, Hannah Hamavid, Cody Horst, et al. 2017. "Factors Associated With Increases in US Health Care Spending, 1996-2013." *JAMA* 318 (17): 1668. <https://doi.org/10.1001/jama.2017.15927>.
- Dimick, Justin B., and Andrew M. Ryan. 2014. "Methods for Evaluating Changes in Health Care Policy: The Difference-in-Differences Approach." *JAMA - Journal of the American Medical Association* 312 (22): 2401–2. <https://doi.org/10.1001/jama.2014.16153>.
- Donahue, Sara M.A., Ken P. Kleinman, Matthew W. Gillman, and Emily Oken. 2010. "Trends in Birth Weight and Gestational Length among Singleton Term Births in the United States: 1990-2005." *Obstetrics and Gynecology* 115 (2 PART 1): 357–64. <https://doi.org/10.1097/AOG.0b013e3181cbd5f5>.
- Done, Nicolae, Bradley Herring, and Tim Xu. 2019. "The Effects of Global Budget Payments on Hospital Utilization in Rural Maryland." *Health Services Research* 54 (3): 526–36. <https://doi.org/10.1111/1475-6773.13162>.
- Dunn, Abe, Scott D. Grosse, and Samuel H. Zuvekas. 2018. "Adjusting Health Expenditures for Inflation: A Review of Measures for Health Services Research in the United States." *Health Services Research* 53 (1): 175–96. <https://doi.org/10.1111/1475-6773.12612>.
- Fenton, Tanis R, and Jae H Kim. 2013. "A Systematic Review and Meta-Analysis to Revise the Fenton Growth Chart for Preterm Infants." *BMC Pediatrics* 13 (1): 59. <https://doi.org/10.1186/1471-2431-13-59>.
- Ferman, Bruno, and Cristine Pinto. 2019. "Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity." *The Review of Economics and Statistics* 101 (3): 452–67. https://doi.org/10.1162/rest_a_00759.
- Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P Newhouse, Heidi Allen, and Katherine Baicker. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." *Quarterly Journal of Economics* 127 (3): 1057–1106. <https://doi.org/10.1093/qje/qjs020>.
- Freedman, Seth. 2016a. "Capacity and Utilization in Health Care: The Effect of Empty Beds on Neonatal Intensive Care Admission." *American Economic Journal: Economic Policy* 8 (2): 154–85. <https://doi.org/10.1257/pol.20120393>.
- . 2016b. "Online Appendix Capacity and Utilization in Health Care: The Effect of Empty Beds on Neonatal Intensive Care Admission A-1." https://assets.aeaweb.org/assets/production/articles-attachments/aej/pol/app/0802/2012-0393_app.pdf.
- Freedman, Seth, Haizhen Lin, and Kosali Simon. 2015. "Public Health Insurance Expansions and Hospital Technology Adoption." *Journal of Public Economics* 121 (January): 117–31. <https://doi.org/10.1016/J.JPUBECO.2014.10.005>.
- Giuriceo, Katherine, Susan Haber, Heather Beil, Walter Adamache, Peter Amico, Christopher Beadles, Olicia Berzin, et al. 2016. "Evaluation of the Maryland All-Payer Model: First Annual Report."

- <https://downloads.cms.gov/files/cmimi/marylandallpayer-firstannualrpt.pdf>.
- . 2018. “Evaluation of the Maryland All-Payer Model: Third Annual Report.” <https://downloads.cms.gov/files/cmimi/md-all-payer-thirdannrpt.pdf>.
- Giuriceo, Katherine, Susan Haber, ScD Heather Beil, Walter Adamache, Peter Amico, Christopher Beadles, Olivia Berzin, et al. 2017. “Evaluation of the Maryland All-Payer Model Second Annual Report.” <https://innovation.cms.gov/Files/reports/md-all-payer-secondannrpt.pdf>.
- Giuriceo, Katherine, Susan Haber, ScD Heather Beil, Marisa Morrison, Leslie Greenwald, Rebecca Perry, Linda Jiang, et al. 2019. “Evaluation of the Maryland All-Payer Model Volume I: Final Report Prepared For.”
- Goodman, David C., Elliott S. Fisher, George A. Little, Thérèse A. Stukel, Chiang-Hua Chang, and Kenneth S. Schoendorf. 2002. “The Relation Between the Availability of Neonatal Intensive Care and Neonatal Mortality.” *Obstetrical & Gynecological Survey* 57 (11): 731–32. <https://doi.org/10.1097/00006254-200211000-00010>.
- Goodman, David C., Cecilia Ganduglia-Cazaban, Luisa Franzini, Therese A. Stukel, Jared R. Wasserman, Megan A. Murphy, Youngran Kim, et al. 2019. “Neonatal Intensive Care Variation in Medicaid-Insured Newborns: A Population-Based Study.” *Journal of Pediatrics* 209 (June): 44-51.e2. <https://doi.org/10.1016/j.jpeds.2019.02.014>.
- Hall, Eric S., Matthew D. Lang, Linda Levin, and Vivek Narendran. 2015. “Variation in Neonatal Inpatient Charges at the State and Local Level.” *Hospital Topics* 93 (2): 27–35. <https://doi.org/10.1080/00185868.2015.1052273>.
- Halpern, Neil A, Debra A Goldman, Kay See Tan, and Stephen M Pastores. 2016. “Trends in Critical Care Beds and Use among Population Groups and Medicare and Medicaid Beneficiaries in the United States: 2000-2010.” *Critical Care Medicine* 44 (8): 1490–99. <https://doi.org/10.1097/CCM.0000000000001722>.
- Harrison, Wade, and David Goodman. 2015. “Epidemiologic Trends in Neonatal Intensive Care, 2007-2012.” *JAMA Pediatrics* 169 (9): 855–62. <https://doi.org/10.1001/jamapediatrics.2015.1305>.
- Harrison, Wade N., Jared R. Wasserman, and David C. Goodman. 2018. “Regional Variation in Neonatal Intensive Care Admissions and the Relationship to Bed Supply.” *Journal of Pediatrics* 192 (January): 73-79.e4. <https://doi.org/10.1016/j.jpeds.2017.08.028>.
- HCUP. Healthcare Cost and Utilization Project. 2019a. “HCUP-US SID Overview.” 2019. <https://www.hcup-us.ahrq.gov/sidoverview.jsp>.
- . 2019b. “HCUP-US Cost-to-Charge Ratio Files.” Agency for Healthcare Research and Quality, Rockville, MD. 2019. <https://www.hcup-us.ahrq.gov/db/state/costtocharge.jsp>.
- Health Care Cost Institute. 2019. “HMI Interactive Report.” 2019. <https://healthcostinstitute.org/research/hmi-interactive#HMI-Price-Index>.

- HRSA. Health Resources & Services Administration. 2019. “HRSA Data Warehouse – Area Health Resource Files.” 2019. <https://datawarehouse.hrsa.gov/topics/ahrf.aspx>.
- Howell, Embry M., Douglas Richardson, Paul Ginsburg, and Barbara Foot. 2002. “Deregionalization of Neonatal Intensive Care in Urban Areas.” *American Journal of Public Health* 92 (1): 119–24. <https://doi.org/10.2105/AJPH.92.1.119>.
- HSCRC. Health Services Cost Review Commission. 2012. “Rate Center Code Descriptions.” <http://www.hscrc.state.md.us/documents/Hospitals/RatesTargets/CenterCodeDescriptions-20120223.pdf>.
- . 2018a. “All-Payer Model Results, CY 2014-2017.” [http://www.hscrc.state.md.us/Documents/Modernization/Updated APM results through PY4.pdf](http://www.hscrc.state.md.us/Documents/Modernization/Updated_APM_results_through_PY4.pdf).
- . 2018b. “Maryland’s All-Payer Hospital Model Results Performance Year Three.” [http://www.hscrc.state.md.us/Documents/Modernization/Maryland APM Performance Report -CY2016_3_9_18.pdf](http://www.hscrc.state.md.us/Documents/Modernization/Maryland_APM_Performance_Report_-CY2016_3_9_18.pdf).
- Jackson, Heide. 2018. “Health Status and Medical Services Utilization: 2013.” www.census.gov/sipp.
- Jorgensen, Anne M. 2012. “A Century of Caring: 100 Years of NICU Nursing.” https://static.abbottnutrition.com/cms-prod/anhi-2017.org/img/history-of-neonatology_tcm1423-102720.pdf.
- Karaca, Zeynal, and Brian Moore. 2013. “Geographic Variation in Hospital Inpatient List Prices in the United States, 2013.” <http://www.nytimes.com/2015/06/02/business/medicare-payments-billing-hospitals-doctors.html>.
- Kastor, John A., and Eli Y. Adashi. 2011. “Maryland’s Hospital Cost Review Commission at 40: A Model for the Country.” *JAMA - Journal of the American Medical Association* 306 (10): 1137–38. <https://doi.org/10.1001/jama.2011.1311>.
- Kaul, Ashok, Stefan Klößner, Gregor Pfeifer, and Manuel Schieler. 2018. “Synthetic Control Methods: Never Use All Pre-Intervention Outcomes Together with Covariates.” *Working Paper*, no. 83790 (March): 1–20.
- Keith K. 2019. “Unpacking The House Medicare-For-All Bill.” Health Affairs Blog. 2019. <https://www.healthaffairs.org/do/10.1377/hblog20190302.150578/full/>.
- KFF. The Henry J. Kaiser Family Foundation. 2014. “Medicare Spending Per Enrollee, by State.” State Health Facts. 2014. <https://www.kff.org/medicare/state-indicator/per-enrollee-spending-by-residence/?currentTimeframe=0&sortModel=%7B%22colId%22:%22PerEnrolleeMedicareSpendingbyResidence%22,%22sort%22:%22desc%22%7D>.
- . 2017. “Distribution of the Total Population by Federal Poverty Level (above and below 100% FPL).” 2017. <https://www.kff.org/other/state->

indicator/population-above-and-below-100-fpl/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.

- . 2019. “Trends in Medicaid Income Eligibility Limits.” 2019. <https://www.kff.org/data-collection/trends-in-medicaid-income-eligibility-limits/>.
- Kilpatrick, Sarah J, Lu-Ann Papile, and George A Macones. 2017. *Guidelines for Perinatal Care*. Am Acad Pediatrics.
- Kowlessar, Niranjana M, Joanna Jiang, and Claudia Steiner. 2013. “Hospital Stays for Newborns, 2011,” 12. <https://www.hcup-us.ahrq.gov/reports/statbriefs/sb163.pdf>.
- Kreif, Noémi, Richard Grieve, Dominik Hangartner, Alex James Turner, Silviya Nikolova, and Matt Sutton. 2016. “Examination of the Synthetic Control Method for Evaluating Health Policies with Multiple Treated Units.” *Health Economics (United Kingdom)* 25 (12): 1514–28. <https://doi.org/10.1002/hec.3258>.
- Kuo, Dennis Z, Jay G Berry, • Matt Hall, • Robert, E Lyle, • Christopher, and J Stille. 2018. “Health-Care Spending and Utilization for Children Discharged from a Neonatal Intensive Care Unit.” *Journal of Perinatology* 38: 734–41. <https://doi.org/10.1038/s41372-018-0055-5>.
- Lantos, John. 2010. “Cruel Calculus: Why Saving Premature Babies Is Better Business than Helping Them Thrive.” *Health Affairs*. <https://doi.org/10.1377/hlthaff.2009.0897>.
- Lau, C., N. Ambalavanan, H. Chakraborty, M. S. Wingate, and W. A. Carlo. 2013. “Extremely Low Birth Weight and Infant Mortality Rates in the United States.” *PEDIATRICS* 131 (5): 855–60. <https://doi.org/10.1542/peds.2012-2471>.
- Long, J. Scott, and Jeremy Freese. 2014. *Stata Bookstore: Regression Models for Categorical Dependent Variables Using Stata, Third Edition*. Stata Press.
- Lorch, S A, P Maheshwari, and O Even-Shoshan. 2012. “The Impact of Certificate of Need Programs on Neonatal Intensive Care Units.” *Journal of Perinatology* 32 (1): 39–44. <https://doi.org/10.1038/jp.2011.47>.
- Lorch, Scott A., Corinne E. Ahlberg, Dylan S. Small, and Michael Baiocchi. 2012. “The Differential Impact of Delivery Hospital on the Outcomes of Premature Infants.” *Pediatrics* 130 (2): 270–78. <https://doi.org/10.1542/peds.2011-2820>.
- Martin, Joyce A, and Fay Menacker. 2004. “National Vital Statistics Reports, Vol. 55, No. 12 (4/19/2007).” Vol. 55. https://www.cdc.gov/nchs/data/nvsr/nvsr55/nvsr55_12.pdf.
- Maryland Department of Health. 2017. “Maryland Department of Health Medical Assistance.” [https://mmcp.health.maryland.gov/Documents/2017 UB04 Hospital Instructions-Final 8FEB2018.pdf](https://mmcp.health.maryland.gov/Documents/2017%20UB04%20Hospital%20Instructions-Final%208FEB2018.pdf).
- Maryland Health Care Commission. 2015. “Annual Report on Selected Maryland Acute Care and Special Hospital Services, Fiscal Year 2015.”

- http://mhcc.maryland.gov/mhcc/pages/hcfs/hcfs_hospital/documents/acute_care/chcf_con_acute_care_license_rpt_2015_revised_20150313.pdf.
- Mcguire, Thomas G. 2000. "Physician Agency." *Handbook of Health Economics*. Vol. 1. [https://doi.org/10.1016/S1574-0064\(00\)80168-7](https://doi.org/10.1016/S1574-0064(00)80168-7).
- Meer, Jonathan, and Harvey S. Rosen. 2004. "Insurance and the Utilization of Medical Services." *Social Science and Medicine* 58 (9): 1623–32. [https://doi.org/10.1016/S0277-9536\(03\)00394-0](https://doi.org/10.1016/S0277-9536(03)00394-0).
- Mortensen, Karoline, Chad Perman, and Jie Chen. 2014. "Innovative Payment Mechanisms in Maryland Hospitals: An Empirical Analysis of Readmissions under Total Patient Revenue." *Healthcare* 2 (3): 177–83. <https://doi.org/10.1016/j.hjdsi.2014.03.002>.
- Murray, Robert. 2009. "Setting Hospital Rates to Control Costs and Boost Quality: The Maryland Experience." *Health Affairs* 28 (5): 1395–1405. <https://doi.org/10.1377/hlthaff.28.5.1395>.
- National Center for Health Statistics. 2012. "Birth Edit Specifications for the 2003 Proposed Revision of the U.S. Standard Certificate of Birth." https://www.cdc.gov/nchs/data/dvs/birth_edit_specifications.pdf.
- New York State Department of Health. 2012. "SPARCS Appendix I- Revenue Codes." 2012. <https://www.health.ny.gov/statistics/sparcs/sysdoc/appi.htm>.
- Papanicolas, Irene, Liana R. Woskie, and Ashish K. Jha. 2018. "Health Care Spending in the United States and Other High-Income Countries." *JAMA*. <https://doi.org/10.1001/jama.2018.1150>.
- Park, James D, Edward Kim, and Rachel M Werner. 2015. "Inpatient Hospital Charge Variability of U.S. Hospitals." *Journal of General Internal Medicine* 30 (11): 1627–32. <https://doi.org/10.1007/s11606-015-3352-0>.
- Patel, Ankit, Rahul Rajkumar, John M. Colmers, Donna Kinzer, Patrick H. Conway, and Joshua M. Sharfstein. 2015. "Maryland's Global Hospital Budgets — Preliminary Results from an All-Payer Model." *New England Journal of Medicine* 373 (20): 1899–1901. <https://doi.org/10.1056/NEJMp1508037>.
- Phibbs, Ciaran S., Susan K Schmitt, Laurence C Baker, Aaron B Caughey, Roderic H Phibbs, and Beate Danielsen. 2007. "Level and Volume of Neonatal Intensive Care and Mortality in Very-Low-Birth-Weight Infants." *New England Journal of Medicine* 356 (21): 2165–75. <https://doi.org/10.1056/NEJMsa065029>.
- Pines, Jesse M., Sonal Vats, Mark S. Zocchi, and Bernard Black. 2019. "Maryland's Experiment With Capitated Payments For Rural Hospitals: Large Reductions In Hospital-Based Care." *Health Affairs* 38 (4): 594–603. <https://doi.org/10.1377/hlthaff.2018.05366>.
- Pope, Chris. 2019. "When Government Sets Hospital Prices: Maryland's Experience | Manhattan Institute."
- Riley, Gerald F. 2009. "Administrative and Claims Records as Sources of Health Care Cost Data." *Medical Care*. <https://doi.org/10.1097/mlr.0b013e31819c95aa>.

- Roberts, Eric T., Laura A. Hatfield, J. Michael McWilliams, Michael E. Chernew, Nicolae Done, Sule Gerovich, Lauren Gilstrap, and Ateev Mehrotra. 2018. "Changes in Hospital Utilization Three Years into Maryland's Global Budget Program for Rural Hospitals." *Health Affairs* 37 (4): 644–53. <https://doi.org/10.1377/hlthaff.2018.0112>.
- Roberts, Eric T., J. Michael McWilliams, Laura A. Hatfield, Sule Gerovich, Michael E. Chernew, Lauren G. Gilstrap, and Ateev Mehrotra. 2018. "Changes in Health Care Use Associated With the Introduction of Hospital Global Budgets in Maryland." *JAMA Internal Medicine* 178 (2): 260–68. <https://doi.org/10.1001/jamainternmed.2017.7455>.
- Roemer, Milton Irwin. 1961. "Bed Supply and Hospital Utilization: A Natural Experiment." *Undefined*.
- Russell, Rebecca B, Nancy S Green, Claudia A Steiner, Susan Meikle, Jennifer L Howse, Karalee Poschman, Todd Dias, et al. 2007. "Cost of Hospitalization for Preterm and Low Birth Weight Infants in the United States." <https://doi.org/10.1542/peds.2006-2386>.
- Salemi, Jason L., Meg M. Comins, Kristen Chandler, Mulubrhan F. Mogos, and Hamisu M. Salihu. 2013. "A Practical Approach for Calculating Reliable Cost Estimates from Observational Data: Application to Cost Analyses in Maternal and Child Health." *Applied Health Economics and Health Policy* 11 (4): 343–57. <https://doi.org/10.1007/s40258-013-0040-2>.
- Schroeder, Steven A, and William Frist. 2013. "Phasing Out Fee-for-Service Payment." *N Engl J Med*. Vol. 21. <https://www.nejm.org/doi/pdf/10.1056/NEJMs1302322>.
- Schulman, Joseph, David Braun, Henry C. Lee, Jochen Profit, Grace Duenas, Mihoko V. Bennett, Robert J. Dimand, Maria Jocson, and Jeffrey B. Gould. 2018. "Association between Neonatal Intensive Care Unit Admission Rates and Illness Acuity." *JAMA Pediatrics* 172 (1): 17–23. <https://doi.org/10.1001/jamapediatrics.2017.3913>.
- Schulman, Marjorie. 2003. "Neonatology and Emerging Trends in Health Insurance." *American Journal of Perinatology* 20 (8): 433–39. <https://doi.org/10.1055/s-2003-45386>.
- Schwartz, R M, R Kellogg, and J H Muri. 2000. "Specialty Newborn Care: Trends and Issues." *J Perinatol* 20 (8 Pt 1): 520–29. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&doctype=Citation&list_uids=11190593.
- Sharfstein, Joshua M., Donna Kinzer, and John M. Colmers. 2015. "An Update on Maryland's All-Payer Approach to Reforming the Delivery of Health Care." *JAMA Internal Medicine* 175 (7): 1083. <https://doi.org/10.1001/jamainternmed.2015.1616>.
- State Health Compare. 2019. "SHADAC Analysis of Prescription Opioid Painkiller Sales in Kilograms per 100,000 People by Oxycodone and Hydrocodone | State

- Health Access Data Assistance Center.” SHADAC, University of Minnesota. 2019. <http://statehealthcompare.shadac.org/table/170/prescription-opioid-painkiller-sales-in-kilograms-per-100000-people-by-oxycodone-and-hydrocodone#1/87,88/25/204>.
- WHO. 2007. “Neonatal and Perinatal Mortality: Country, Regional and Global Estimates.” <https://doi.org/10.2307/41110883>.
- Wing, Coady, Kosali Simon, and Ricardo A Bello-Gomez. 2018. “Designing Difference in Difference Studies: Best Practices for Public Health Policy Research Keywords.” *Annu. Rev. Public Health* 39: 453–69. <https://doi.org/10.1146/annurev-publhealth>.
- Wolfers, Justin. 2006. “Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results.” *American Economic Review*. <https://doi.org/10.1257/aer.96.5.1802>.
- Woolhandler, Steffie, and David U. Himmelstein. 2019. “Single-Payer Reform - ‘Medicare for All.’” *JAMA - Journal of the American Medical Association* 321 (24): 2399–2400. <https://doi.org/10.1001/jama.2019.7031>.
- Xu, Jiaquan, Sherry L Murphy, Kenneth D Kochanek, Brigham Bastian, and Elizabeth Arias. 2018. “National Vital Statistics Reports Volume 67, Number 5 July 26, 2018, Deaths: Final Data for 2016.” *National Vital Statistics Reports*. Vol. 67. <https://www.cdc.gov/>.