

States in the 1990s. To establish causal estimates, I propose a new identification strategy that combines the synthetic control method and the instrumental variable estimator. I then put the estimated effects in a spatial general equilibrium model and calculate the welfare impacts on different agents of the local economy. I find that military personnel contractions significantly reduced local employment levels, but as people migrate, the incidence of welfare impacts was mainly on landowners, not on workers.

Chapter 2: The Effect of Aid on Growth: Evidence from a Quasi-Experiment
(with Sebastian Galiani, Steve Knack, and L. Colin Xu)

Whether foreign aid promotes economic growth in recipient countries is one of the most important yet most debated questions in the study of economic growth. The second chapter studies the causal effect of foreign aid on economic growth by exploiting the large discontinuous reduction in aid that occurs as a country passes an exogenously-given income threshold. We find a positive and sizable causal effect of foreign aid: a one percentage point increase in the aid-to-GNI ratio raises annual economic growth by 0.35 percentage point.

Chapter 3: Industrializing from Scratch: the Persistent Effects of China's
"Third Front Movement" (with Jingting Fan)

The third chapter studies the effect of a large-scale industrialization effort in China known as the "Third Front Movement" on long-run development of regional economies. The Movement provides a unique policy experiment to study the important question of whether temporary government subsidies in the nascent industrial sector can permanently push a rural economy into a new development path. We find that decades after the Movement ended, industrialization and urbanization levels remained much higher in local economies that received large subsidies from the Movement, and the effects are mainly driven by the fast-growing non-state sector.

ESSAYS ON ECONOMIC SPILLOVERS, LABOR MARKETS, AND
ECONOMIC DEVELOPMENT

by

Ben Zou

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
2015

Advisory Committee:

Professor Judith K. Hellerstein, Co-chair

Professor Sebastian Galiani, Co-chair

Professor Ethan Kaplan

Professor Ginger Z. Jin

Professor Margaret M. Pearson

© Copyright by
Ben Zou
2015

Dedication

To my parents

Acknowledgments

I am deeply indebted to my advisors. Judy Hellerstein provided outstanding mentoring at all stages of my PhD. I am deeply grateful to her for bringing me into the profession. Sebastian Galiani provided invaluable advice on my research and brought me in to work with him on multiple projects. The second chapter is coauthored with him. Ethan Kaplan guided me through different stages of my dissertation through timely feedback, encouragement, and many good discussions. Ginger Jin provided invaluable advice and support on all my research. I would like to thank Margaret Pearson for serving on my dissertation committee. I benefited enormously from discussions with Ron Chan, John Chao, Daisy Dai, Jingting Fan, Jessica Goldberg, Raymond Guiteras, John Ham, Soohyung Lee, Peter Murrell, John Shea, Lixin Tang, Lesley Turner, Sergio Urzua, John Wallis, Shulin Wee, and Haishan Yuan. I would also like to thank participants at various seminars, particularly at the University of Maryland, for helpful suggestions. I thank the University of Maryland Graduate School and the U.S. Department of Labor for financial support for part of this dissertation.

Contents

1	Introduction	1
2	CHAPTER 1: THE LOCAL ECONOMIC IMPACTS OF MILITARY PERSONNEL CONTRACTIONS	6
2.1	Introduction	6
2.2	Background	11
2.2.1	Military Personnel and Local Economies	11
2.2.2	Post-Reagan Military Contractions	12
2.3	A Spatial Equilibrium Model	14
2.3.1	Overview	14
2.3.2	Household Problem	16
2.3.3	Housing Market	17
2.3.4	Local Businesses	18
2.3.5	Equilibrium Conditions	21
2.3.6	Welfare Impacts	22
2.3.7	Implication for Empirical Study	24
2.4	Econometric Approach	24
2.4.1	Estimation Model	24
2.4.2	A Generalized Synthetic Control Approach	27
2.5	Data and Sample	32
2.5.1	Data	32
2.5.2	Sample	33
2.6	Empirical Results	33
2.6.1	Comparing Specifications	33
2.6.2	Local Labor and Housing Markets	35
2.6.3	Sectoral Composition	38
2.6.4	Welfare Impacts	39
2.7	Adjustment of Local Economies	40
2.7.1	Estimating Dynamic Effects	40
2.7.2	Mobility	43

2.8	Conclusions	45
2.9	Appendices	48
2.9.1	Appendix A: Sources and Construction of Variables	48
2.9.2	Appendix B: Details of the Model	49
2.9.3	Appendix C: Two-Step Estimation of Panel Model with En- dogenous Treatment	54
2.9.4	Appendix D: Statistical Inferences for the Two-Step Estimation	59
2.9.5	Appendix E: Robustness and Heterogeneous Effects	61
3	CHAPTER 2: THE EFFECT OF AID ON GROWTH: EVIDENCE FROM	
	A QUASI-EXPERIMENT	85
3.1	Introduction	85
3.2	Previous Aid-Growth Studies	88
3.3	Data and Sample	89
3.4	The IDA Threshold and Foreign Aid	90
3.5	Foreign Aid and Economic Growth	94
3.5.1	Econometric Models	94
3.5.2	Baseline Results	98
3.5.3	Measurement Error in Aid	102
3.5.4	Bunching	103
3.6	Robustness Checks	104
3.6.1	Omitted Variables	104
3.6.2	Timing of Crossing	105
3.6.3	Other Robustness Checks	106
3.7	External Validity	107
3.8	Conclusions	108
4	CHAPTER 3: INDUSTRIALIZATION FROM SCRATCH: THE PERSIS-	
	TENT EFFECTS OF CHINA’S “THIRD FRONT” MOVEMENT	128
4.1	Introduction	128
4.2	Background	134
4.3	Data and Sample	137
4.3.1	Data Sources	137

4.3.2	Sample and variables	138
4.3.3	Summary Statistics	140
4.4	A Case Study of Shiyan	141
4.4.1	A Natural Comparison	142
4.4.2	A Synthetic Comparison	143
4.4.3	Comparing Shiyan and Xiangxi with Synthetic Shiyan	145
4.5	Long-Run Effects of Temporary Subsidies to Industrialization	146
4.5.1	Baseline Specification	146
4.5.2	Local Economic Outcomes in the Post-Reform Years	148
4.6	Mechanisms	152
4.7	Conclusions	157
4.8	Appendices	158
4.8.1	Appendix A. Sources of Data	158
4.8.2	Appendix B. Details of Variable Construction	160

List of Tables

2.1	Summary Statistics	70
2.2	Military Personnel Contractions and Civilian Employment, 1988-2000	71
2.3	Military Contractions and County Outcomes, 1988-2000	72
2.4	Changes in Civilian Employment between 1988 and 2000 by Sector .	73
2.5	Dynamic Effects of Military Personnel Contractions	74
2.6	Military Personnel Contractions and Procurement Contractions . . .	75
2.7	Simulation Results	76
2.8	Bias and Rejection Rate by Estimation Model	77
2.9	Variance and Correlation Matrix from Bootstrapping	78
2.10	Panel Regressions with County Fixed Effects, 1980-2000	79
2.11	Dynamic Effect of Military Personnel Contractions- Dynamic panel with county fixed effects	80
2.12	Influential Counties	81
2.13	Commuting Zones	82
2.14	Heterogeneous Effects by the Size of the Cut	83
2.15	Heterogeneous Effects by Population Density	84
3.1	Sample Countries and Years of Crossing the IDA Threshold	112
3.2	Summary Statistics	113
3.3	IDA Threshold and Aid	114
3.4	IDA Threshold and Aid - Placebo Threshold at 50% of the True Level	115
3.5	Baseline Results	116
3.6	Alternative Specifications	117
3.7	Narrowing Periods	118
3.8	Adding Covariates	119
3.9	Model Robustness Checks	120
3.10	External Validity	121
3.11	IDA Threshold and Aid with Quadratic Log Initial Income Levels	122
3.12	First Stage - Policies as Outcome Variables	123
3.13	First Stage - Policies as Outcome Variables	124
3.14	Functional Form of Aid	125
3.15	Effects of Aid on Investment	126
3.16	Construction and Sources of Key Variables	127

4.1	Characteristics of Firms by Year of Opening	167
4.2	Shiyan and Its Comparisons	168
4.3	Post-Reform Industrial Employment in the Non-State Sector (2004) .	169
4.4	Prefecture-Level Outcomes in 2000	170
4.5	Initial Conditions versus Continuing Spillovers	171
4.6	Mechanism - Conditional on Early 1980s Outcomes	172
4.7	Within-Sector and Input-Output Spillovers	173
4.8	Migration and Urbanization	174
4.9	Summary Statistics	175
4.10	Shiyan and Its Comparisons (More Results)	176
4.11	Prefecture-Level Outcomes in 2000 - Robustness	177
4.12	Data Sources	178

List of Figures

2.1	Military Personnel Contractions between 1988 and 2000	66
2.2	Illustration of Workers' Welfare	67
2.3	San Diego and Synthetic San Diego	68
2.4	Cumulative Effects by Grouping Years	69
3.1	IDA Threshold 1987-2010	110
3.2	Histogram of Income and McCrary Test of Bunching	111
4.1	Third Front Region	162
4.2	Intensity of Investment by Region	163
4.3	Effects over Time - Shiyan and Its Comparisons	164
4.4	Varying Cutoff Urban Rate in 1964	165
4.5	Effects over Time	166

1 Introduction

Economic conditions vary greatly across locations. In December, 2014, the unemployment rate in Gary, Indiana was 10.5%. Thirty miles away, the unemployment rate in Chicago was 6.2%. In 2011, the state of West Virginia had the second lowest median household income in the United States, while its neighbor Maryland had a median household income that was about 80% higher, and was the highest in the country. In developing countries, regional economic disparities are even greater. The richest province in China has a per capita GDP that is more than 3 times that of the poorest province.

Many factors affect economic differences across space, from locational fundamental characteristics such as geography, to transitory factors such as weather shocks. Due to these factors, it is unrealistic that the local economies will be exactly the same across space, nor should spatial economic differences necessarily be a concern. As long as people can move to different places at no cost, spatial economic disparity as measured by income or population simply reflects the heterogeneous choices of people: a farmer chooses to live in an agricultural state over an industrial center; a person who has an idiosyncratic preference for variety and a small disutility from congestion chooses to live in crowded cities.

However, people face various restrictions to arbitrage geographic economic differences. Migration is costly; information on different locations is often limited. In fact, controlling for demographic characteristics, residents in economically depressed regions still work fewer hours and earn less than their counterparts in locations with better economic conditions. A local economy may also fail to reach its production potential for various reasons. For example, in some developing countries, some (local) economies may be stuck in bad equilibria due to lack of basic infrastructure (Rosenstein-Rodan, 1943).

The idea that market forces may not generate desirable local economic out-

comes have motivated various location-based government policies. By targeting residents and businesses within certain geographic areas, these policies aim at promoting local economies and benefiting their residents (Gottlieb and Glaeser, 2008). These policies, often in the form of tax credits for local businesses, employee subsidies, or infrastructure investment, are widespread across the world. In the United States, it is estimated that in 2011, state and local governments spent more than 80 billion dollars on tax incentives alone in order to attract or retain local businesses (Story et al., 2012). Chapter 1 of this dissertation studies the local economic impacts of the removal of employers from the local economy. Internationally, bilateral and multilateral donors donate 140 billion dollars each year in the form of development aid in the hope of pulling developing countries out of income traps and promoting economic growth (Chapter 2). Throughout history, governments launch large-scale comprehensive development policies aiming at fundamentally changing regional economic landscape. Examples of these policies include the Tennessee Valley Authority in the United States, China’s massive industrialization campaign in its hinterland (Chapter 3), and the development of Northeast Brazil.

The efficacy and efficiency of these policies depend on the objectives of these policies, and how agents respond to incentives. Thus the evaluation of these policies may well be different from case to case. This dissertation consists of three studies of different government interventions in local economies, at the county, region, and country level. All three chapters estimate causal impacts of the policies on local economic conditions and discuss the mechanisms behind the impacts. . The first chapter, titled “The Local Economic Impacts of Military Personnel Contractions”, studies the impacts on various aspects of the county economies of military personnel contractions that took place in the United States in the 1990s, and the welfare impacts on local economic stakeholders such as workers, landowners, and business owners. The second chapter, titled “The Effect of Aid on Growth: Evidence from

a Quasi-Experiment”, estimates the causal effect of international aid on economic growth of recipient developing countries by proposing and estimating a novel instrumental variable. The third chapter, titled “Industrializing from Scratch: the Persistent Effects of China’s “Third Front Movement”, studies a large-scale industrial buildup in China and its impact on the long-run regional economic development.

The first chapter focuses on a period between 1988 and 2000, during which the number of military personnel stationed in the United States shrank by 30 percent, and hundreds of local economies with military bases were affected. As military personnel pull out of a locale, demands for local labor, housing, and locally-traded goods decline. In this chapter, I jointly estimate the impacts of this personnel reduction on the equilibrium quantities and prices of local labor markets, local housing markets, and local product markets. In order to establish causal identification, this chapter proposes and estimates a two-step empirical model combining synthetic control and instrumental variables methods. I find sizable effects of military personnel contractions on civilian employment levels and the numbers of private businesses. But this employment reduction translates into out-migration of local civilian residents to other jurisdictions, resulting in a drop in rental prices that is larger than the drop in wages. Then I build a simple spatial equilibrium model that describes the equilibrium conditions and derives expressions for welfare calculation. Relating them to my econometric estimates, I show that the welfare cost on workers is small while that on landowners is sizable.

Whether foreign aid promotes economic growth in recipient countries is one of the most important and most debated questions in the study of economic growth. However, the literature on aid and growth has not found a convincing method for identifying the causal effects of aid. The second chapter of this dissertation secures a causal estimate by using a new instrumental variable. The instrumental variable is based on the fact that since 1987, eligibility for aid from the International Develop-

ment Association (IDA) has been based partly on whether or not a country is below a certain threshold of per capita income. We find evidence that other donors tend to reinforce rather than compensate for reductions in IDA aid following threshold crossings. Overall, aid as a share of gross national income (GNI) drops about 59 percent on average after countries cross the threshold. Focusing on the 35 countries that have crossed the income threshold from below between 1987 and 2010, this chapter finds a positive, statistically significant, and economically sizable effect of aid on growth. A one percentage point increase in the aid to GNI ratio from the sample mean raises annual real per capita growth in gross domestic product by approximately 0.35 percentage points. We further show that the main channel through which aid promotes growth is by increasing physical investment.

In poor economies with low levels of industrialization, can temporary government subsidies on the industrial sector promote long-run economic development? The evaluations of government subsidies on industries are often confounded by the fact that subsidies often go to sectors or locations that are selected based on particular expectations of economic returns to the subsidies. The third chapter of this dissertation sheds light on this important yet difficult question by studying a large-scale industrialization campaign in China known as the “Third Front Movement”. The Movement, initiated in the mid-1960s and ended in the late 1970s, involved establishing large-scale industrial plants in the remote and mountainous regions in western China. It was motivated by national defense considerations, and these industrial plants were scattered in places with small existing industrial presences. The Movement provides a unique policy experiment to study the important question of whether temporary government subsidies can permanently push a local economy into a new equilibrium development level. We find that decades after the Movement ended, industrialization and urbanization levels remained much higher in local economies that received large subsidies from the Movement. The effects are

stronger when the subsidy went to advanced manufacturing industries.

These three essays focus on studying government policies aimed at local economies. They develop econometric approaches that are suitable for evaluating each of these policies, and use conceptual frameworks that take into account the rational responses of economic agents to these policy interventions. In combination, this dissertation contributes to our current understanding of local economies and location-based policies.

2 CHAPTER 1: THE LOCAL ECONOMIC IMPACTS OF MILITARY PERSONNEL CONTRACTIONS

2.1 Introduction

Economic activities in a local geographic area interact with each other: shocks to one sector have direct and indirect effects on other agents in the local economy. Economists have long been interested in identifying the size of local economic spillover effects and the mechanisms through which these effects work.¹ The findings of studies on spillover effects have important policy implications as governments at all levels across the world spend billions of dollars on various place-based policies aiming at promoting economic opportunities in certain localities.

In order to fully capture the total impacts on the various stakeholders of the local economy, it is important to simultaneously model and estimate the behavior and interactions of many agents. Most existing studies focusing on local labor market outcomes, usually in terms of the number of jobs created or lost, are likely to miss important parts of the story.² Fewer job opportunities depress employment and wages but also make housing and local services more affordable. Restricting attention to the local labor market overlooks this compensating effect. Moreover, the effects of decreased job opportunities on workers' levels of economic activity may also differ from the effects on welfare: if displaced workers are freely mobile and can easily find jobs elsewhere with competitive wages, their welfare losses are likely to be small. Finally, local shocks also affect the welfare of landowners and the profitability of local firms. The welfare incidence matters for policy. A progressive government, for example, may care more about the welfare of the workers than that of the landowners.

This chapter studies the local economic impacts of the contractions of a special industry – the US military – on workers, landowners and firms. I focus on the post-

¹Adam Smith (1776) notes that local economic links help workers and firms to specialize. Marshall (1890) remarks that co-location of related industries promote productivity through economic linkages. Recent studies on local economic spillover effects date back to the economic downturn in the early 1980s when there were large demand shocks to the manufacturing industry in the United States (Topel, 1986; Blanchard and Katz, 1992; Bound and Holzer, 2000).

²Many existing studies estimate a local job multiplier, the additional jobs created or lost due to exogenous changes in employment in one sector (e.g., Bartik, 1991; Blanchard and Katz, 1992; Black et al., 2005; Moretti, 2010b).

Reagan military personnel cuts between 1988 and 2000, a period during which the size of the US military shrank by 30 percent.³ The bulk of the cuts took place in locations hosting major military bases. Declining military presence reduces demand for housing and local non-tradable goods, which drives down local housing prices and the price of the non-tradable goods. Some non-tradable sector firms go out of business. Meanwhile, an over-supply of workers relative to job opportunities drives down local wages, and some households choose to leave the local economy in favor of better economic opportunities. The relative strengths and net effects of these intertwined channels depend on the parameters of the local economy and are the interests of this empirical study.

There are two challenges to establishing the causal effect: the problems of omitted variables and simultaneity. The omitted variables problem means that there might be location-specific secular trends driven by local fundamentals that may confound with military contractions. For example, if large military bases are frequently located in places with faster economic growth, the effects of military contractions will be under-estimated if the underlying economic trajectories are not taken into consideration. Simultaneity would occur if, in deciding the size of military contraction in each place, the Department of Defense (DoD) avoided large cuts in places with unfavorable transitory shocks in order to minimize negative economic impacts. Failing to correct for this simultaneity problem would also make the OLS estimate biased downward.

I propose a shift-share instrumental variable to address the simultaneity problem. This type of instrument, also known as a Bartik instrument (Bartik, 1991), is widely used in the related literature. The instrumental variable is the interaction between the location's historical military presence and the contemporaneous nationwide contraction in military personnel. The instrument solves the simultaneity problem as long as the national change is not driven by local idiosyncratic shocks and historical military presence is uncorrelated with the transitory shocks. However, this instrument does not solve the omitted variable problem because the historical military presence and the secular trend are both functions of local fundamentals: places with some unobserved advantages are likely to attract large military bases

³In a typical year without major wars, the Department of Defense (DoD) spends about 40 percent of its budget on compensation for military personnel, 30 percent on procurement, and the rest on operational costs. During this period, all these components declined substantially. Section 2.7.2 uses variations in procurement as a robustness check.

and have better trajectories of economic development. Therefore, the identification of the causal effect using the shift-share instrument hinges on partialling out the secular trend.

I use two different approaches to partial out the secular trend. In the first approach, I include a host of pre-determined county characteristics as covariates. This conventional approach is easy to implement but imposes restrictive parametric assumptions and has the risk of mis-specifying the secular trend. In the second approach, I use a non-parametric method similar to the synthetic control approach used in Alberto Abadie and Hainmueller (2009). From a large pool of counties without military bases that serve as potential comparisons, I construct for each county with military bases a “synthetic control” that best matches the outcome trajectories in the pre-treatment period. I use the synthetic control’s post-treatment trajectories of the outcomes as the counterfactuals for counties with military bases. This approach allows me to account for flexible secular trends without knowing their exact functional forms.

The identification can be thought of as having two steps. The first involves partialling out the secular trends, and the second estimates the partial equation using the Bartik instrument. Potential mis-specification in the parametric approach in the first step may still invalidate the Bartik instrument. I show in a simulation exercise how mis-specified parametric models could lead to substantially biased estimates, while the synthetic control approach obtains estimates close to the true value.

I find sizable effects of military personnel contractions on the levels of local economic activities. I find that, over the 12-year period, cutting one military worker results in the loss of 1.2 civilian jobs and 32,000 dollars in civilian earnings. The effects are concentrated in the non-tradable sectors while the tradable sectors are barely affected, a result often found in the related literature (e.g., Black et al., 2005). I find a large migration response: on average 2.4 civilians leave the local economy for every civilian job loss. The high migration response results in small impacts on local wages but big impacts on rental prices. The average size of the impact in the sample is equal to a reduction in the ratio of military personnel to the total local population by one percentage point. This impact reduces local wages by 0.47 percent, or 141 dollars annually, and barely affects the civilian employment to population ratio. In contrast, local rental prices drop by 1.3 percent.

To understand the reduced-form results and calculate welfare impacts on the stakeholders of the local economy, I incorporate a two-sector model in a simple spatial equilibrium framework with workers, landowners, firms, and their interactions. In the model, each location is a small open economy, and capital and labor mobility are costless. Workers have heterogeneous preferences over locations so that some do not move when real wages decline in the local economies where they reside. There are two sectors in the local economy, producing tradable and non-tradable goods, respectively, and workers are interchangeable between the two sectors. The housing market is competitive. The model endogenously determines the local economy's population, employment, wages, housing rental prices, and prices of local non-tradable goods. Military personnel contractions reduce demand for local non-tradable goods and housing. Intuitively, businesses and workers in the non-tradable sectors are directly affected. Displaced workers can work in the tradable sector with lower wages, or migrate out of the local economy for a more competitive wage. For a particular shock, the welfare of workers depends on the flexibility of their location choices and the technology used in the housing, tradable and non-tradable sectors. The welfare of the landowners, on the other hand, depends on changes in rental prices and local population.

To calculate welfare impacts, I first derive expressions for changes in welfare for workers, landowners, and profits for firms as functions of military personnel contractions. I then relate these expressions to the reduced-form estimates. The welfare loss for workers is negligible: as a result of reducing the military personnel to population ratio by one percentage point, workers' utility drops by a mere 0.02 percent. In fact, the utility of the tradable firms increases by 0.3 percent thanks to declining local wages. Landowners bear most of the welfare loss, as rental revenue drops by 3.6%. This result is not surprising since the large migration response suggests that the local labor supply is elastic, but it has different policy implications than more conventional approach to assess welfare which only measures the local economic impact via the change in the number of jobs.

Finally, I investigate how quickly local economies adjust to shocks. Using a year-by-year panel estimation, I find that local economies respond to military personnel contractions quickly. The accumulated effects on employment, earnings, and businesses in the first two years after a particular military cut are close to the magnitude of the accumulated effects over the 12-year period.

This chapter has relevant policy implications. The DoD is the largest employer among all federal agencies. After a decade of expansions since 9/11, the US military is again making substantial budget cuts. In general, the non-welfare part of federal government spending is projected to decline (Congressional Budget Office, 2014). Many local communities that rely heavily on federal spending are concerned that their local economies may be deeply affected. The findings in this chapter suggest that workers and local economies exhibited substantial resilience during similar shocks in the 1990s. The size of local economic activity might be smaller as demand shrinks, but a smaller economy does not necessarily mean a worse one. On the other hand, landowners are likely to be most negatively affected, as land cannot move and housing stocks are slow to adjust.⁴

This chapter contributes to the literature on local economic dynamics due to exogenous local shocks. Most existing studies estimate a “local job multiplier” based on a partial equilibrium framework. Black et al. (2005) find that losing one job in the coal mining industry causes the loss of 0.35 jobs in the service sector in the same county. Moretti (2010b) finds that an additional job in the tradable sector creates 1.6 jobs in the non-tradable sector in the same metropolitan area. Other papers look at local employment impacts from other sources, such as Chodorow-Reich et al. (2012); Wilson (2012); Serrato and Wingender (2014); Shoag (2012) on the effects of government expenditure, and Autor et al. (2013) on the effects of competition from China. The post-Reagan military personnel contractions were among the largest declines in employment in a single industry in the US history and arguably the largest cuts of government employment. The magnitude of the employment impact found in this chapter falls within the range of existing studies.

A partial equilibrium framework may lead to incomplete understanding of local economic impacts. To obtain a fuller picture, one needs to model all the relevant local economic agents and their interactions (Bartik, 1991; Gottlieb and Glaeser, 2008). New developments in spatial equilibrium models extend the Rosen-Roback framework (Rosen, 1979; Roback, 1982,9) and allow for welfare analysis on different agents (Moretti, 2010a; Kline and Moretti, 2014b). Using this type of models, recent papers show that the welfare implications differ from those based on partial equilibrium frameworks. Taking into consideration house price differences

⁴There is a recent strand of thought that recommends that depressed local economies “shrink to greatness.” For example, one recipe for Detroit is for the city to restore functionality an area suitable for the volume of its current economic activity (Glaeser, 2010).

across metropolitan areas, Moretti (2013) shows that the real wage difference between skilled and unskilled workers is smaller than would be indicated by using nominal wages. Glaeser and Gyourko (2005) argue that the durability of housing during local economic downturns provides a natural safety net for residents who choose to stay. Notowidigdo (2013) points out that government benefit programs that kick in during times of local economic distress have a similar role. Diamond (2013), on the other hand, argues that endogenous amenities lead to welfare impacts larger than what nominal wages imply. This chapter adopts a similar framework to study the case of military personnel contractions, where reductions in military presence directly hit the non-tradable sector and the housing sector, and the effects spill over to other sectors via the shared local labor market and people's migration decisions.

Variation in military spending has been used in the macroeconomics literature for estimating the fiscal multiplier at the national or subnational level (e.g., Nakamura and Steinsson, 2014; Barro and Redlick, 2011; Ramey and Shapiro, 1998). Surprisingly, there are few studies on the impacts of military spending on local economies. Guthrie (1995); Hooker and Knetter (2007); Hultquist and Petras (2012); aus dem Moore and Spitz-Oener (2012) are among the few that study local impacts of military personnel contractions. These papers focus on local labor market outcomes and, in general, do not address potential endogeneity concerns.

The rest of the chapter proceeds as follows: Section 2.2 introduces the setting and historical background. Section 2.3 presents the model. Section 2.4 describes the empirical approach. Section 2.5 describes the data and the sample. Section 2.6 presents the empirical results, relates the model and the empirical results, and calculates welfare impacts. Section 2.7 investigates how fast local economies adjust to new equilibria. Section 2.8 summarizes the findings.

2.2 Background

2.2.1 Military Personnel and Local Economies

The vast majority of military personnel live in or around military bases. In 1987, 353 major military bases in the US were located in 381 counties. The personnel located in these counties accounted for over 90 percent of total US military personnel. These counties are spread across the United States, from large coastal metropolises to sparsely populated deserts. This geographic dispersion is partly due to national

defense considerations: the military should be somewhat evenly distributed across the territory such that it can quickly react to threats to national security. Economic and other practical considerations also affect location choices for military bases. For example, Naval bases must be close to large bodies of water, while testing and training areas, due to their demand for large pieces of land, are often located in places with low population density and low land values.

On average, counties with military bases (henceforth called “base counties”) are larger and more densely populated than counties without military bases (“non-base counties”). In 1980, base counties accounted for about 12 percent of the total county-level jurisdictions in the United States, but contained about 50 percent of the country’s population. Their economic performance was also better: private sector employment in base counties grew by 22 percent between 1980 and 1987, almost twice as much as it did in non-base counties. However, the differences in averages mask important heterogeneity within each group. As the military bases are located across the country, the local economies where they are located have many different types of geographic and economic characteristics. It is therefore a potential econometric concern that cuts in military personnel may be correlated with the underlying characteristics of the local economy.

The presence of military bases generates three types of direct impact on the local economy. First, local military bases employ civilian workers, many of whom work on the bases. In 1990, the DoD employed almost one million civilian workers, making it the largest employer of civilian workers among federal government agencies.⁵ Second, the military bases create demand for goods and services from local contractors. Third, military personnel create demand for housing and local non-tradable goods. Military contractions directly reduce the demand for local labor, non-tradable goods, and housing, which in turn affects other parts of the local economy.⁶

2.2.2 Post-Reagan Military Contractions

The size of the military as measured by the number of men and women in uniform in the United States has been declining since the end of WWII. This trend paused in

⁵The 2012 Statistical Abstract. Table 499. url: <http://www.census.gov/compendia/statab/2012/tables/12s0499.pdf>. Last accessed in Oct, 2014.

⁶I ignore the effects on local labor supply of new veterans. I find that military personnel contractions in the sample period do not increase the number of veterans in the county.

the 1980s when the Reagan Administration significantly expanded military spending while the number of military personnel stayed stable.⁷ Political gridlock in deciding which military bases to cut contributed to the pause, as members of Congress and local politicians viewed letting military bases be slashed in their jurisdictions under their watch as political suicide.

By the end of the Reagan Administration, it became clear that military contractions were necessary. The international geopolitical situation changed dramatically in the late 1980s as the winding down of the Cold War prompted a reduction in military capacity in the United States. Domestically, the federal fiscal situation after large tax cuts and spending hikes throughout the Reagan Administration also made military cuts necessary. To circumvent political gridlock, in 1988, the Base Re-Alignment and Closure (BRAC) Act was passed. An independent commission with members jointly nominated by the president and Congress became responsible for selecting military installations to be re-aligned or permanently closed, based on a list of the DoD's recommendations. The BRAC insulated individual members of Congress from the political penalties of potential shutdowns of military bases in their jurisdictions. Military values and cost-saving were the main criteria for choosing military bases to be re-aligned or closed, although potential economic impacts were also an explicit consideration.⁸

The top graph in figure 2.1 shows the trajectory of the number of active duty military personnel between 1975 and 2010. Over two-thirds of the 535 major military bases were affected by the four rounds of BRAC in 1988, 1991, 1993, and 1995. Many bases that were not chosen for closure by the BRAC also experienced substantial declines in personnel as parts of their operations were cut or moved. The declining trend did not stop until the 9/11 terrorist attacks in 2001.⁹ Between 1988 and 2000, the number of military personnel dropped by over 30 percent, discharging about half a million military personnel. It was one of the largest negative shocks generated by a single industry in the past 50 years of US history.¹⁰ I refer to this

⁷The expansion was mainly concentrated in military procurement for developing new weapons, but the number of military personnel also stopped declining during this period.

⁸The military value under consideration includes the current and future mission capabilities, operational readiness; the availability and condition of land, facilities, and associated airspace; the ability to accommodate contingency, mobilization, surge, etc.

⁹Although there was another round of BRAC in 2005, the size of active duty military personnel has been fairly stable at around one million since 2000.

¹⁰A comparable figure is that between 1977 and 1987, the United States shed about 500,000 jobs in the auto industry and 350,000 jobs in the steel industry (Feyrer et al., 2007).

episode of military personnel contractions as “post-Reagan military contractions”.

The post-Reagan military contractions affected counties with military bases since most military personnel were stationed in these counties. As the bottom graph of figure 2.1 shows, between 1990 and 2000, the number of military personnel in base counties dropped from 1.47 million to 1.05 million, a 28.5 percent decline. In these counties, the military personnel to population ratio dropped from 2.61 percent in 1988 to 1.68 percent in 2000. In contrast, the non-base counties saw negligible military personnel contractions, as military personnel dropped slightly from 0.11 percent of the population in 1988 to 0.08 percent of the population in 2000. To estimate the effects of military personnel contractions on local economies, I focus on base counties. The non-base counties, with negligible military presence, serve as potential comparisons.

Although not a priority, the potential impact on local economies was an explicit consideration for the DoD and the BRAC Commission in deciding where to cut military personnel and by how much. During the process of deciding which bases would be closed or realigned, objections from local communities citing economic concerns were numerous, and there were indeed cases in which the BRAC Commission rejected large cuts proposed by the DoD due to local economic concerns.¹¹ In addition, considerations not directly related to local labor market conditions, such as cost-saving considerations and military values, can be confounded with economic outcomes. For example, since good labor market conditions drive up local prices, if other factors remain constant, shrinking the military presence in these areas would lead to sizable savings for the DoD. These explicit and implicit considerations for local economic conditions raise concerns about the endogeneity of military personnel contractions.

2.3 A Spatial Equilibrium Model

2.3.1 Overview

In order to describe the effects of military personnel contractions on local economies, this section presents a spatial equilibrium model of local economies. The model

¹¹For example, the proposed closure of the submarine base in New London, CT, was rejected by the 2005 BRAC Commission partly because “[...] the Commission found the argument of overall economic impact compelling”. (2005 BRAC Commission Report to the President)

serves three purposes. First, it describes the mechanisms through which military personnel contractions affect various stakeholders in the local economy. Second, its equilibrium conditions suggest a vector of outcomes of interest, which I will estimate empirically in sections 2.4 through 2.6. Third, it allows me to derive expressions for welfare changes of the stakeholders in the local economy, which I bring to the data in section 2.6.4.

This model adapts from the framework for local labor market equilibrium in Kline and Moretti (2014b). The model features many locations. Each location is a small open economy in which households, landowners, and firms interact with each other.¹² Capital is elastically supplied nationwide. The key feature of this framework is that the households, besides deriving utility from consumption, have heterogeneous preferences over locations. Shocks to the local economy will be inframarginal for some households with strong preferences for the locality in making migration decisions. These households will therefore bear the full welfare consequences of changes in local wage, price, and rent. This feature distinguishes this kind of model from the standard Rosen-Roback framework (Rosen, 1979; Roback, 1982,9) where workers are perfectly mobile, and migration fully arbitrages differences in local real wages across locations.

Military personnel contractions directly reduce the demand for locally-traded goods, but nationally-traded goods are not directly affected. In order to model this, I extend this framework by introducing a simple two-sector model. Firms in the non-tradable sector, facing a competitive market, produce goods and services that are traded within the local economy. The equilibrium of non-tradable goods is obtained within the local economy with the local price endogeneously determined. Each firm in the tradable sector produces a differentiated good that is sold in the national market. Firms in the tradable sector are indirectly affected by the military personnel contractions through changes in the conditions of the local labor market they share with firms in the non-tradable sector. Introducing a two-sector model in the spatial equilibrium framework thus adds an inter-sectoral wedge in local labor supply besides the wedge generated by migration.¹³

¹²Landowners are assumed to be absent. Throughout the chapter, the term “household” is interchangeable with the term “worker”.

¹³Kovak (2013) includes a local non-tradable sector in the analysis of trade shocks on local economies. But his model does not allow heterogeneity in preference for locations. Yoon (2014) builds a two-sector model with idiosyncratic individual preference over locations in a dynamic model of location choices. Unlike the model introduced in this chapter, the changes in sectoral

2.3.2 Household Problem

Each household i chooses where to live and how much to consume. Households consume tradable and non-tradable goods and housing. Households all have the same productivity and each provides one unit of labor inelastically.¹⁴ For a household that lives in location c , its utility maximizing problem can be written as

$$\begin{aligned} \max_{h_{ic}, X_{ic}^N, X_{ic}^T} u_{ic} &= \ln A_c + \alpha \ln h_{ic} + \beta \ln X_{ic}^N + (1 - \alpha - \beta) \ln X_{ic}^T + e_{ic}, \\ \text{s.t.}, r_c h_{ic} + p_c X_{ic}^N + p_T X_{ic}^T &= w_c \end{aligned}$$

where A_c is the dollar value of amenities in location c , which are freely available to all its residents. h_{ic} , X_{ic}^N , and X_{ic}^T are, respectively, the amount of housing, the amount of non-tradable goods, and the amount of the tradable goods consumed by household i . The tradable good, X^T , is a composite good with many varieties of tradable goods

$$X^T = \left(\int_{j \in J} x_j^{(\sigma^T - 1)/\sigma^T} dj \right)^{\sigma^T / (\sigma^T - 1)},$$

where $j \in J$ is one variety of the partially substitutable tradable goods, and $\sigma^T > 1$ is the elasticity of substitution between any two varieties.¹⁵ p_T is the price of the composite tradable good, which is standardized to 1. Both rental price, r_c , and non-tradable goods price, p_c , are determined by local housing and non-tradable goods markets equilibria and differ across locations. The utility is in the Cobb-Douglas form. α , β , and $1 - \alpha - \beta$ represent the shares of income spent on housing, non-

composition are driven by nationwide factor-biased technological change in the two sectors.

¹⁴Appendix B.1 provides a simple extension of the household problem that allows for unemployment. In that extension, each household independently draws an idiosyncratic utility from leisure (or distaste for work). The labor force participation decision is binary and is made by comparing utility from working and that from not working. The local unemployment rate is thus determined by the marginal worker who is indifferent between working and not. Given the same distribution of idiosyncratic preference for leisure, the local unemployment rate is in turn determined by the local real wage level. Local unemployment rates differ only because local real wage rates differ. In the empirical section, I find that there is no effect on local unemployment rates, which suggests that inelastic supply of labor is not a restrictive assumption here.

¹⁵This constant elasticity of substitution (CES) demand function is widely used in the labor and trade literature. $\sigma^T > 1$ indicates that x_j and $x_{j'}$ are substitutes. See Suarez-Serrato and Zidar (2014) for a recent application in the spatial general equilibrium framework.

tradable goods, and tradable goods, respectively.

e_{ic} is the idiosyncratic utility household i derives from living in location c . e_{ic} is assumed to be *i.i.d* and follows a type I extreme value distribution with dispersion σ^W .¹⁶ Solving the household's problem, household i 's indirect utility from living in location c is

$$v_{ic} = u_c + e_{ic} \quad (2.1)$$

u_c , which is equal to $a_0 + \ln w_c - \alpha \ln r_c - \beta \ln p_c + \ln A_c$ is the deterministic term common to each household that lives in location c , where a_0 is a constant. u_c can be also thought of as a measure of real wages adjusted for local living expenses and amenities.

Each household chooses a location to live in such that v_{ic} is maximized. The population size (and labor supply) in location c can be expressed as¹⁷

$$\ln N_c = \frac{1}{\sigma^W} (a_0 + \ln w_c - \alpha \ln r_c - \beta \ln p_c + \ln A_c) + \frac{1}{\sigma^W} a_C. \quad (2.2)$$

Denote c' as a location other than location c . $a_C = \ln \sum_{c'} \exp(u_{c'}/\sigma^W)$, which is a constant. Local population is determined by the local real wage and the distribution of preference across locations. The inverse of the dispersion of idiosyncratic preference across locations, $1/\sigma^W$, is the elasticity of local labor supply with respect to local real wages u_c . Intuitively, if σ^W is large, households have strong preferences over different locations, and local labor supply is inelastic. After a negative shock, as many households do not migrate out of the local economy, real wages drop and the remaining households bear a large share of the welfare loss. In equilibrium, local real wages can vary substantially across locations. Alternatively, if σ^W is small, households do not have strong preferences across locations, and local labor supply is elastic. After a negative shock, local real wages do not change by much in the new equilibrium and neither does the welfare of the remaining households.

2.3.3 Housing Market

The Cobb-Douglas utility function predicts that each household spends a constant share α of its income on housing; local residents' total spending on housing is $N_c \alpha w_c$.

¹⁶The CDF is $F(e_{ic}) = \exp(-\exp(-(e_{ic})/\sigma^W))$

¹⁷See Appendix B.2 for derivation.

Denote $m_c^H > 1$ as the demand shift due to the military presence.¹⁸ The aggregate housing demand in location c is

$$H_c^D = N_c \alpha w_c m_c^H / r_c. \quad (2.3)$$

The housing sector in location c has productivity κ_c . Housing supply in location c responds positively to local rental price but is restricted by geographic characteristics, land use regulations, and other costs, which are governed by location-specific elasticity η_c ,

$$H_c^S = \kappa_c r_c^{\eta_c}. \quad (2.4)$$

The local housing market equilibrium is obtained by combining equation 2.3 and equation 2.4 and can be expressed in log forms

$$(1 + \eta_c) \ln r_c = \ln N_c + \ln w_c + \ln m_c^H + a_H, \quad (2.5)$$

where $a_H = \ln \alpha - \ln \kappa_c$. Equation 2.5 is intuitive: the local rental price is higher when the local population (N_c) is larger, local wage (w_c) is higher, households spend a larger share of income on housing (α), and when housing is hard to produce either because of natural or policy restrictions (low η_c) or low productivity (low κ_c).

2.3.4 Local Businesses

2.3.4.1 Firms in the Tradable Sector

Each firm in the tradable sector produces one variety of good $j \in J$. Since all j 's are symmetric, without loss of generality, I assume that there is only one firm producing durable goods in location c and it produces one particular variety j . The production function is in the Cobb-Douglas form with constant returns to scale and labor's share equal to h_T . It is written as

¹⁸ m_c^H is modeled as multiplicative to the demand generated by local residents for tractability of the model. To be consistent with the model, in the empirical part, military presence is specified as the number of military personnel as a ratio of population.

$$x_j = B_{Tc}(N_c^T)^{h_T}(K_c^T)^{1-h_T}, \quad (2.6)$$

where B_{Tc} is the total factor productivity.

In a world of differentiated goods, each firm is a monopolistic competitor in the national market. It faces a downward sloping demand curve and earns a positive profit. The demand for good j can be written as

$$x_j = \frac{I_j}{p_j^{\sigma^T}}, \quad (2.7)$$

where I_j is the nation's total spending on x_j , which is a constant; $\sigma^T > 1$ is the elasticity of substitution between any two varieties, and p_j is the national price for j . Since j is only one of many varieties in the tradable goods market, changes in the production of x_j do not affect the price of the composite tradable good, which is still at price $p_T = 1$. The firm is a price-taker in factor markets; it solves the following profit-maximizing problem

$$\max_{N_c^T, K_c^T} \pi_j = p_j x_j - w_c N_c^T - \rho K_c^T. \quad (2.8)$$

Solving this problem, labor demand in the tradable sector is¹⁹

$$\ln N_c^T = [(1 - h_T)(\sigma^T - 1) - \sigma^T] \ln w_c + a_{TL}. \quad (2.9)$$

As a standard result of the CES production function, the log profit can be expressed as

$$\ln \pi_j = \ln \varphi_j + (\sigma^T - 1)[\ln B_{Tc} - h_T \ln w_c - (1 - h_T) \ln \rho], \quad (2.10)$$

where φ_j is a constant.

¹⁹ a_{TL} is a constant. Appendix B.3 provides detailed derivations for the problem of the firms in the tradable sector.

2.3.4.2 Firms in the Non-tradable Sector

Firms producing non-tradable goods face a competitive market.²⁰ Each firm's production function is also in the Cobb-Douglas form with constant returns to scale

$$X_c^N = B_{Nc}(N_c^N)^{h_N}(K_c^N)^{1-h_N}, \quad (2.11)$$

where B_{Nc} is the total factor productivity of the local non-tradable sector. The firm is a price-taker, while the output price is determined by the equilibrium of local non-tradable goods. The firm maximizes its profit by solving the following problem

$$\max_{N_c^N, K_c^N} \pi_c^N = p_c B_{Nc} (N_c^N)^{h_N} (K_c^N)^{1-h_N} - w_c N_c^N - \rho K_c^N. \quad (2.12)$$

Solving this problem, the implicit demand for labor in the non-tradable sector can be written as²¹

$$\ln X_c^N = \ln N_c^N + (1 - h_N) \ln w_c - (1 - h_N) \ln \rho + a_{NX}. \quad (2.13)$$

Local demand for non-tradable goods can be expressed as

$$X_c^N = \frac{\beta w_c N_c m_c^N}{p_c}, \quad (2.14)$$

²⁰The assumption that the firms producing tradable goods are monopolistic competitors while the firms producing non-tradable goods are perfectly competitive rests on both economic and technical explanations. Economically, the differentiable goods in the tradable sector are a result of the consumer's "love for variety." On the other hand, firms in the non-tradable sector are often small-scale service establishments that provide homogeneous services and fit the environment of perfect competition. Technically, in order to close the model, I need to specify these market structures for both sectors. If the firm in the tradable sector earns zero profit, its labor demand is flexibly adjusted such that the total local labor supply will not change after the shock. This is because local labor is supplied relatively more inelastically than capital. A firm in a perfectly competitive market facing exogenously given prices accommodates changes in labor supply by freely adjusting its capital input. In order to generate labor mobility across regions, I need to introduce a wedge such that the firm in the tradable sector faces other restrictions in adjusting its production.

²¹Detailed steps for the solution of this problem, as well as other conditions for the non-tradable sector listed below, are included in Appendix B.4.

where β is the share of income spent on non-tradable goods. $\beta w_c/p_c$ is the amount of non-tradable goods demanded by each civilian household. $m_c^H > 1$ is the demand shifter for the non-tradable goods, generated by the military presence in the local economy.

The market for the non-tradable goods is cleared within the local economy

$$\ln N_c^N - \ln N_c = h_N \ln w_c + \ln m_c^H - \ln p_c + a_{NL}. \quad (2.15)$$

Finally, the zero profit condition in the non-tradable sector yields

$$h_N \ln w_c = \ln p_c + b_N. \quad (2.16)$$

2.3.5 Equilibrium Conditions

The equilibrium of the local economy can be described in 5 equations: (1) local labor supply in equation 2.2, (2) local housing market equilibrium in equation 2.5, (3) labor demand in the tradable sector in equation 2.9, (4) labor demand in the non-tradable sector in equation 2.15, (5) zero profit condition for firms in the non-tradable sector in equation 2.16. We are interested in the effects of changes in military presence, Δm_c^H , Δm_c^N , on other outcome variables, differentiating these equilibrium conditions yields

$$\sigma^W \Delta \ln N_c = \Delta \ln w_c - \alpha \Delta \ln r_c - \beta \Delta \ln p_c \quad (2.17)$$

$$(1 + \eta_c) \Delta \ln r_c = \Delta \ln N_c + \Delta \ln w_c + \Delta \ln m_c^H \quad (2.18)$$

$$\Delta \ln N_c^T = [(1 - h_T)(\sigma^T - 1) - \sigma^T] \Delta \ln w_c \quad (2.19)$$

$$\Delta \ln N_c^N - \Delta \ln N_c = h_N \Delta \ln w_c + \Delta \ln m_c^N - \Delta \ln p_c \quad (2.20)$$

$$h_N \Delta \ln w_c = \Delta \ln p_c. \quad (2.21)$$

2.3.6 Welfare Impacts

In this subsection I derive expressions for the impacts of military personnel contractions on the welfare of workers, landowners, and firms in the tradable sector.²²

2.3.6.1 Workers/Households

Figure 2.2 illustrates the impacts on workers as the local real wage (u_c) drops. The horizontal axis represents the whole population of the nation, aligned by increasing e_{ic} , household i 's preference for location c , from left to right. The vertical axis shows the utility of household i living in location c or somewhere else c' .

The solid upward sloping line shows that the utility of living in location c increases with e_{ic} . The downward sloping line in dash-dots shows that the utility of living somewhere else drops as e_{ic} increases. The intersection of the two lines determines the marginal household, with $e_{ic} = e_{ic}^*$, which is indifferent between either location. Every household that has $e_{ic} > e_{ic}^*$ chooses to live in location c . Location c has a population equal to $1 - F(e_{ic}^*) = N_c$, where $F(\cdot)$ is the cumulative distribution function for e_{ic} .

Now consider a decline in u_c to u'_c due to military personnel contractions. The new utility of living in c as a function of e_{ic} moves down to the dashed line. Some households originally living in location c that have e_{ic} just above e_{ic}^* will migrate out. The new marginal household, with the idiosyncratic preference for location e_{ic}^{**} , will be determined by the intersection of the dash-dotted line and the dashed line. Population in location c drops to $1 - F(e_{ic}^{**}) = N'_c$.

Households that migrate and those that stay bear different welfare incidences. Households that choose to migrate out of location c have idiosyncratic preferences such that $e_{ic} \in (e_{ic}^*, e_{ic}^{**})$. These households suffer from a loss in utility, as they have to move to a less desirable location. For the household with $e_{ic} = e_{ic}^*$, welfare loss is zero; for the household with $e_{ic} = e_{ic}^{**}$, the loss is equal to $\Delta u_c = u_c - u'_c$; for households with $e_{ic}^* < e_{ic} < e_{ic}^{**}$, the welfare loss is somewhere in between. $\Delta \ln N_c$

²²Firms in the non-tradable earn zero profit, so they do not bear welfare incidence.

share of location c 's original residents choose to leave. The total loss for movers is illustrated in the graph as area B, which can be approximated by $\Delta \ln N_c \cdot \Delta u_c / 2$. Households that have idiosyncratic preferences for location c such that $e_{ic} > e_{ic}^{**}$ choose to stay. Each of these households bears a drop in utility that is equal to Δu_c . $(1 - \Delta \ln N_c)$ share of residents choose to stay. The total welfare loss for remaining households is illustrated in the graph as area A and can be expressed as $(1 - \Delta \ln N_c) \cdot \Delta u_c$. The total welfare change for households can be approximated as

$$\Delta V^W = (1 - \Delta \ln N_c) \cdot \Delta u_c + \frac{1}{2} \Delta \ln N_c \cdot \Delta u_c.$$

Recall that $\Delta u_c = \Delta \ln w_c - \alpha \Delta \ln r_c - \beta \Delta \ln p_c$, and plug in equation 2.21 for $\Delta \ln p_c$. We have the welfare change for workers

$$\Delta V^W = \left(1 - \frac{1}{2} \Delta \ln N_c\right) \cdot [(1 - \beta h_N) \Delta \ln w_c - \alpha \Delta \ln r_c] \quad (2.22)$$

2.3.6.2 Landowners

The welfare change for landowners is equal to the change in aggregate rents

$$\Delta V^H = \Delta \ln(r_c H_c) = \Delta \ln r_c + \Delta \ln H_c, \quad (2.23)$$

where H_c is the number of equilibrium housing units.

2.3.6.3 Firms in the Tradable Sector

The welfare change for the firm in the tradable sector is the change of its profit (equation 2.10)

$$\Delta V^T = \Delta \ln \pi_j = -(\sigma^T - 1) h_T \Delta \ln w_c. \quad (2.24)$$

Notice that $-(\sigma^T - 1) h_T < 0$, the firm in the tradable sector *gains* from a negative demand shock to the non-tradable sector, since local labor becomes cheaper.

2.3.7 Implication for Empirical Study

The model provides guidance for empirical studies. The equilibrium conditions and welfare implications suggest a vector of outcomes of interest $\{\Delta \ln p_c, \Delta \ln w_c, \Delta \ln r_c, \Delta \ln H_c, \Delta \ln N_c, \Delta \ln N_c^N, \Delta \ln N_c^T, \Delta H_c\}$.²³ We investigate all these outcomes in the empirical section. The model also involves a vector of parameters $\{\alpha, \beta, h_T, h_N, \sigma^T, \sigma^W, \eta_c\}$. I borrow from the macro-level data and existing studies to determine the values of these parameters whenever they are used.

2.4 Econometric Approach

2.4.1 Estimation Model

The impacts of the post-Reagan military personnel contractions on local economic outcomes over the period of 1988 and 2000 can be modelled empirically as

$$\Delta y_{kc} = \beta_k \Delta mil_c + \Delta \xi_{kc}. \quad (2.25)$$

$\Delta mil_c = mil_{c,2000} - mil_{c,1988}$ is the change in military presence in county c , where $mil_{c,t} = Mil_{ct}/Pop_{c,1980}$ is the number of military personnel in county c in year t scaled by its population in 1980, such that the impact is always proportional to the size of the local economy. Since population is also an outcome variable of interest, I use population data from 8 years prior to the cuts to avoid possible endogeneity. 1980 is also a census year, so the population count is more accurate and the division error is reduced. $\Delta y_{kc} = y_{kc,2000} - y_{kc,1988}$ is the change in outcome k of county c .²⁴ y_k may represent each of the following: (1) civilian employment divided by 1980 population, $emp_c = Emp_c/Pop_{c,1980}$; (2) earnings divided by 1980 population, $inc_c = Inc_c/Pop_{c,1980}$; (3) civilian population divided by 1980 population, $pop_c = CivPop_c/Pop_{c,1980}$; (4) log median rental price, $\ln r_c$; (5) private sector business establishments divided by 1980 population, $est_c = Est_c/Pop_{c,1980}$; (6) log average county wage, $\ln w_c$; (7) occupied housing units divided by 1980 population, $h_c = H_c/Pop_{c,1980}$; (8) civilian employment by sector s divided by 1980 population, $emp_{sc} = Emp_{sc}/Pop_{c,1980}$. These are the outcomes of interest as suggested by the

²³We do not observe price for local non-tradable good p_c , but it can be handily expressed by a function using equation 2.21.

²⁴Some outcomes are measured between 1990 and 2000.

model in section 2.3.

$\Delta\xi_{kc}$ is the composite error term that includes everything other than military personnel contractions that affects the dependent variable. I separate it into two components,

$$\Delta\xi_{kc} = \Delta\gamma_k\lambda_c + \Delta\varepsilon_{kc}.$$

The first term captures the secular trend of outcome k determined by time-invariant county characteristics λ_c . The term $\Delta\gamma_k\lambda_c$ exists in this long-differenced specification when there are location-specific trends in the outcome variables. For example, a county with a seaport is likely to be on a different labor market trajectory from a county in the desert. $\Delta\varepsilon_{kc}$ captures other unobservable contemporaneous shocks pertinent to the outcome. An example for $\Delta\varepsilon_{kc}$ is a weather shock in the period when the long difference is taken. I assume that $\Delta\gamma_k\lambda_c$ and $\Delta\varepsilon_{kc}$ are orthogonal to each other.

Both λ_c and $\Delta\varepsilon_{kc}$ are potentially unobservable. The OLS estimation of equation 2.25 can be biased since both $\Delta\gamma_k\lambda_c$ and $\Delta\varepsilon_{kc}$ can be endogenous, though for different reasons. The first is a omitted variable problem. Location choices for military bases are not random; particular locational characteristics both attract military bases and affect trends in economic performance. Larger cuts in military personnel are also concentrated in places with large military bases, so if these places are growing at a different rate, $\Delta\gamma_k\lambda_c$ and Δmil_c can be correlated. The second is a simultaneity problem. The DoD may avoid large cuts in places experiencing unfavorable idiosyncratic shocks in order to minimize the economic and political cost of military contractions, so $\Delta\varepsilon_{kc}$ and Δmil_c can be correlated.

When locational characteristics have no effect on the outcome variable or the effects are time-invariant such that long-differencing cancels out the effects, $\Delta\gamma_k\lambda_c = 0$. In this case, we can use a simple instrument to address the simultaneity problem. The instrumental variable I propose is the shift-share predictor widely used in the literature.²⁵ Specifically, I instrument the actual cut in military personnel in each location with the product of its pre-determined military presence and the size of the

²⁵Also known as the Bartik instrument due to Bartik 1991.

nation-wide cut. That is

$$\Delta mil_c^{IV} = mil_{c,1987} \cdot \frac{NtlMil_{2000} - NtlMil_{1988}}{NtlMil_{1988}}, \quad (2.26)$$

where $mil_{c,1987} = Mil_{c,1987}/Pop_{c,1980}$ is the military personnel to population ratio in county c the year before the post-Reagan military contractions. $NtlMil_t$ is the total number of military personnel nationwide in year t . $(NtlMil_{2000} - NtlMil_{1988})/NtlMil_{1988}$ is the percent change in nationwide military personnel between 1988 and 2000. Therefore, the instrumental variable is the predicted size of military personnel contractions in county c had the DoD adopted a simple rule of cutting military personnel everywhere by the same proportion. The instrument is valid as long as neither the pre-determined military presence nor the national trend is correlated with the idiosyncratic shock $\Delta \varepsilon_{kc}$. The nationwide military cuts were unlikely to be driven by local economic shocks in some particular locations. The post-Reagan military contractions were motivated by national political and fiscal situations, and each location is small compared with the whole nation.²⁶

The instrument is invalid when $\Delta \gamma_k \lambda_c \neq 0$ because Δmil_c^{IV} can be correlated with $\Delta \gamma_k \lambda_c$. To see this more clearly, notice that since the second term of Δmil_i^{IV} is the same for all observations in the long-difference equation, the variation in the instrument only comes from $mil_{c,1987}$ (denote as $\Delta mil_c^{IV'}$), which might be correlated with λ_c . For example, the fact that Washington, DC is the capital city (λ_c), leads to both large presence of military personnel ($mil_{c,1987}$) and the fact that it attracts private sector economic activities seeking political connection, and thus the city has promising trends in economic performance ($\Delta \gamma_k \lambda_c$). Therefore, the instrument is only valid conditional on the county-specific secular trend $\Delta \gamma_k \lambda_c$.²⁷

I use two approaches to partial out the county-specific secular trends. First, in a standard approach, I include observable pre-determined county characteristics in equation 2.25 in place of $\Delta \gamma_k \lambda_c$. That is, I estimate the following equation by

²⁶I test for this possibility in the Appendix by excluding local economies that were most likely to have influenced national policy. The results are robust.

²⁷In general, the Bartik instrument based on pre-determined sectoral composition can be correlated with time-invariant local fundamentals. As long as the effects of the fundamentals are not constant, differencing or controlling for location dummies cannot get rid of their effects, making the Bartik instrument invalid and the two-stage least squares estimate inconsistent. This point is worth stressing since many papers that use the Bartik instrument do not make this assumption explicit.

the Two-Stage Least Squares estimator using $\Delta mil_c^{IV'}$ as the instrument for Δmil_c

$$\Delta y_{kc} = \beta_k \Delta mil_c + \mathbf{X}_{kc} \cdot \gamma_k + \varepsilon_{kc}. \quad (2.27)$$

Each observation in the regression is a county with military bases in 1987. \mathbf{X}_{kc} is a vector of pre-determined county characteristics.²⁸ The identifying assumption is that the choice of the pre-determined characteristics correctly specifies the secular trends. This approach has a few limitations. First, it imposes parametric assumptions: it assumes that secular trends in all counties can be described by the same data-generating process. Since military bases are spread across the United States in all kinds of economic and geographic situations, a single data-generating process may fail to work for every county. More importantly, the paucity of information on county characteristics which I collect from population censuses and County Business Patterns may fail to correctly specify the secular trends, which may lead to inconsistent estimates.

2.4.2 A Generalized Synthetic Control Approach

In this subsection, I propose an alternative approach in order to partial out secular trends. This non-parametric approach eliminates $\Delta \gamma_k \lambda_c$ by constructing a counterfactual based on a large number of counties that did not have military bases in 1987. The idea is simple: since there was little change in the size of the military presence between 1980 and 1987 (see figure 2.1), the changes in outcome variables were on average driven by the secular trends. If I can find for each base county a comparison county with the same trajectories of the outcome variables in the period prior to the treatment, I can use the comparison county's outcomes in the period of treatment

²⁸These variables are constructed from various data sources including population censuses, Regional Economic Accounts from the Bureau of Economic Analysis, County Business Patterns, and County Data Books. \mathbf{X} includes state dummies and metropolitan status; civilian employment, earnings level (scaled by 1980 population) and growth from 1980 to 1987; log median housing price in 1980, 1990, and their difference; number of private business establishments (scaled by 1980 population) and growth from 1980 and 1987; demographic (racial, educational, age) and industrial (2-digit sectors) composition in the latest available year; and other social and economic indicators such as crime rate, number of physicians per 10,000 people, population density, road density, area, and terrain. When appropriate, \mathbf{X} also includes the quadratic forms of all aforementioned variables. I alternatively use a Bayesian Model Averaging (BMA) approach to select the most relevant covariates (Hoeting et al., 1999). The results from using \mathbf{X} and those from the BMA are very similar.

to represent the secular trends of the county with military bases.

The comparison is constructed as the weighted average of non-base counties. Formally, denote $c \in I$ as a base county, and $j \in J$ as a non-base county. For each county c , there is a set of weights, $w_{cj} \in \mathbf{W}_{cj}$, such that

$$\Delta\gamma_{kc}\lambda_c = \sum_{j \in J} w_{cj} \Delta\gamma_{kj}\lambda_j. \quad (2.28)$$

Note that there is no k subscript for the weights as the same set of weights should match the secular trend of any outcome variable k . This means that the constructed comparison county is the same in the pre-treatment period in all dimensions of interest. The weights are well behaved in the sense that they are bounded between 0 and 1 ($w_{cj} \in [0, 1]$, $\forall c, j$) and sum up to 1 ($\sum_j w_{cj} = 1$).

Equation 2.28 is the identification assumption, as we do not observe λ_c . In practice, I calculate weights for each base county c such that the following distance is minimized

$$w_{cj} = \underset{w_{cj}}{\operatorname{argmin}} \left\| \mathbf{Z}_c - \sum_{j \in J} w_{cj} \mathbf{Z}_j \right\|. \quad (2.29)$$

\mathbf{Z} is a vector of variables for which the distance is minimized. \mathbf{Z} includes the pre-treatment trajectories of the local economic conditions.²⁹

The underlying identification assumption is that by matching on the pre-treatment trajectories of the outcome variables, we are capturing the underlying mechanisms of the secular trends. This approach makes no assumption about the underlying determinants of the secular trends, and it allows for arbitrary functional forms of secular trends for each base county.

This approach is similar in spirit to the synthetic control method (?). In the standard synthetic control approach, a counterfactual, called the synthetic control, is constructed based on the past trajectory of the outcome variable up to the period of treatment. The treatment effect is obtained by subtracting the outcome of the

²⁹I pick the important outcome variables as predicted by the model. Specifically, these variables are $(Emp_{c,1984} - Emp_{c,1980})/Pop_{c,1980}$, $(Emp_{c,1987} - Emp_{c,1980})/Pop_{c,1980}$, $(Inc_{c,1984} - Inc_{c,1980})/Pop_{c,1980}$, $(Inc_{c,1987} - Inc_{c,1980})/Pop_{c,1980}$, $r_{c,1990} - r_{c,1980}$, $p_{c,1990}^h - p_{c,1980}^h$. In fact, the results are robust to alternative set of predictors. All the predictors are given the same weight in the loss function. The results are quantitatively similar when the inverse of standard deviation is used as weights. As in all matching-based methods, there is a tradeoff between the number of variables to be matched on and the average matching quality for each variable.

synthetic control from the real outcome. I extend this approach in three important dimensions. First, instead of focusing on one outcome variable, I match on a vector of outcome variables. Second, the synthetic control assumes that the treatment is binary and exogenous, while in this case the treatment is continuous and endogenous. Third, the standard synthetic control model has only one or a small number of treated units; here I have hundreds, so I pool pairs of base counties and their synthetic controls and estimate using weighted 2SLS. Despite the differences, with some abuse of the terminology, I call this way of constructing comparing groups a “synthetic control approach”.

This approach has a few advantages over the first approach, which simply includes pre-determined covariates. First, this non-parametric approach allows for a county-specific secular trend for each treated unit, which is more flexible than parametric specifications. Since military bases are located in virtually all types of counties, the individual matching approach is arguably more precise than pooled regressions. Second, by constructing a counterfactual based on the trajectories of outcome variables in the pre-treatment period, the chance of mis-specification is lower. In this case, mis-specifying the covariates not only reduces the precision of the estimation, but also invalidates the instrument and results in inconsistent estimates. In Appendix C, I show using a simulation exercise that mis-specifying the secular trend in a parametric way leads to large biases, while non-parametric matching based on pre-treatment trajectories of outcome variables leads to reduced bias in a finite sample. Finally, this approach finds a synthetic county that is similar in many important dimensions of interest, which is conceptually consistent with a “counterfactual” county.

I denote a county group g as the collection of the base county and the counties making up its synthetic control, and estimate the following equation using Weighted Two-Stage Least Squares estimator

$$\Delta y_{khg} = \beta_k \Delta mil_{hg} + \theta_g + \varepsilon_{khg}, \quad (2.30)$$

with the first stage equation

$$\Delta mil_{hg} = \eta \Delta mil_{hg}^{IV} + \zeta_{hg},$$

where county $h \in I \cup J$ is either a base county or a non-base county and θ_g is a dummy variable that is equal to 1 if county h belongs to county group g . I use weights from the synthetic control and assign counties with military bases weights equal to 1. Each county group thus has a total weight equal to 2. $\Delta mil_h^{IV'} = mil_{h,1987}$ is used as the instrumental variable for Δmil_{hg}^{IV} .

The standard errors of estimating equation 2.30 should be adjusted for multiple reasons. First, a non-base county can be used multiple times as parts of the synthetic control for different base counties. Second, the fact that a base county and its synthetic control had similar secular trends suggests that they are likely to be affected by similar shocks. Finally, estimation of equation 2.30 involves using weights in the synthetic control that are constructed from a previous step. Standard errors should be adjusted to reflect the uncertainty in the way the weights are constructed. I report two sets of standard errors. First, I cluster the standard errors by county (h) and county group (g), using the multi-way clustering technique proposed by Cameron et al. (2011). The two-way clustered standard errors utilize the prior knowledge about the error structure, but do not take into account of the uncertainty of the synthetic control approach in the first step. Alternatively, I bootstrap the whole procedure to conduct statistical inferences.³⁰ The bootstrapped standard errors incorporate variability from the construction of the weights, but can be less efficient since it does not fully incorporate prior knowledge about the error structure. Nevertheless, both methods give very similar statistical inference.

To illustrate this novel identification strategy that combines synthetic control and weighted 2SLS, figure 2.3 shows the case of San Diego County, which has the largest number of military personnel in the United States. The four panels in figure

³⁰One complete procedure includes (1) re-sampling with replacement separately for the sample of base counties and the sample of non-base counties, (2) constructing synthetic controls for each base county from the pseudo-sample of the non-base counties, and (3) estimating equation 2.30 and clustering the standard error by county group and county. The standard error from the bootstrap-se procedure is the standard deviation of the estimated coefficients from bootstrapped pseudo-samples. Abadie and Imbens (2008) show that bootstrapping fails for simple nearest neighbor matching. The intuition is that, by restricting the closest comparison units to a fixed small number, bootstrapping fails to generate meaningfully different pseudo-samples. They conjecture that bootstrapping works for kernel density matching since it allows somewhat different comparison units to enter the sample, though the convergence rate is slower. My approach is similar in spirit to kernel density matching in the sense that more similar comparison units are given higher weights and less similar comparison units are given smaller weights. In Appendix D, I show Monte Carlo simulation evidence that bootstrapping leads to a rejection rate that is close to the true rate. In Appendix D, I also compare statistical inferences based on the bootstrap-se procedure and that based on the bootstrap-t procedure.

2.3 show changes in military presence and changes in civilian employment, total civilian earnings, civilian population, and private business establishments for both San Diego and “synthetic San Diego”, with the levels in 1980 standardized to 1. Between 1980 and 1987, the ratio of military personnel to 1980 population in San Diego stayed at around 0.08. This number gradually dropped to below 0.06 in 2000. Synthetic San Diego, on the other hand, had military presence close to zero throughout the period. San Diego and its synthetic control exhibit similar trajectories in all four sequences before 1988. After 1988, San Diego experienced slower growth in employment, earnings, population, and the number of business establishments than its synthetic control. The difference between the changes in outcome variables in San Diego and in its synthetic control is attributed to the net changes of military presence in San Diego. Any possible correlation between the size of military cut and idiosyncratic shocks is addressed by using the Bartik instrument.³¹

Figure 2.3 shows that the synthetic control for San Diego closely resembles San Diego in pretreatment trends. In table 2.1, I report the overall quality of synthetic controls for all base-counties. Each row shows the changes in the outcomes of interest prior to the military cuts. Column 1 reports the changes for base counties, and column 2 reports the changes for non-base counties. On average, base counties were on a better growth path than non-base counties, prior to the cuts. For example, between 1980 and 1987, civilian employment grew by 19 percent in base counties, and only 11 percent in non-base counties. Column 4 reports the p -value for the differences between column 1 and column 2, which shows that the differences are always statistically significant at the 1% level. Column 3 shows that the weighted averages of the synthetic counties are very similar to those of the base counties (column 1); the differences are statistically insignificant at any conventional levels, regardless of whether the comparison is made on average (column 5) or within each county pair (column 6).

³¹Mathematically, the weighted 2SLS estimation of equation 2.30 is equivalent to estimating the following equation $(\Delta y_{kc} - \sum_j w_{cj} \Delta y_{kj}) = \beta_k (\Delta mil_c - \sum_j w_{cj} \Delta mil_j) + \Delta \tilde{\varepsilon}_{kcg}$, using the 2SLS with $(\Delta mil_c^{IV'} - \sum_j w_{cj} \Delta mil_j^{IV'})$ as the instrument. Each observation is a county with military bases. Estimating this equation gives results very similar to those from equation 2.30.

2.5 Data and Sample

This chapter uses multiple sources of publicly available data. This section briefly introduces data sources and the construction of the main sample. Appendix A has a more detailed description of how the key variables are constructed.

2.5.1 Data

The Department of Defense (DoD) publishes the distribution of military bases in its annual Base Structural Report (BSR). The BSR from 1989 reports a snapshot of 535 major military bases with at least 250 personnel as of February 1988, eight months before the enactment of the Base Re-alignment and Closure Act. The geographic boundaries for most of these military bases are obtained from a digitalized map called “Military Installations, Ranges, and Training Areas Boundaries”, prepared by the DoD and published as part of the National Transit Database in 2011. For the remaining bases, I searched for historical maps as well as web resources including Google Maps and Wikipedia to determine their location and boundaries. I then overlaid the GIS map with the 1989 military bases onto a GIS map of US counties from 1990 Census. I consider a base county to be any county in which at least parts of its territory intersect with a military base.

The number of military personnel by county in each year comes from the Regional Economic Accounts (REA) published by the Bureau of Economic Analysis, a panel dataset at the county level with annual local economic characteristics. The REA collects data from another DoD annual publication called the Distribution of Personnel, which gives detailed information on DoD employees on military bases and in major cities and converts this into the number of military personnel by county.³²

The outcomes of interest in this chapter capture aspects of the size of the population, and changes in the strengths of the local labor market, local housing market, and local businesses. County population counts are county-level tabulations of decennial censuses, with the number in non-census years interpolated. The REA also reports employment and earnings by 2-digit sectors, from which I calculate total civilian sector employment and earnings. County median rental prices and units of

³²The Distribution of Personnel reports both military personnel and civilian employees who work on bases. Unfortunately, the REA only separately reports the number of military personnel. Civilian employees of the DoD are grouped into a broader category called federal civilian workers.

occupied housing also come from county-level tabulations of decennial population censuses. These variables are based on self-reported values of individual households responding to the census. The annual number of private business establishments by 2-digit sectors comes from County Business Patterns.

2.5.2 Sample

This chapter treats counties as local economies since they are the level of aggregation for most of the key variables used in this chapter.³³ Due to concerns that military personnel contractions can spillover across county borders, in one of the robustness checks I use commuting zones as the alternative definition of local economies.

The REA groups some neighboring small counties into new county-level units. Throughout the chapter I use the REA definition of counties and convert county-level variables from other sources accordingly. To avoid cases in which the military is the only economic activity, I drop counties that had fewer than 5,000 civilians in 1980. I further drop about 10 percent of base counties that actually experienced *increases* in military presence between 1988 and 2000. Since the effects can be asymmetric for expansions,³⁴ dropping a small number of counties that experienced expansions helps with the interpretation of the estimation results. The final sample has 335 base counties. There are also 2,429 non-base counties prior to matching.

2.6 Empirical Results

2.6.1 Comparing Specifications

I estimate various specifications of the baseline model in equation 2.25 and equation 2.30, with and without using the Bartik instrument. Table 2.2 reports these estimates using changes in civilian employment between 1988 and 2000 as the outcome variable. Since both the outcome variable and the key explanatory variable Δmil_c , the 1988-2000 change in military personnel, are scaled by the 1980 population, the coefficient associated with Δmil_c can be conveniently interpreted as the number of

³³Existing studies have used various definitions of local economies, including counties (e.g., Black et al., 2005), commuting zones (e.g., Autor et al., 2013), metropolitan areas (e.g., Moretti, 2010b), and states (e.g., Blanchard and Katz, 1992).

³⁴For example, Glaeser and Gyourko (2005) and Notowidigdo (2013) explain why positive shocks and negative shocks have asymmetric effects.

civilian jobs lost that is associated with cutting one military person.

I first estimate equation 2.27 using OLS on a sample of only base counties. I control for a vector of pre-determined county-level economic and demographic characteristics. The results are reported in column 1 of table 2.2, which shows that cutting one military worker is associated with losing 0.8 jobs in the civilian sector. The estimate is statistically significant at 1% level. The OLS estimate may be inconsistent due to simultaneity bias and omitted variable bias. If the DoD avoided large cuts in counties with negative idiosyncratic shocks, or counties experiencing large cuts had otherwise faster economic growth, the OLS estimate is likely to be biased downwards. Column 2 estimates the same equation as in column 1 but uses the initial military presence ($\Delta mil_c^{IV'}$) as the instrument for the change in number of military personnel. The estimated coefficient is 1.2, and it is statistically significant at the 1% level. The first stage is strong with F statistics of about 60. Consistent with the hypothesized source of endogeneity, the 2SLS estimate is larger than the OLS estimate, although a simple Hausman test shows that the difference is not statistically significant.³⁵

If the covariates included in the regression in column 1 and column 2 misspecify the secular trend, the estimates could still suffer from omitted variable bias and the instrumental variable would be invalid. Column 3 in table 2.2 uses the generalized synthetic control method introduced earlier to purge out the secular trends and estimate equation 2.30 by weighted 2SLS. The weights constructed from the synthetic control approach are used in the regression. The number of observations increases to 19,787 just because each component of the synthetic control is included as an observation, no matter how small a weight it bears in the regression. The weights in each county group sum to 2. The estimated coefficient is about 1.3, similar to that in column 2. I report two sets of standard errors. Standard errors in parentheses are two-way clustered at the county level and the county-group level. Standard errors in curly brackets are the standard deviation of 100 bootstrapped estimates. The two-way clustered standard errors are very similar to those in column 2. Both two-way clustered standard errors and the bootstrapped standard errors give the same statistical inference and are significant at the 1% level.

Column 4 re-estimates equation 2.30 using weighted 2SLS while controlling for the same set of covariates as in column 2. If the specifications in column 2 and

³⁵When conducting the Hausman test, I use conventional standard errors.

column 3 are both valid, including valid controls will not change the point estimate but only improve the precision of the model. I find an estimated coefficient of around 1.3, and the clustered standard error is slightly smaller than that in column 3. The result in column 4 lends more confidence to the findings in column 2 and column 3. Equation 2.27 and equation 2.30 rely on different identification assumptions and use different sources of variation, so the fact that both yield very close results is reassuring. For the remainder of the chapter I use the specification in column 3 as the baseline for its flexibility in assumptions and good quality in matching, as argued and demonstrated in subsection 2.4.2.

The number of additional jobs created or destroyed in the local economy as a result of adding or losing one job is called the local job multiplier. Table 2.2 column 3 shows that the local multiplier is about 1.3. This number falls within a wide range of estimates from previous studies. Using coal price fluctuations as exogenous shocks, Black et al. (2005) estimate that one coal-mining worker brings an additional 0.35 workers to the local labor market within a four-year period. One possible reason for the small effect might be that fluctuations in coal price was perceived as temporary shocks, so local businesses do not adjust on the extensive margin. Moretti (2010b) finds that one additional job in the manufacturing sector creates another 1.6 jobs in the service industry at the MSA level over a decade, a number much closer to mine. This number rises to 2.5 additional jobs, if the manufacturing job is filled by a skilled worker.

2.6.2 Local Labor and Housing Markets

Using the same specification as in column 3 of table 2.2, table 2.3 reports the estimates on a vector of outcomes regarding local labor and housing markets. Columns 1 through 5 report the effects of military personnel contractions on the levels of local economic activity. Column 1 repeats the estimate on changes in civilian employment. Columns 2 through 5 report the estimates on civilian earnings (in thousands of 2000 dollars), civilian population, the number of private business establishments, and the number of occupied housing units. All these outcome variables are divided by the county population in 1980. Therefore, the coefficients can be interpreted as the changes in levels of each outcome due to cutting one military worker.

The military personnel contractions have sizable effects on levels of civilian employment, earnings, population, and business establishments. Specifically, cut-

ting one military worker causes a reduction of 1.3 civilian jobs (column 1) and 31,000 dollars in earnings (column 2) of civilian workers. It also causes a drop of 3.1 civilians in the local population (column 3). This is a large migration response relative to the impacts on the local labor markets. The result suggests that essentially every individual who lost a job left the local economy with his or her family.³⁶ The number of local business establishments also declines with local employment. For every 10 military jobs cut, a local business is also lost.³⁷ All of these estimates are statistically significant.

The military to population ratio in counties with military bases dropped by about one percentage point, on average, during the sample period. At the bottom of each column, I convert the effects to percent changes. Reducing the military to population ratio by one percentage point reduces civilian employment by 2.3 percent, total civilian earnings by 1.8 percent, and the civilian population by 2.9 percent. The fact that the percent decline in employment is greater than the percent decline in total earnings suggests that the jobs lost were relatively low-paying.

Column 6 reports the effect on changes in the ratio of civilian employment to civilian population. Given the large migration response to job losses, it is not surprising that employment to population ratios are not seriously affected. In fact, the estimate is small, statistically insignificant, and has the wrong sign: reducing military to population ratio by one percentage point increases the civilian employment to population ratio by 0.3 percentage points.³⁸

³⁶According to the 1990 census, each military person has on average 0.7 dependents living in the same household. About 60 percent of the civilian population works in the civilian sector. So a loss 1.2 civilian jobs is associated with 2 civilians. Therefore, cutting one military worker directly affects about 2.7 civilians. A reduction of 3.1 civilians is larger than that back-of-the-envelope calculation.

³⁷The effects are concentrated in small firms with fewer than 25 employees.

³⁸A natural alternative measurement is unemployment rate. The county-level unemployment rates are from the Local Area Unemployment Series (LAUS) published by the Bureau of Labor Statistics (BLS). Using the unemployment rate has three disadvantages. First, county-level unemployment rates are only available from 1990. Second, county-level unemployment rates are model based, and the model has changed since 2000. Third, since unemployment rates are based on state unemployment insurance claims, it is possible that displaced workers who have migrated are still counted as unemployed in their original counties. Using county unemployment rates as the outcome variable, I find that reducing the military to population ratio by one percentage point increases the unemployment rate between 1990 and 2000 by 0.14 percentage points. The estimate is statistically significant at the 5% level but is very small in magnitude. The average unemployment rate was around 5.4 during the period. About 60% of the population is in the workforce, so an increase in the unemployment rate of 0.14 percentage points roughly corresponds to 0.08 unemployed workers per 100 people. As the same cut reduces the civilian population by 1.3 per 100 people, this suggests that the vast majority of the displaced workers are not unemployed by

The next two columns report the effects on log local prices. Therefore, the coefficients can be conveniently interpreted as semi-elasticities. Column 7 reports the effects on local average wages. Due to data limitations, I calculate local average wages based on county-level variables adjusted by demographic characteristics. First, I calculate annual raw average wages by dividing total county wages and salary earnings with total county wages and salary employment, both from the REA. I then regress log raw county average wages on a vector of county demographic characteristics. These characteristics include racial compositions (white, black, and other), percent of adults with college degrees, and the quadratic terms of these variables. County demographic characteristics are drawn from decennial censuses. Demographic characteristics in non-census years are interpolated.

Column 7 shows that cutting the military to population ratio by one percentage point reduces the log wage by 0.47 percent. The effect is not statistically significant and is quite small in magnitude: for an average job paying 30,000 dollars a year, a 0.47 percent decline is a loss of 141 dollars per year. Although average wages are potentially measured with a lot of error, the estimated effect is consistent with the estimated effects on other results. First, effects on civilian employment and civilian earnings suggest that the effects on raw average wages are small.³⁹ Second, the large migration response suggests that workers are rather mobile; large changes in wages will be arbitrated out by large out-migration. Column 8 reports the effects on the median rental price. Cutting the military to population ratio by one percentage point reduces log median rent by 1.3 percent. The decline in rental prices is a result of both lower wages and smaller population. Since wages do not drop by much while many people leave the local economy, it is not surprising that the effect on rental prices is larger than that on wages.⁴⁰

the end of the sample period.

³⁹Note that if workers are homogeneous, changes in log wages can be expressed in terms of log changes in total civilian employment and log changes in total civilian earnings, that is $\Delta \ln Wage = \Delta \ln Earning - \Delta \ln Employment$. From column 1 and column 2 in table 2.3, we have the $\Delta \ln Wage$ due to a one percentage point decline in the military to population ratio equal to $(-1.796 - (-2.345)) = 0.549$. That is, the wage would *increase* by 0.5 percent. Comparing this result with that in column 5 also suggests that military personnel contractions affect low-skilled workers more severely.

⁴⁰An alternative measure of local housing market conditions is the median housing price. Since conceptually the housing price is just the present discounted value of future rental prices, the percent change in rental prices and housing prices due to the same shock should be the same. Using the same specification, I estimate an effect that cutting the military to population ratio by one percentage point reduces local housing prices by 0.9 percent. Smaller responses in housing prices than in rental prices are often found in the literature, especially when housing prices are

Throughout table 2.3, I report two sets of standard errors: the two-way clustered standard errors are in parentheses, and the bootstrapped ones are in curly brackets. Both sets of standard errors are very similar and give the same statistical inference, but the bootstrapped standard errors are always slightly smaller. Unless otherwise specified, for the rest of the chapter I report only two-way clustered standard errors so as to be conservative about statistical inference.

2.6.3 Sectoral Composition

The model in section 2.3 incorporates the tradable and non-tradable sectors and gives different predictions for them under military personnel contractions. According to equation 2.20, labor demand in the non-tradable sector declines since local wages (w_c), local population (N_c), and the demand shifter for local non-tradable goods (m_c^N) all decline. In contrast, according to equation 2.19, employment in the tradable sector may in fact increase as the firm hires more workers when local labor costs become lower. I therefore investigate separately the impact of military personnel cuts on different sectors. Moreover, I also investigate the effects on the number of federal civilian workers since those employed by the DoD are directly affected by the downsizing of military operations.

Table 2.4 reports the results of re-estimating equation 2.30 using changes in civilian employment in each 2-digit sector as the outcome variables. Column 1 shows that for cutting every four military jobs cut, one federal civilian job is also lost. In the sample base counties in 1988, federal civilian workers account for 2.5% of the total population. Therefore, a one percentage point drop in the military to population ratio reduces the average level of federal civilian employment by about 10%, which is shown in the last row of column 1. Column 2 reports the effect on employment in the manufacturing sector, a typical tradable sector, where the military personnel contractions have a small and negative effect and the estimate is not statistically significant. So there is no evidence that local firms in the tradable sector expand their employment. Columns 3 through 5 show the employment responses in non-tradable sectors. I find large and statistically significant effects in the construction and the retail sectors, and a sizable effect on the service sector, although it is not self-reported. Because houses are only infrequently transacted, self-reported housing prices may not reflect current housing market conditions (Greenstone and Gallagher, 2008; Busso et al., 2013). Therefore, I use rental prices instead of housing prices.

statistically significant. These findings are in agreement with previous studies that have also found that the impacts are concentrated in the local non-tradable sectors and that the tradable sector is not affected (e.g., Black et al, 2005).

2.6.4 Welfare Impacts

I can calculate the magnitude of the welfare changes for workers, landowners, and firms by plugging the empirical results found in the previous section into the expressions for welfare analysis in section 2.3.6, i.e., equations 2.22, 2.23, and 2.24. In order to calculate the welfare impacts, I first need to calibrate the model parameters involved in these expressions. I choose the values of the relevant parameters from national accounts as well as other studies. The share of expenditure on housing, α , and the share of income spent on non-tradable goods, β , come from the BEA National Income and Product Accounts Tables, $\alpha = 0.18$, $\beta = 0.46$.⁴¹ The share of labor in the non-tradable sector production, h_N , and the share of labor in the tradable sector, h_T , both take the value of 0.7.⁴² The constant elasticity of substitution among differential goods is taken from the trade literature: $\sigma^T = 2.2$.⁴³ With a one percentage point drop in the military to population ratio, the changes in welfare for each agent are

$$\begin{aligned}\Delta V^W &= -0.02\% (0.72\%) \\ \Delta V^H &= -3.6\% (0.39\%) \\ \Delta V^T &= 0.3\% (0.22\%).\end{aligned}$$

That is, the welfare of workers drops by 0.02 percent and the welfare of landowners drops by 3.6 percent, while the welfare of firms in the tradable sector increases

⁴¹Source: Section 2: Personal Income and Outlays. Table 2.3.5: Personal Consumption Expenditures by Major Type of Product for 2012 full year. Expenditure on housing includes “Housing and utilities” (line 15). Expenditure on non-tradable goods includes all service items except for housing (lines 15-21). URL: <http://www.bea.gov/iTable/iTable.cfm?reqid=9&step=1&acrdn=2#reqid=9&step=3&isuri=1&903=65>. Last accessed on Aug 24, 2014.

⁴²This value is the commonly used labor’s share of national income, e.g., (e.g., Krueger, 1999).

⁴³Source: Broda and Weinstein (2006). The median value of 3-digit SITC code for tradable goods. URL: http://www.columbia.edu/~dew35/TradeElasticities/ElasticitiesBrodaWeinstein90-01_SITCRev3_3-digit.xls. Last accessed Aug 24, 2014.

by 0.3 percent. Standard errors are reported in the parentheses for each welfare calculation.⁴⁴ Only the welfare impact on landowners is statistically significant.

The average decline in the military to population ratio was about one percentage point for counties in the sample over the 1988 to 2000 period. The aggregate welfare impact is small for workers and firms in the tradable sector but substantial for landowners.⁴⁵

2.7 Adjustment of Local Economies

2.7.1 Estimating Dynamic Effects

The results thus far show the long-run effects of military contractions.⁴⁶ It is also interesting to investigate how quickly local economies adjust to the shocks and achieve new equilibria. For example, although there is little welfare loss for workers in the new equilibrium, if it takes a long time for the local economy to arrive at the new equilibrium, the welfare cost in the transition might still be large.

The existing literature has not reached a consensus on how long it takes for new equilibria to be achieved. Blanchard and Katz (1992) and Feyrer et al. (2007) find that employment levels and unemployment rates in US states, MSAs, and counties recover from negative shocks within a decade. Bartik (1991) finds a somewhat longer period of recovery. Still other studies (e.g., Yoon, 2014) find that negative shocks can have lasting impacts on local economies.

Ideally, we would have panel data and a one-time shock so that we could capture the dynamic effects of the shock using the impacts of the shock on outcome variable changes over many periods both before and after the shock. In the setting of this study, however, the shocks continued for over a decade and were followed

⁴⁴The variance-covariance matrix for the empirical results is obtained by bootstrapping. Denote $b_k = \{\hat{\beta}_k^1, \dots, \hat{\beta}_k^B\}$ as the vector of estimated coefficients for outcome k from bootstrap. The bootstrapped standard error for the original estimate $\hat{\beta}_k$ is the standard deviation of b_k , σ_{b_k} . Denote b_l as the vector of estimated coefficients for another outcome l , the bootstrapped correlation coefficient between outcome k and outcome l is $cov(b_k, b_l) / \sigma_{b_k} \cdot \sigma_{b_l}$. Table 2.9 shows the bootstrapped standard errors for the full vector of outcome variables and their correlation coefficients. The standard errors for welfare calculations are then obtained using the Delta Method.

⁴⁵Appendix E provides a host of robustness checks and estimates heterogeneous effects.

⁴⁶To be precise, the effects shown here are the changes in county outcomes over a 12-year period as a result of military contractions that took place anytime in between. I loosely call these estimates long-run effects. Most of the cuts in military personnel took place between 1988 and 1995, and as I will show in this subsection, the full effects of a particular cut was usually realized within 3 years.

by a different treatment period resulting from the anti-terrorism wars prompted by 9/11. Therefore, including many periods of leads and lags is not feasible. Instead of depicting the full trajectory of dynamic effects, this subsection will address the less demanding question of how many years it takes for a shock's effects to become negligible.

I propose two alternative approaches. First, I estimate a panel model with a few leads and lags. I intend to identify the limits of leads and lags that have any effect on the outcome in the current period. Second, I run long-differenced equations with varying numbers of years in between. As we increase the length of the time period for which the difference is taken, the estimate will eventually converge as the the local economy achieves the new equilibrium, since the inclusion of more years does not make a difference.

In the first approach, I estimate the following regression

$$\Delta y_{khgt} = \sum_{s=t-T}^{t+T} \beta_{ks} \Delta mil_{hgs} + \theta_{gt} + \Delta \varepsilon_{khgt}, \quad (2.31)$$

where, as earlier, a county h is either a base county or a non-base county. g refers to the county group constructed from the synthetic control approach, and t indexes year. $\Delta y_{khgt} = y_{khgt} - y_{khgt-1}$ is the first-differenced outcome k in county h . The outcomes include civilian employment, earnings, and private businesses. Dynamics of rental price, population, and units of occupied housing are not included since they are only available decennially. $\Delta mil_{hgs} = (Mil_{hgs} - Mil_{hgs-1})/Pop_{h,1980}$ is the first differenced number of military personnel in the county, scaled by its 1980 population. I include past and future cuts up to two years ($T = 2$) in order to capture the dynamic effects.⁴⁷ $\Delta \varepsilon_{khgt}$ is the error term.

Δmil_{hgs} is instrumented using the Bartik instrument as discussed earlier

$$\Delta mil_{hgs}^{IV} = mil_{h,1987} \cdot \frac{NtlMil_s - NtlMil_{s-1}}{NtlMil_{s-1}}. \quad (2.32)$$

$mil_{h,1987} = Mil_{h,1987}/Pop_{h,1980}$ is the initial military presence. $(NtlMil_s - NtlMil_{s-1})/NtlMil_{s-1}$

⁴⁷Ideally, I would like to include many leading and lagging terms in order to capture the full dynamic effects. However, including more terms creates two problems. First, the number of observations will be smaller for years between 1988 and 2000. Second, including more leads and lags increases the likelihood of incurring multi-colinearity problems.

is the annual percent change in nationwide military personnel. In a dynamic panel setting, the variation of the instrument comes from both cross-county and within-county. Finally, the standard errors are two-way clustered at the county group-year level and at the county level.

Column 1 in table 2.5 shows the dynamic effects on employment. Changes in employment are mainly effected by changes in military personnel contractions that takes place in the current year (t) and those took place in the year before ($t - 1$). Both estimates are sizable and statistically significant. The sum of the effects in these two years (1.35) is close to the overall effect found in the long-differenced specification (1.26). Although the sample does not allow for the inclusion of more lags, a comparison of the dynamic results here with those from the long-differenced specification suggests that a military personnel cut has a permanent effect on the level of employment: the effect takes place within a couple of years after the cut, and there is no evidence of recovery.

Since most of the cuts were planned and anticipated, future cuts could have effects on current employment as people respond to future cuts by changing current behavior. Contrary to that prediction, I find that military cuts in the next the the year after ($s = t + 1$ and $s = t + 2$) have small and statistically insignificant effects. The dynamic effects on civilian sector earnings (column 2) and business establishments (column 3) exhibit similar patterns, although the estimates of the effects on earnings are not statistically significant. I report the Angrist-Pischke partial F statistics to assess the power of instruments for each endogenous variable. The first stages are strong.

The dynamic panel models with leads and lags potentially suffer from colinearity, the many IV problem, and loss of observations with the inclusion of more leading and lagging terms. In the second approach, I estimate a series of long-differenced equations with increasing year gaps and see how long it takes the estimates to converge. Suppose the treatment takes place in year 0; this approach builds the intuition that if the local economy has achieved a new long-run equilibrium, say in year n , we would find the same result by comparing the local economy in year 0 with that in year n , or with any year after n . The smallest year gap for which we observe the long-run effect is likely to be the amount of time it takes the local economy to achieve the new equilibrium. Formally, I estimate variations of the following equation

$$\Delta y_{kcl} = \beta_{kl} \Delta mil_{cl} + \theta_{gl} + \Delta \varepsilon_{kcl}, \quad (2.33)$$

where $\Delta y_{kcl} = y_{kc,1988+l} - y_{kc,1988}$ and $\Delta mil_{cl} = mil_{c,1988+l} - mil_{c,1988}$, and θ_{gl} are dummies for county-group-year-group dummies. The Bartik instrument, Δmil_{cl}^{IV} is constructed in the same manner. The standard errors are two-way clustered at the county-group-year-group level and the county level. Recall that the cuts in military personnel during the sample period were continuous, which means that cuts in two different periods, Δmil_{cl} and $\Delta mil_{cl'}$, are likely to be positively correlated. For the same reason, Δmil_{cl}^{IV} is also correlated with $\Delta mil_{cl'}$. $\Delta mil_{cl'}$ will be the omitted variables in equation 2.33, and the 2SLS estimate of equation 2.33 will be upward biased. The correlation will be smaller when the two periods are further away or when l is larger. Therefore, when $\hat{\beta}_{kl}$ and $\hat{\beta}_{kl'}$ ($l' > l$) are similar, it is evidence that when we look at a long enough period of time, the omitted variable bias is not too much of a concern.

Figure 2.4 shows the effects with l ranging from 1 to 12. For all three outcome variables, the effects achieve a plateau in about 3 years. This result is consistent with those in table 2.5 and shows that local economies are relatively quick to adjust to shocks.

2.7.2 Mobility

The quick adjustment of local economies and the large effects on population after negative shocks are consistent with the high mobility of the US population. During the 1990s, about 3% of the population moved across state borders and about 5% of the population moved across county borders every year (Molloy et al., 2011). Low skilled workers, however, are less mobile and more likely to be affected by local shocks. As a result, the negative demand shocks tend to have longer effects on local economies (Glaeser and Gyourko, 2005; Notowidigdo, 2013; Yoon, 2014).

One explanation for the high mobility found in this chapter is that people displaced due to military personnel contractions are more mobile than typical displaced workers. Military workers have little connections to the localities where they serve. Once discharged, they are likely to leave with their dependents, some of whom work in the local economy. Veterans, who tend to cluster near military bases in order to benefit from amenities such as commissaries and post exchanges, also tend to

leave as these amenities pull out along with the declining military presence.⁴⁸ Some civilian workers employed by the DoD have special skills and need to migrate to find good fits. All of these factors may contribute to the large migration response after military personnel contractions.

I test whether there is something special about military contractions in terms of the resulting migration response by estimating the effects of a different type of shock that is arguably more similar to conventional demand shocks. The post-Reagan military contractions involved not only cuts in military personnel contractions, but also cuts in military procurement contracts awarded to local companies. Unlike military personnel contractions, people affected by military procurement reductions are not directly associated with military operations. If indeed military personnel contractions generate particularly high migration responses, we would expect the migration response for each job loss due to military procurement contractions to be lower.

The inclusion of cuts in military procurement contracts to local contractors also serves as a test to check whether military procurement contractions are correlated with personnel contractions. If that is the case, the impacts of personnel contractions found in previous sections are overestimated, and including procurement contractions will significantly reduce the estimates. That said, it is unlikely to be the case, since the spatial distribution of changes in military personnel and those in procurement are not correlated.⁴⁹

I use contract-level data from the Federal Procurement Data System (FPDS) and identify military procurements by restricting procurements to those ordered by the Department of Defense. Since 1969, every contract worth more than 25,000 dollars was recorded. Crucially, the files show the amount of each contract and the county address of the primary contract awardee. I aggregate the amount of procurement to the county-year level. Changes in procurement contractions during the post-Reagan military contractions are measured as $\Delta proc_c = (Proc_{c,98-00} - Proc_{c,86-88}) / Pop_{c,1980}$, where $Proc_{c,86-88}$ is the average value of procurement awarded

⁴⁸I find that reducing one military worker increased about 0.2 new veterans in the county who served in the 1990s. Since cutting one military worker on net creates about 4 new veterans, the probability of new veterans staying in the county where they served is as low as 5% (0.2/4). I find that the number of older veterans decrease as the military pulls out. On net, I find that the total number of veterans do not increase: the estimate is small and statistically insignificant.

⁴⁹The correlation coefficient is around 0.2. The bulk of the military contracts are for weaponry and equipment. The awardees of these contracts are usually large manufacturing companies located in industrial clusters.

to companies located in county c between 1986 and 1988.⁵⁰ $Pop_{c,1980}$ is the county population in 1980. I plug this term in equation 2.30 and instrument $\Delta proc_c$ with the Bartik instrument

$$\Delta proc_c^{IV} = \frac{Proc_{c,83-85}}{Pop_{c,1980}} \times \frac{Proc_{98-00} - Proc_{86-88}}{Proc_{86-88}},$$

where $Proc_{86-88}$ and $Proc_{98-00}$ are the total amounts of procurement awarded at the beginning and the end of the period, each of which is averaged across three consecutive years.

Table 2.6 reports the results. First note that including procurement contractions does not change the estimates on personnel contractions. The two components of post-Reagan military cuts, though connected at the national level, have separate effects on local economies. Angrist-Pischke partial F-statistics are reported for each endogenous variable, and the instruments are strong throughout. Procurement contractions have expected effects, although the coefficients are not always precisely estimated. Most importantly, the migration response is similar for both personnel contractions and procurement contractions. For the former, for each job loss in the civilian sector there are 2.4 civilians leaving the county (3.2/1.3); for the latter, this ratio is about 2.1. This is evidence that the large response in out-migration is not due to the special characteristics of military personnel contractions.

2.8 Conclusions

This chapter studies the local economic impacts of military personnel contractions in the United States between 1988 and 2000. The contractions had sizable effects on the levels of economic activities in counties with historical military presence. Cutting each military worker resulted in the loss of an additional 1.3 civilian jobs and 0.1 private business establishments, most of which were in the non-tradable sectors and from small businesses. However, local economies adjusted quickly. By year 2000, 2.4 civilian residents had left the county for each civilian job loss. As a result, declines in local wages were small relative to those in rental prices. Quantifying

⁵⁰The reason for using a 3-year average is that procurement is a flow variable and is bumpy from year to year. In contrast, the number of military personnel is a stock variable. Changes in military procurement from 1986-1988 compared to 1998-2000 capture the gradual contraction in procurement in the 1990s.

the welfare impacts based on a simple spatial equilibrium framework, I show that negative welfare shocks to workers due to lower local labor demand were largely compensated by substantive declines in local cost of living, while landowners suffer from a large decline in total rents.

Many local economies rely heavily on military spending. There are increasing concerns in these communities in a period of substantial cuts in military spending: the budget sequestration of 2013 planned for a 42 billion dollar cut in military expenditure. In a more recent military budget plan, the former Secretary of Defense Chuck Hagel proposed to cut the Army to pre-WWII levels.⁵¹ This chapter, by exploiting a similar policy in the past, finds that these cuts have distributional effects. On the one hand, the labor market quickly adjusts via labor migration, and the welfare of workers is not affected very much. On the other hand, landowners suffer a large decline in rental revenues as local demand for housing drops and housing stocks are slow to adjust. I also find no effects on government revenue or expenditure, or on Congress incumbents' probability of being re-elected.⁵²

Policies targeting particular places in order to promote local economies are popular across the world. State and local governments in the United States spend billions of dollars every year trying to attract new businesses or retain existing ones through various incentives (Story, 2012; Story et al., 2012). At the national level, these policies are only justifiable by agglomeration effects in production or frictions in labor mobility, such that the gains from locating in a larger local economy is greater than the gains if the business is located somewhere else. This chapter does not explicitly test the effects of the military operations on the productivity of nearby firms, but it finds small frictions in migration across local economies. The effects of these policies are likely to be consolidated in local land prices instead of being reaped by workers.

This chapter makes a few simplifying assumptions. In reality, some households contain both workers and landowners. It would be interesting to see the heterogeneous effects on renters and homeowners. Renters have fewer constraints to move, but they also incur a smaller welfare shock if they choose to stay. Another assumption I made is that workers are interchangeable across sectors. The wealth effects due to changes in housing values and the tradeoff between space mismatch

⁵¹The Washington Post, Feb 24, 2014. <http://www.washingtontimes.com/news/2014/feb/24/chuck-hagel-shrink-army-pre-wwii-levels-report/>

⁵²Results available upon request.

and sector mismatch may both affect households' migration decisions, which in turn have implications for impacts on wages and welfare. The simple model introduced in this chapter can be extended to relax these assumptions. Finding individual data on home ownership, industry, and migration, and empirically testing the predictions of the enriched model would be interesting future work.

Empirically, I demonstrate that the validity of the widely used shift-share instruments hinges on the assumption that past levels of the treatment variable are not correlated with unobservable secular trends in the outcome variable. Because the location choices of many economic activities are endogenous and can be correlated with expected economic trajectories, this assumption can be violated. To solve this problem, I develop a novel two-step identification strategy combining the synthetic control and the two-stage least square methods. Using information from a pre-treatment period, the synthetic control approach constructs a counterfactual and purges out the unobservable secular trends, conditional on which the shift-share instrument is valid. This approach has general applications in cases where parallel trends are not guaranteed and the instrument is only conditionally valid.

2.9 Appendices

2.9.1 Appendix A: Sources and Construction of Variables

Key explanatory variable

$mil_{ct} = Mil_{ct}/Pop_{c1980}$ is the military personnel in county c in year t as a ratio to the county's population in 1980. $NtlMil_t$ is the total number of military personnel in the United States (not including those on overseas military bases) in year t . $\Delta mil_c = mil_{c,2000} - mil_{c,1988}$ is the long difference between 1988 and 2000. $\Delta mil_{ct} = mil_{ct} - mil_{ct-1}$ is one-year difference.

Outcome variables

$CivEmp_{ct}/Pop_{c1980}$ is the civilian sector employment in county c in year t divided by the county's population in 1980.

$CivInc_{ct}/Pop_{c1980}$ is the civilian sector earnings (in thousand dollars, 2000 constant dollar) in county c in year t divided the county's population in 1980.

$Wage_{ct}$ is the demographic-adjusted average civilian sector wage in county c in year t . It is calculated as follows. First, average wage per job is calculated as $CivInc_{ct}/CivEmp_{ct}$. It is then regressed on a vector of county-level demographic characteristics. The demographic-adjusted average wage is the residual from the regression.

$CivPop_{ct}/Pop_{c1980}$ is the civilian population in county c in year t as a ratio of the county's population in 1980. Civilian population is total population minus military workers.

$Estab_{ct}/Pop_{c1980}$ is the number of private business establishments in county c in year t divided by the county's population in 1980. Private business establishments are from County Business Patterns.

$urate_{ct}$, the unemployment rate in county c in year t , is from Local Area Unemployment Statistics reported by the Bureau of Labor Statistics. County unemployment rates are model-based and are only available since 1990.

$CivEmp_{ct}/CivPop_{ct}$ is the civilian employment to civilian population ratio in county c in year t .

$MedRent_{ct}$ is the median rental price in county c in year t , and is from county-level tabulations in decennial censuses in 1980, 1990, and 2000.

Other variables

Other county characteristics are from various issues of County Data Books

as well as population censuses. These variables include share of white, share of black, share of women, share of population aged between 25-64, share of college graduates, share of high school graduates, share of high school dropouts, share of urban population, county geographic and climatic characteristics, such as population density, average temperature in January, average annual precipitation and snowfall, census regions, etc.

In order to describe the demographic characteristics of the military personnel, including educational attainment, marital status, residence status (live on or off base), characteristics of military spouses (living arrangement, working status, etc), and number of dependents, I use publicly available 5% individual level data of decennial censuses in 1980, 1990, and 2000 from IPUMS. IPUMS data report MSA and state but do not have county identifier. I impute county-level military demographic characteristics by MSA, or by state when the county is not part of any MSA.

Military procurement is from Department of Defense Form DD350, available from 1966 to 2006. DD350 includes all DoD procurement contracts more than 25,000 dollars awarded to the primary contractor. I aggregate the procurement at the county-year level.

2.9.2 Appendix B: Details of the Model

B.1: Household Problem with Unemployment

In this appendix I extend the household problem in section 2.3.2 to allow for (voluntary) unemployment.⁵³ Each household now makes a discrete choice of labor market participation. d_{ic} is a dummy variable which is equal to 1 if the household chooses to work, 0 otherwise. Workers have the same productivity if they choose to work. If household i chooses to work, it earns a local wage \tilde{w}_c but derives no utility from leisure; if it chooses not to work, it receives a pecuniary benefit which is standardized to 0, but enjoys a utility from leisure, denoted as l_i .⁵⁴ Equally productive workers facing the same labor market conditions have different job market participation decisions due to idiosyncratic preference for leisure. Household i 's problem becomes:

⁵³See also Busso et al. (2013) for a similar treatment of unemployment. Kline and Moretti (2013) introduces involuntary unemployment in the model with matching frictions.

⁵⁴ l_i can also be interpreted as distaste for work.

$$\begin{aligned} \max_{h_{ic}, X_{ic}^N, X_{ic}^T, d_{ic}} \quad & u_{ic} = \ln A_c + \alpha \ln h_{ic} + \beta \ln X_{ic}^N + (1 - \alpha - \beta) \ln X_{ic}^T + l_i \cdot (1 - d_{ic}) + e_{ic}, \\ \text{s.t.}, \quad & r_c h_{ic} + p_c X_{ic}^N + p_T X_{ic}^T = d_{ic} \tilde{w}_c \end{aligned}$$

I assume the distribution of l_i to be independently distributed across locations and is uncorrelated with locational characteristics. I assume that the cumulative density function (c.d.f.) is $L(\cdot)$. The problem for a non-working household and for a working household can be solved in the same way as the problem in section 2.3.2. Denote $(h_{ic}^w, X_{ic}^{Nw}, X_{ic}^{Tw})$ as the optimal consumption bundle for a working household and $(h_{ic}^n, X_{ic}^{Nn}, X_{ic}^{Tn})$ as the optimal consumption bundle for a non-working household. All these choices are functions of a vector of local prices and model parameters. $u_c^l = \ln A_c + \alpha \ln h_{ic}^l + \beta \ln X_{ic}^{Nl} + (1 - \alpha - \beta) \ln X_{ic}^{Tl}$, $l \in \{w, n\}$ is thus the real wage a household gets if it chooses to work, or not. The decision for whether to work is:

$$d_{ic} = \begin{cases} 0, & \text{if } u_c^w < u_c^n + l_i \\ 1, & \text{if } u_c^w \geq u_c^n + l_i. \end{cases}$$

Denote the marginal household with idiosyncratic locational preference for leisure l_i^* which is indifferent between working or not working, this household has utility $u_c^w = u_c^n + l_i^*$. Therefore, households with $l_i < l_i^*$ choose to work, while households with $l_i > l_i^*$ choose not to work. The unemployment rate in location c is

$$urate_c = 1 - L^{-1}(l_i^*).$$

Since l_i^* is a function of local prices $\{\tilde{w}_c, r_c, p_c\}$, which are jointly determined by the local economic equilibrium, the unemployment rates are different across locations. However, the unemployment rate is determined by other local economic conditions. Having unemployment in the model in this way does not alter the intuition or mechanism of the local economic equilibrium model introduced in in section 2.3.

B.2: Local Labor Supply

The population size in location c is captured by the preference of the marginal household, e_{ic}^* , in the distribution of e_{ic} . Assume the total population nationwide is 1. Following the standard result of a type-I extreme value distribution, population size in location c can be written as:

$$N_c = \frac{\exp(u_c/\sigma^W)}{\sum_{c'} \exp(u_{c'}/\sigma^W)}.$$

Take log on both sides, we have equation 2.2.

B.3: The Tradable Sector

From the firm's profit maximizing problem in equation 2.8:

$$\max_{N_c^T, K_c^T} \pi_j = p_j x_j - w_c N_c^T - \rho K_c^T,$$

plug in the expression for p_j , and the production function, we have

$$\begin{aligned} \pi_j &= (I_j)^{1/\sigma^T} x_j^{(\sigma^T-1)/\sigma^T} - w_c N_c^T - \rho K_c^T \\ &= (I_j)^{1/\sigma^T} (B_c^T (N_c^T)^{h_T} (K_c^T)^{h_T})^{(\sigma^T-1)/\sigma^T} - w_c N_c^T - \rho K_c^T \end{aligned}$$

First order conditions with regard to N_c^T and K_c^T are:

$$\begin{aligned} \frac{\partial \pi_j}{\partial N_c^T} &= 0 \\ \Rightarrow (I_j)^{1/\sigma^T} \frac{\sigma^T-1}{\sigma^T} x_j^{-1/\sigma^T} h_T B_{Tc} \left(\frac{K_c^T}{N_c^T}\right)^{1-h_T} &= w_c \end{aligned} \quad (2.34)$$

$$\begin{aligned} \frac{\partial \pi_j}{\partial K_c^T} &= 0 \\ \Rightarrow (I_j)^{1/\sigma^T} \frac{\sigma^T-1}{\sigma^T} x_j^{-1/\sigma^T} (1-h_T) B_{Tc} \left(\frac{K_c^T}{N_c^T}\right)^{-h_T} &= \rho \end{aligned} \quad (2.35)$$

Stacking equation 2.34 and equation 2.35, we have the marginal rate of substitution between the two factors:

$$\frac{K_c^T}{N_c^T} = \frac{w_c (1 - h_T)}{\rho h_T} \quad (2.36)$$

Plug equation 2.36 back into equation 2.34, we derive the expression describing the demand for labor in the tradable sector as a function of factor prices and model parameters:

$$\begin{aligned} h_T (I_j)^{1/\sigma^T} \frac{\sigma^T - 1}{\sigma^T} B_{Tc}^{(\sigma^T - 1)/\sigma^T} \left(\frac{1 - h_T}{h_T}\right)^{(\sigma^T - 1)(1 - h_T)/\sigma^T} \rho^{-(\sigma^T - 1)(1 - h_T)/\sigma^T} w_c^{(\sigma^T - 1)(1 - h_T)/\sigma^T - 1} \\ = (N_c^T)^{h_T/\sigma^T} \end{aligned}$$

Take logs on both sides yields equation 2.9

$$\ln N_c^T = [(1 - h_T)(\sigma^T - 1) - \sigma^T] \ln w_c + a_{TL},$$

where

$$\begin{aligned} a_{TL} = & \sigma^T \ln h_T + \ln(I_j) + \sigma^T \ln\left(\frac{\sigma^T - 1}{\sigma^T}\right) + (\sigma^T - 1) \ln B_{Tc} \\ & + (\sigma^T - 1)(1 - h_T) [\ln(1 - h_T) - \ln(h_T)] - (1 - h_T)(\sigma^T - 1) \ln \rho \end{aligned}$$

Notice that $[(1 - h_T)(\sigma^T - 1) - \sigma^T] < 0$. The demand for labor in the tradable sector decreases with local wage.

Similarly, plug equation 2.36 back into equation 2.35, we derive the expression for demand for capital in the tradable sector as a function of factor prices and model parameters:

$$\begin{aligned} (1 - h_T)^{1 + h_T(\sigma^T - 1)/\sigma^T} h_T^{h_T(\sigma^T - 1)/\sigma^T} \rho^{-[1 + h_T(\sigma^T - 1)/\sigma^T]} \left(\frac{\sigma^T - 1}{\sigma^T}\right) B_{Tc}^{(\sigma^T - 1)/\sigma^T} w_c^{[h - (\sigma^T - 1)]/\sigma^T} \\ = (K_c^T)^{1/\sigma^T} \end{aligned}$$

Take logs on both sides

$$\ln K_c^T = -(\sigma^T - 1)h_T \ln w_c + a_{TK}, \quad (2.37)$$

where

$$\begin{aligned} a_{TK} = & [(\sigma^T - 1)h_T - \sigma^T] \ln \rho + \ln I_j + \sigma^T \ln \frac{\sigma^T - 1}{\sigma^T} - \ln B_{Tc} \\ & + h_T \ln \left(\frac{1 - h_T}{h_T} \right) + \sigma^T \ln(1 - h_T) + \sigma^T \ln B_{Tc} + \sigma^T h_T \ln \frac{1 - h_T}{h_T}. \end{aligned}$$

B.4: The Non-Tradable Sector

In order to derive the labor demand function for the non-tradable sector, we start from the non-tradable firm's profit function in equation 2.12:

$$\max_{N_c^N, K_c^N} \pi_c^N = p_c B_{Nc} (N_c^N)^{h_N} (K_c^N)^{1-h_N} - w_c N_c^N - \rho K_c^N$$

First order conditions with regard to N_c^N and K_c^N are:

$$\begin{aligned} \frac{\partial \pi_c^N}{\partial N_c^N} &= 0 \\ \Rightarrow p_c B_{Nc} h_N \left(\frac{K_c^N}{N_c^N} \right)^{1-h_N} &= w_c \end{aligned} \quad (2.38)$$

$$\begin{aligned} \frac{\partial \pi_c^N}{\partial K_c^N} &= 0 \\ \Rightarrow p_c B_{Nc} (1 - h_N) \left(\frac{K_c^N}{N_c^N} \right)^{-h_N} &= \rho \end{aligned} \quad (2.39)$$

Stacking equation 2.38 and equation 2.39, we have the marginal rate of substitution (MRS) between the two factors:

$$\frac{K_c^N}{N_c^N} = \frac{w_c}{\rho} \frac{1 - h_N}{h_N} \quad (2.40)$$

Plug the MRS condition back into the production function of local non-tradable goods, we have the supply of local non-tradable goods as a function of local prices and model parameters

$$\begin{aligned} X_c^N &= B_{Nc} N_c^N \left(\frac{K_c^N}{N_c^N} \right)^{1-h_N} \\ &= B_{Nc} N_c^N \left(\frac{w_c}{\rho} \frac{1-h_N}{h_N} \right)^{1-h_N}. \end{aligned}$$

Take logs on both sides and we get equation 2.13, where $a_{NX} = \ln B_{Nc} + (1 - h_N)[\ln(1 - h_N) - \ln(h_N)]$.

Taking logs on both sides of equation 2.14 and combining it with equation 2.13 yields equation 2.15:

$$\ln N_c^N - \ln N_c = h_N \ln w_c + \ln m_c^H - \ln p_c + a_{NL},$$

where $a_{NL} = \ln \beta - \ln B_{Nc} - (1 - h_N)[\ln(1 - h_N) - \ln \rho - \ln h_N]$.

The implications for the zero-profit condition for firms in the non-tradable sector can be derived as follows. Plugging equation 2.40 into equation 2.12 and requiring the maximized profit to be zero yields:

$$\begin{aligned} \pi_c^N &= N_c^N \left[p_c B_{Nc} \left(\frac{1-h_N}{h_N} \right)^{1-h_N} - w_c - w_c \frac{1-h_N}{h_N} \right] = 0 \\ \Rightarrow h_N \ln w_c &= \ln p_c + b_N, \end{aligned}$$

where $b_N = \ln B_{Nc} + (1 - h_N) \ln(1 - h_N) + h_N \ln h_N - (1 - h_N) \ln \rho$.

2.9.3 Appendix C: Two-Step Estimation of Panel Model with Endogenous Treatment

C1. Data Generating Process

Consider the following data-generating process

$$y_{it} = \beta_0 + \beta_1 X_{it} + f(\lambda_i, t) + \varepsilon_{it}, \quad (2.41)$$

where i is the index for units and t is the index for time periods, $t \in [1, T]$. X_{it} is the treatment of interest. $f(\cdot)$ is an unknown function of fixed effect λ_i and t . λ_i is unknown to the econometrician.

There is an observable predetermined status $D_i = \{0, 1\}$ which is a function of λ_i and v_i , $D_i = g(\lambda_i, v_i)$. v_i is *iid* and uncorrelated with ε_{it} , $E[v_i, \varepsilon_{it}] = 0$.

X_{it} depends on D_i and is correlated with unobserved determinants of the outcome, and is potentially measured with error:

$$X_{it} = \begin{cases} 0 & , \text{ if } D_i = 0 \text{ or } t < t_s \\ \rho_1 \tilde{X}_{it-1} + e_{it} + h(\lambda_i, \varepsilon_{it}, t) + \epsilon_{it} & , \text{ if } D_i = 1 \text{ and } t \geq t_s \end{cases}$$

The specification of X_{it} requires some clarification. First, $\rho_1 > 0$ means that X_{it} exhibits positive serial correlation. e_{it} is classical measurement error with *iid* distribution and mean zero. \tilde{X}_{it} is the deterministic part of X_{it} , $E(\tilde{X}_{it}, \varepsilon_{it}) = 0$. $h(\lambda_i, \varepsilon_{it}, t)$ specifies the endogeneity structure of X_{it} . ϵ_{it} is an *iid* error term, $E[\epsilon_{it}, \varepsilon_{it}] = 0$, $E[\epsilon_{it}, \lambda_i] = 0$, $E[\epsilon_{it}, \tilde{X}_{it}] = 0$. Endogeneity of the treatment, X_{it} , comes from three sources. First, it conditions on D_i , which is determined by λ_i . Second, it is correlated with the unobserved fixed effects λ_i . Third, it is also correlated with the contemporaneous error term ε_{it} .

X_{it}^{IV} is a candidate instrumental variable for X_{it} .

$$X_{it}^{IV} = \begin{cases} 0 & , \text{ if } D_i = 0 \text{ or } t < t_s \\ \rho_2 \tilde{X}_{it-1} + \tilde{e}_{it} + \tilde{\epsilon}_{it} & , \text{ if } D_i = 1 \text{ and } t \geq t_s \end{cases}$$

\tilde{e}_{it} is also classic measurement error. X_{it}^{IV} is uncorrelated with ε_{it} , $E(X_{it}^{IV}, \varepsilon_{it}) = 0$. But it is conditional on D_i , in other words, $E(X_{it}^{IV}, \lambda_i) \neq 0$. Thus the validity of X_{it}^{IV} as an IV hinges on whether $f(\lambda_i, t)$ can be purged out from equation 2.41.

Consistency of the conventional panel model hinges on the correctly specifying $f(\cdot)$. Consider the simplest case where $f(\lambda_i, t) = \lambda_i$, the data generating process reduces to:

$$y_{it} = \beta_0 + \beta_1 X_{it} + \lambda_i + \varepsilon_{it}.$$

In a panel fixed effect model, the equation is first demeaned:

$$(y_{it} - \bar{y}_i) = \beta_1(X_{it} - \bar{X}_i) + \varepsilon_{it},$$

where the condition $\bar{\varepsilon}_i = 0$ is used. $(X_{it}^{IV} - \bar{X}_i^{IV})$ is thus a valid IV for $(X_{it} - \bar{X}_i)$ in the demeaned equation, as $E[(X_{it}^{IV} - \bar{X}_i^{IV}), \varepsilon_{it}] = 0$. But in general, if $f(\cdot)$ is not a constant, demeaning does not cancel out λ_i and X_{it}^{IV} is invalid.

We construct a synthetic control for each unit i with $D_i = 1$ from a the “donor pool” of units with $D_i = 0$ such that the distance between the trajectory of y_{it} in $t \in [1, t_s]$ and that of its synthetic control y_{it}^{synth} is minimized by solving the following problem:

$$\min_{\alpha_{i_j}} \|\mathbf{Z}_{it} - \sum_j \alpha_{i_j} \mathbf{Z}_{jt}\|,$$

where $\alpha_{i_j} \in [0, 1]$ if $D_j = 0$, and $\sum_j \alpha_{i_j} = 1$. $\alpha_{i_i} = 1$ if $D_i = 1$. $\mathbf{Z}_{it} = \{y_{i\tau}, \tau \in [1, s]\}$ is a set of past realizations of the outcome variable before the treatment takes place in period s . According to the data generating process as specified in equation 2.41, under the optimal weights α_{i_j} , the distance $\|f(\lambda_i, t) - \sum_j \alpha_{i_j} f(\lambda_j, t)\|$ is also minimized.

If the matching is perfect, $f(\lambda_i, t) - \sum_j \alpha_{i_j} f(\lambda_j, t) = 0$. Counties in the same synthetic control group ($D_i = 1$ and $D_j = 0$) with appropriate weights can be written as

$$\begin{aligned} y_{it} &= \beta_0 + \beta_1 X_{it} + f(\lambda_i, t) + \varepsilon_{it} \\ \alpha_{i_j} y_{jt} &= \alpha_{i_j} [\beta_0 + \beta_1 X_{jt} + f(\lambda_j, t) + \varepsilon_{jt}] \end{aligned}$$

We can purge the secular trend by taking difference between the treated unit and its synthetic control.

$$y_{it} - \sum_j \alpha_{i_j} y_{jt} = \beta_1(X_{it} - \sum_j \alpha_{i_j} X_{jt}) + (\varepsilon_{it} - \sum_j \alpha_{i_j} \varepsilon_{jt})$$

We can estimate the following transformation of equation 2.41:

$$y_{hgt} = \beta_0 + \beta_1 X_{hgt} + \omega_{hgt} + \varepsilon_{hgt}, \quad (2.42)$$

where county h can be a treated unit i or an untreated unit j , group g includes a unit i with $D_i = 1$ and units j with $D_j = 0$. ω_{igt} is a dummy variable indicating a group-time period. Each equation is weighted by α_h . Demeaning within the group-time period, we have the following

$$(y_{hgt} - \frac{1}{2} \sum_h \alpha_h y_{hgt}) = \beta_1 (X_{hgt} - \frac{1}{2} \sum_h \alpha_h X_{hgt}) + (\varepsilon_{hgt} - \frac{1}{2} \sum_h \alpha_h \varepsilon_{hgt}).$$

Thus $(X_{hgt}^{IV} - (1/2) \sum_h \alpha_h X_{hgt}^{IV})$ is a valid instrumental variable for $(X_{hgt} - (1/2) \sum_h \alpha_h X_{hgt})$ because

$$E[(X_{hgt}^{IV} - (1/2) \sum_h \alpha_h X_{hgt}^{IV}), (X_{hgt} - (1/2) \sum_h \alpha_h X_{hgt})]$$

To sum up, we have a case here where the instrument is only conditionally valid. If the unobserved heterogeneity affects the outcome in the sample period, our instrument is only valid when conditional on the potentially time-varying effects of the unobserved heterogeneity. When the functional form of the time-varying heterogeneous effect is unknown, the synthetic control approach provides a non-parametric way to partial out the effects of the heterogeneous effect.

C2. Analogue

I have a sample of counties (i and j) between 1980 and 2000 (t). Military cuts (X_{it}) started in 1988 (t_s). Only counties with military bases (D_i) are affected. Military bases have been in a county since many periods earlier. Whether having military bases (D_i) and economic outcome (y_{it}) are both determined by some unobservable characteristics of the county (λ_i). The effect of λ_i on the outcome variable can also vary with time in an unknown functional form ($f(\lambda_i, t)$). The size of the military cut in a particular year is correlated with the time-invariant county characteristics and potentially time-varying idiosyncratic shock ($h(\lambda_i, \varepsilon_{it}, t)$), and is measured with error (e_{it}).

C3. Simulation

I simulate the data and demonstrate that the two-step procedure proposed above reduces estimation bias without knowing the functional form of $f(\lambda_i, t)$, while

the bias of the standard panel data approach relies on correctly specifying $f(\lambda_i, t)$.

I assume that the unobservable time-invariant county characteristics λ_i is drawn from a Poisson distribution with mean 5 in a sample of $N = 200$. Define $D^* = \lambda_i + v_i$, where $v_i \sim N(0, 5)$. I then sort D^* from the largest to the smallest and assign $D_i = 1$ for the 50 units with the largest D^* . v_i is drawn from a distribution with relatively large standard error such that the group with $D_i = 1$ and the group with $D_i = 0$ have enough overlapping in λ_i . As in a conventional matching approach, overlapping is crucial for the validity of the estimator.

I draw $T = 30$ periods. The treatment period starts since $t_s = 11$. Equation 2.41 is parameterized as follows: $\varepsilon_{it} \stackrel{i.i.d.}{\sim} N(0, 2)$, $\beta_0 = 3$, $\beta_1 = 1$. $f(\lambda_i, t) = \gamma_1 \lambda_i + \gamma_2 \lambda_i \cdot t$. The endogenous variable, X_{it} , is parameterized as follows:

$$X_{it} = \begin{cases} 0 & , \text{ if } D_i = 0 \text{ or } t < t_s \\ \rho_1 \tilde{X}_{it-1} + d_e \cdot e_{it} + h(\lambda_i, \varepsilon_{it}, t) + \epsilon_{it} & , \text{ if } D_i = 1 \text{ and } t \geq t_s \end{cases}$$

$e_{it} \stackrel{i.i.d.}{\sim} N(0, 5)$, ϵ_{it} is from an extreme value distribution. $X_{i1} = 0$. $\rho_1 = 0.95$. Therefore, the shock is most likely to be centered around zero, but there are chances when a large shock takes place, it tends to persist. The high serial correlation captures the fact that annual cuts during the Post-Reagan military personnel contractions were rather smooth. The endogeneity structure of the treatment is specified as $h(\lambda_i, \varepsilon_{it}, t) = d_\lambda \cdot \rho_\lambda \cdot t + d_\varepsilon \cdot \varepsilon_{it}$. d_λ , d_ε , d_e are dummy variables that turn on and off endogeneity due to, respectively, unobservable local characteristics, unobservable contemporaneous shocks, and measurement error.

C4. Simulation Results

Results using different estimating approaches when d_λ , d_ε , d_e take different values are reported in table 2.7. Columns in panel A are estimated using Two-Stage Least Squares (2SLS) estimator in a conventional panel model with unit dummies (hereafter called panel FE model). Columns in panel B are estimated using a two-step approach which involves generating groups and sample weights using the synthetic control approach in the first step and estimate using weighted 2SLS in the second step. Column 1 shows the case with no endogeneity, in which case both panel FE and the synthetic matching methods give consistent results, and the estimated coefficients are close to the true value. When there is secular trend based on

unobservable characteristics, panel FE estimate is biased due to mis-specification (panel A, column 2), and estimating the model using 2SLS with the instrument does not make the estimation correct (panel A, column 3). In contrast, the two-step approach, no matter whether the model is estimated using weighted least squares or weighted 2SLS, gives estimates close to the true value (panel B, column 2 and column 3). Column 4 and column 5 estimate a model with both unobservable secular trend and endogenous contemporaneous shocks. Panel FE models, either estimated using OLS or 2SLS, are biased. For the synthetic control approach, the weighted least square estimate is also biased (panel B, column 4), but the two-step weighted 2SLS estimate is close to the true parameter (panel B, column 5). Column 6 and column 7 report results when the endogeneity comes from the unobservable county characteristics, column 8 and column 9 report results when the endogeneity comes from both sources, panel FE models all have results far off from the true value, while the two-step weighted 2SLS estimator obtains results that are similar to the true value.

2.9.4 Appendix D: Statistical Inferences for the Two-Step Estimation

The two-step estimation involves the weights and group dummies constructed from the first step to be used in the second step estimation. The standard errors for the second step estimation do not take into account of the uncertainty in matching in the first step. I use bootstrap to conduct statistical inference. In each bootstrap, a pseudo sample is drawn and the two-step procedure, including constructing the synthetic groups with weights and the subsequent estimation, is carried out. Abadie and Imbens (2006) and Abadie and Imbens (2008) note that bootstrap does not work in simple nearest neighbor matching with fixed number of matches since the bootstrapped samples do not achieve the asymptotic distribution of the population. They postulate that the bootstrapping in matching estimates that allows a smoother reweighing function, which include the two-step approach introduced here, are likely to remain valid.

I simulate the data using parameters specified table 2.7, column 2, specifically,

$$y_{ht} = 3 + \beta_1 X_{ht} + (\lambda_h + 0.1 \cdot \lambda_h \cdot t) + \varepsilon_{ht},$$

$$X_{ht} = \begin{cases} 0 & , \text{ if } D_h = 0 \text{ or } t < t_s \\ 0.95 \cdot X_{ht-1} + \epsilon_{it} & , \text{ if } D_h = 1 \text{ and } t \geq t_s , \\ 0 & , \text{ if } t = 0 \end{cases}$$

$$D = \mathbf{1}\{D^* = \lambda_h + v_h \geq \bar{d}\},$$

where $\epsilon_{ht} \stackrel{i.i.d.}{\sim} N(0, 2)$, $v_h \sim N(0, 5)$, ϵ_{ht} is from an extreme value distribution. The sample size is $N = 200$. \bar{d} is a fixed value such that there are 50 units with $D_h = 1$. $\beta_1 = 0$. In each simulation, ϵ_{ht} and v_h are simulated and other variables are generated according to the data generating process.

Since X_{ht} is not correlated with ϵ_{ht} , I estimate the model using weighted least squares as in table 2.7, panel B, column 2. In each bootstrap, I redraw with replacement units with $D_h = 0$ and $D_h = 1$ separately such that the pseudo samples have the same number of units as in the original samples. Then I conduct synthetic control and construct the groups (w_{hg}) and weights (α_{hg}). Finally, I estimate the following regression using weighted least squares:

$$y_{hgt} = \beta_0 + \beta_1 X_{hgt} + \omega_{gt} + \epsilon_{hgt} .$$

I conduct three sets of statistical inference based on pairwise bootstrapping. The first is the conventional t -statistics with standard errors clustered at the group and unit level. The second approach involves using the standard deviation of the bootstrapped $\hat{\beta}_1$ as the bootstrapped standard error, then calculating the Wald statistic and conduct statistical inference. The third approach involves bootstrapping the Wald statistic, a pivotal statistic, so that the bootstrapping procedure also provides asymptotic refinement. Specifically, denote the Wald statistics from the original sample for the null hypothesis $H_0 : \beta_1 = 0$ as $w_0 = \hat{\beta}_1 / se(\hat{\beta}_1)$. Then for each bootstrap, calculate the Wald Statistics $w^b = (\hat{\beta}_1^b - \hat{\beta}_1) / se^b(\hat{\beta}_1)$. I rank w^b 's from the smallest to the largest. Denote the 2nd percentile as w_{p2}^b and the 98th percentile as w_{p98}^b . If $w^b \in [w_{p2}^b, w_{p98}^b]$ then we cannot reject H_0 . If $w^b \notin [w_{p2}^b, w_{p98}^b]$ then we reject H_0 .⁵⁵

⁵⁵The reason for using the 96th confidence interval instead of the conventional 95th is because I only conduct 100 bootstraps.

I use three different estimation models. The first model uses the correctly specified model - in this case panel OLS with county fixed effect and county-specific time trend. The second model uses a mis-specified model, in this case panel OLS with unit fixed effects. The third estimation uses the proposed two-step approach. For each estimation model, I calculate the average bias of the estimation, and the rejection rate at 4% level based on (1) the cluster-robust standard errors from OLS, (2) Wald statistics using $\hat{\beta}_1^0/sd(\hat{\beta}_1^b)$, where $sd(\hat{\beta}_1^b)$ is the standard deviation of estimates from bootstrapped sample, and (3) whether the $w^b \notin [w_{p2}^b, w_{p98}^b]$. Table 2.8 reports the results.

Findings based on table 2.8 are as follows. First, OLS with the incorrect specification is clearly biased and inconsistent. The two-step identification strategy gives estimates that are close to the real value. The degree of bias depends on the effects of the unobserved heterogeneity can be purged, which in turn relies on the match quality. The quality of match relies on the overlapping of the distribution of λ_h from the group of units with $D_h = 1$ and the group of units with $D_h = 0$. As in the simulation exercise, $D_h = g(\lambda_h, v_h)$. v_h is *iid* and uncorrelated with ε_{ht} , $E[v_h, \varepsilon_{ht}] = 0$. Matching quality will be better if the standard deviation of v_h is large comparing with the standard deviation of λ_h , which generates sufficient overlap of λ_h among the treated units and the untreated units. Second, the correctly specified OLS estimate is less biased than the two-step approach based on the synthetic control, but not by much. The standard errors are larger in the synthetic control approach than that in OLS due to many more control variables. Thus the statistical inference based on the Wald statistic from cluster-robust standard errors in the OLS is under-rejected.

2.9.5 Appendix E: Robustness and Heterogeneous Effects

E1. Robustness Checks

Panel model with county dummies

A conventional approach to deal with unit-specific time trend in a differenced equation is to include set of unit dummies. Notice that it is impossible to include county dummies in the baseline model in equation 2.25, as each county has only one observation. However, we can include the pre-treatment period, and estimate the following equation:

$$\Delta y_{cp} = \beta \Delta mil_{cp} + \lambda_c + \tau_p + \Delta \varepsilon_{cp}, \quad (2.43)$$

where each c is a base county. p indicates one of the two periods, the pre-treatment period between 1980 and 1987, and the post-treatment period between 1988 and 2000. λ_c is a set of county dummies and τ_p is a dummy indicating the post-treatment period. The instrument for Δmil_{cp} is

$$\Delta mil_{cp}^{IV} = (Mil_{c,1980}/Pop_{c,1980}) \cdot (NtlMil_{\bar{p}} - NtlMil_{\underline{p}})/NtlMil_{\underline{p}},$$

\underline{p} and \bar{p} indicates the first year and the last year in the period, respectively.

Equation 2.43 allows the secular trends to be county-specific, but assumes that the trends to be linear throughout both periods. The model is identified by the deviation in the changes and the predicted changes in military personnel (IV). Notice that the first component, the pre-existing military presence, is time-invariant, the model is identified by the variation in the second component, the changes in national total military personnel. If the variation of the instrumental variable mainly comes from the differences in pre-existing military presence, the first stage is likely to be weaker in estimating equation 2.43. Table 2.10 shows that it is indeed the case. The first stage F statistics is only about 3. Although the estimated coefficients stay largely the same as in table 2.25, the standard errors are much larger and many of the coefficients are not statistically significant.

Parallel to equation 2.31, I estimate a first-differenced panel model with leads and lags in military cuts in order to study the dynamic effects:

$$\Delta y_{ct} = \beta \Delta mil_{ct} + \lambda_c + \tau_t + \Delta \varepsilon_{ct}, \quad (2.44)$$

where Δ indicates first differences. Similarly, the instrumental variable for Δmil_{ct} is

$$\Delta mil_{ct}^{IV} = (Mil_{c,1980}/Pop_{c,1980}) \cdot (NtlMil_t - NtlMil_{t-1})/NtlMil_{t-1}.$$

The results are reported in table 2.11 exhibit the same patterns: the estimates are smaller and somewhat less precisely estimated as those in table 2.5.

Influential counties

One assumption needed for the validity of the instrumental variable is that the national military contractions were not driven by the unobservable economic shocks in some particular local economies. A hypothetical, though unlikely, example of this case is that some local interest groups regard the existence of military bases as detrimental to their local economies, and they lobby for national military cuts. To rule out this possibility, I identify three sets of counties that are most likely to have the incentive and capability to have influenced the national policies: base counties in the Washington, DC area, counties with the largest military bases, and counties that are from districts that were formerly represented by key decision makers in the BRAC commission.⁵⁶ The sample in panel A of table 2.12 excludes the 27 base counties in the DC area. The sample in panel B excludes the top 5 percent counties with the largest military bases. The sample in panel C excludes 25 counties that potentially have political influences. Overall the results are remarkably similar to the baseline results in table 2.3.

Geographic aggregation

People may work, live, and consume in different counties, the boundaries of local labor markets can be larger than counties. The impacts of military personnel contractions in a larger geographic area can be either larger or smaller than those using counties as local economies. The reason is both economic and statistical. In terms of economics, the impacts are likely to be larger in a larger geographic aggregation if there are spillover effects across county borders. On the other hand, a larger geographic aggregation may internalize local economic impacts, which leads to smaller estimates. For example, in the extreme case, if we define the entire United States as a local labor market, there will be little migration response to demand shocks. Statistically, using larger geographic aggregation may either alleviate or exacerbate the measurement error problem. Aggregation cancels out measurement error, while converting some county-level outcomes into larger geographic aggregations incur additional measurement error. Military contractions at a larger geographic level are also more likely to be endogenous, as a large local economy can influence national policies. Finally, larger aggregations result in a smaller donor pool, the matching bias from the synthetic matching approach is likely to be larger and the two-step identification strategy more likely to lead to inconsistent estimates.

⁵⁶These people include Speaker of the House, House Minority Leader, members of the BRAC commission.

I use 1990 commuting zones as an alternative definition of local economies. The commuting zones are developed by the Department of Agriculture according to commuting patterns. The 741 commuting zones in 1990 cover the entire United States. On average a commuting zone is about four times as large as a county. I call a “base commuting zone” if the commuting zone intersected with at least one military base in 1987, and the remaining commuting zones form the donor pool. Commuting zone level outcomes are either aggregated from county level variables (for employment, income, population, housing units, etc) or are averages using county population as weights (for median rental price, etc). I identify 208 commuting zones with military bases and saw military presence decline during the 1988-2000 period. I then construct a synthetic control for each base commuting zone from the donor pool. The instrument for military contractions in each commuting zone is the a Bartik instrument constructed in the same way as in equation 2.26. Finally I estimate equation 2.30 using the commuting zone sample.

Table 2.13 shows the results using the same specification and outcome variables as in table 2.3. The estimated impacts of military personnel contractions are in many cases much larger than those at the county level. For example, cutting one military worker causes the commuting zone to lose about 2.9 civilian jobs, more than twice the estimate at the county level. The migration response is smaller due to larger geographic area: for each civilian job loss, 1.3 civilians leave the commuting zone. For one percentage point decrease in the military to population ratio, the average wage drops by 0.95 percent, but this is not statistically significant. The rental price drops by 3.7 percent, also much larger than the effect found using the sample of counties.

E2. Heterogeneous Effects

By the size of the cut

So far the results show that although military personnel contractions have sizable effects on the levels of local employment and population, the welfare impacts are small, and local economies adjust quickly. This finding may mask the non-linear effects of negative shocks. In particular, a local economy may be resilient to small shocks, but large shocks can kick off chain effects causing the local economy to end up on a downward spiral.

In this subsection I investigate non-linearity in the treatment effects by dividing the sample by the size of the cut. Panels A through C in table 2.14 reports the

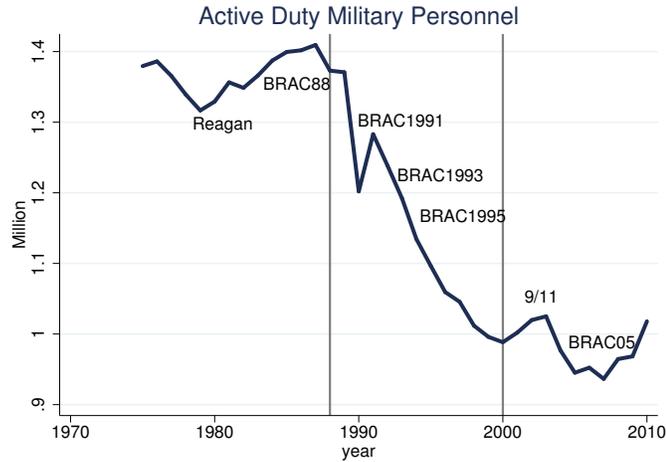
results of replicating table 2.3 using base counties (and their synthetic controls) that experience military personnel contractions above the 25th, 50th, and 75th percentile of the distribution. For base counties that are above the 25th percentile, the average size of military personnel contractions accrue to 1 percent of its 1980 population. For base counties that are above the 75th percentile, that it is 3 percent, or close to 6 percent of its 1980 workforce. It is a substantial shock by any standard. Somewhat surprisingly, all three panels give very similar results, which are also close to those in table 2.3. It seems that the effects of military contractions are close to linear.

By population density

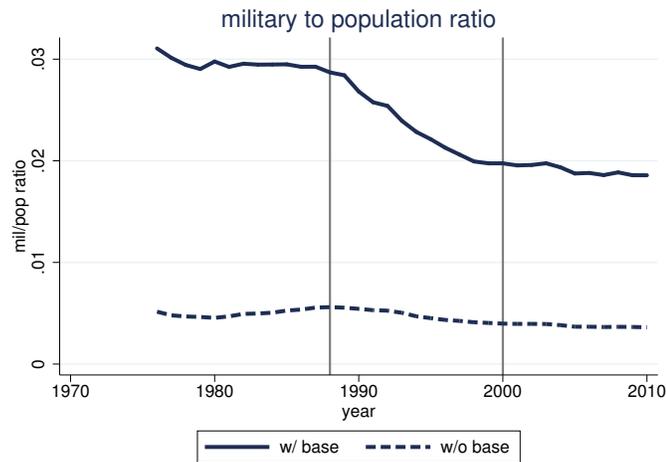
Local economies differ in various dimensions, such that a same degree of military cuts may have different effects across local economies. I investigate the heterogeneous effects by population density of the local economy. I use population density to proxy for the richness of local labor markets in terms of alternative job opportunities. Studies have found that urban local labor markets are more resilient to shocks, we would expect that local labor market outcomes would be less negatively affected in counties with high urban rate.

I divide base counties into two groups, one with the 1987 population density above median and the other below. I repeat the estimates in table 2.3 separately for the two samples. Table 2.15 shows the results. The first stage is equally strong in both sub-samples, so the models are well identified. The impacts on employment and population in more densely populated counties are slightly smaller than that in less densely populated counties (column 1 and column 3), and the impacts on civilian earnings, wage and local businesses are larger in less densely populated counties (column 2, column 4, and column 7). Thus it seems that local labor markets in more densely populated areas weather shocks better. Rental price also drops less in more densely populated counties.

Figure 2.1: Military Personnel Contractions between 1988 and 2000
 A: Change in size of military personnel by county type



B: Change in military to population ratio by county type



Note: Data for figure A are from the annual report of the Distribution of Personnel from the Department of Defense. Active-duty military that deployed in the 50 states and the District of Columbia are included. Figure B shows the change in military to population ratio by whether the county had a major military base in 1987.

Figure 2.2: Illustration of Workers' Welfare

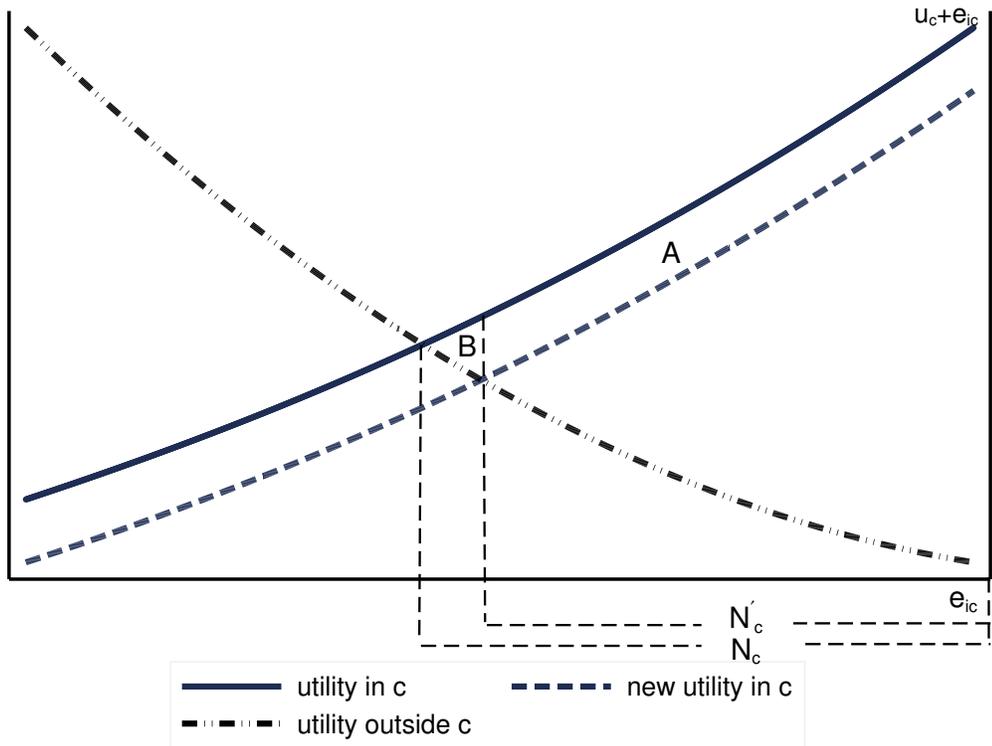
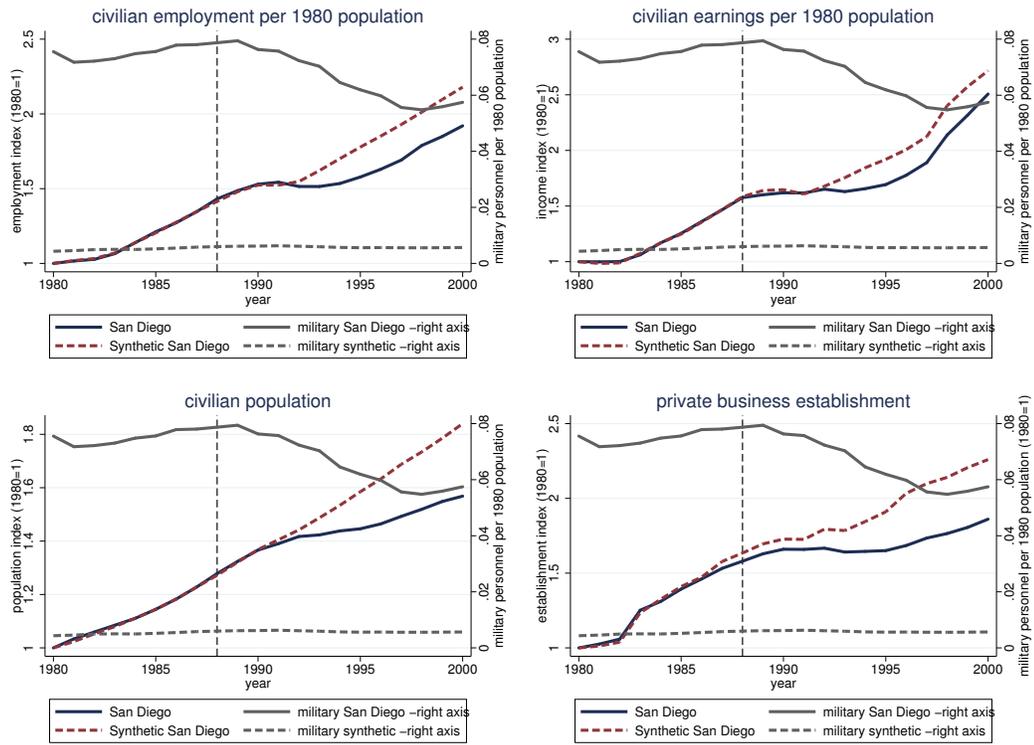


Figure 2.3: San Diego and Synthetic San Diego



Note: This figure shows an example of San Diego County, California and its synthetic control. The first graph shows the trajectory of civilian sector employment. The second figure shows the trajectory of civilian sector earnings. The third figure shows the trajectory of civilian population. The fourth figure shows the trajectory of private sector establishment. All trajectories are standardized at 1 in 1980. In each panel the gray line shows the trajectory of the military personnel to population ratio.

Figure 2.4: Cumulative Effects by Grouping Years

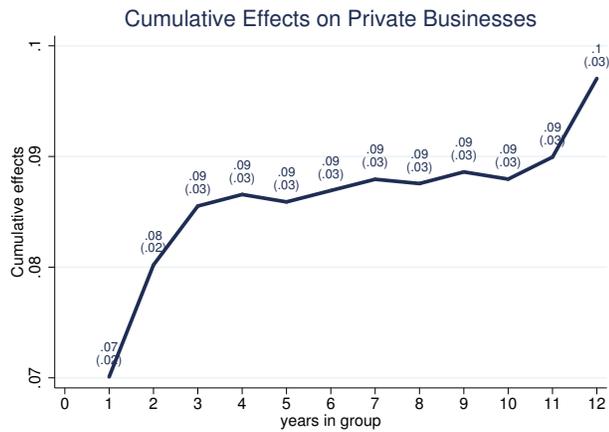
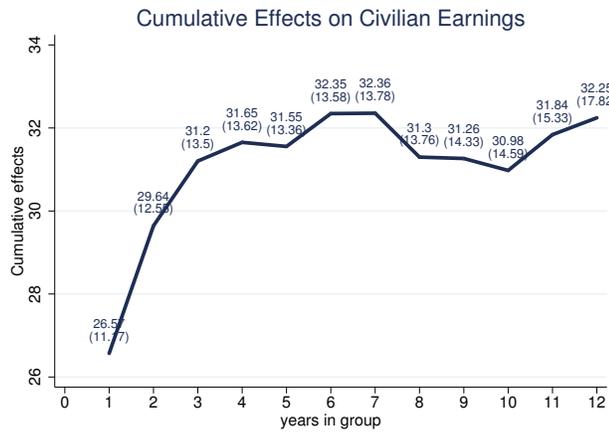
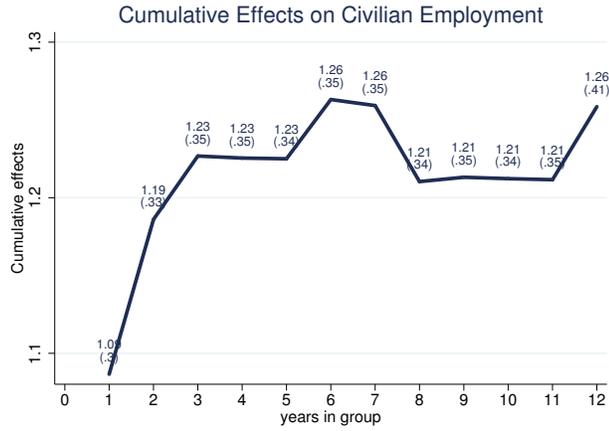


Table 2.1: Summary Statistics

	No bases			<i>p</i> – value		
	Bases (1)	No bases (2)	synth (3)	(1)-(2) (4)	(1)-(3) (5)	(1)-(3) (6)
% change in civilian employment, 1980-1987	.192 (.186)	.109 (.196)	.191 (.204)	0	.909	.671
% change in civilian payroll, 1980-1987	.228 (.234)	.093 (.25)	.226 (.265)	0	.946	.792
% change in civilian population, 1980-1987	.078 (.098)	.019 (.099)	.075 (.109)	0	.827	.391
% change in private establishment, 1980-1987	.340 (.186)	.211 (.099)	.341 (.191)	0	.795	.69
log change in median housing price, 1980-1990	.508 (.262)	.368 (.213)	.505 (.309)	0	.916	.775
log change in median rental price, 1980-1990	.618 (.157)	.552 (.172)	.62 (.202)	0	.902	.763
County-group dummies					N	Y
Number of counties	335	2764	1737	2764	1691	1691
Total weights	335	2764	335	2764	670	670

Note: There are, in total, 2764 counties in the sample that had a civilian population of at least 5000 in 1980. In the first 3 columns, I tabulate average county characteristics for counties with military bases (column 1), counties with no military bases (column 2), and reweighted counties with no military bases on synthetic approach (column 3). Standard deviations are in parentheses in the first 3 columns. Column 4 reports the *p*-value of the mean difference between columns 1 and 2. Column 5 reports the *p*-value of the mean difference between columns 1 and 3. Column 6 reports the *p*-value of the difference between counties with military bases and those without while restricting the comparison to be within each treated county and its synthetic cohorts. Predicted changes in county characteristics between 1988 (1990 for housing market outcomes) and 2000 are obtained by running a set of pooled seemingly unrelated regressions using counties without military bases.

Table 2.2: Military Personnel Contractions and Civilian Employment, 1988-2000

	(1)	(2)	(3)	(4)
Δmil_c	0.793*** (0.225)	1.188*** (0.343)	1.259*** (0.405) {0.338}	1.266*** (0.343) {0.283}
Covariates	X	X		X
N	335	335	19,787	19,787
Total weights	335	335	670	670
Model equation no.	(2.27)	(2.27)	(2.30)	(2.30)
Estimation method	OLS	2SLS	2SLS	2SLS
Weights			X	X
First stage F-statistics		58.634	70.688	68.908

Note: The dependent variable is the change in civilian sector employment between 1988 and 2000, scaled by 1980 population. For weighted regressions, weights constructed from the synthetic control, w_{cj} , are used. For 2SLS estimations, first stage F-statistics are reported. In columns 1 and 2, robust standard errors reported in the parentheses are robust to heteroskedasticity. In columns 3 and 4, standard errors in parentheses are first clustered at the county group level, then clustered at the county level. Significance levels are marked according to the standard errors in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. For columns 3 and 4, bootstrapped standard errors are reported in curly brackets.

Table 2.3: Military Contractions and County Outcomes, 1988-2000

	(1)	(2)	(3)	(4)
	civ emp	civ earning	civ pop	private busi estab
Δmil_c	1.259*** (0.405) {0.338}	32.245* (17.815) {15.243}	3.095*** (0.810) {0.691}	0.097*** (0.032) {0.025}
percent change with $\Delta mil_c = -0.01$	-2.345	-1.796	-2.903	-3.78
	(5)	(6)	(7)	(8)
	occupied housing units	emp/pop	log wage	log median rent
Δmil_c	0.902*** (0.183) {0.162}	-0.263 (0.227) {0.222}	0.468 (0.296) {0.259}	1.336*** (0.192) {0.179}
percent change with $\Delta mil_c = -0.01$	-2.227	0.438	-0.468	-1.336

Note: Each column uses the outcome variable as indicated in the column headline. In columns 1 through 6 and column 8, the outcome variable is long-differenced between 1988 and 2000. In columns 1 through 4, outcome variables are scaled by 1980 population. Percent changes are reported for columns 1 through 4. Percentages are calculated based on the estimated coefficients and 1988 average levels among base counties. Rental and housing prices use contemporaneous prices. Earnings are in thousands of dollars and are denominated in 2000 dollars. See text for details about the additional covariates. All columns are estimated using the weighted 2SLS estimator using the predicted military personnel contractions as the instrumental variable, and weights, w_{cj} , are constructed from the synthetic control, as in Column 3 of Table 2.2. The first-stage F statistic is 70.688. There are 19,787 observations in the sample and 335 county groups. Standard errors are first clustered at the county-group level, then clustered at the county level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors are in curly brackets.

Table 2.4: Changes in Civilian Employment between 1988 and 2000 by Sector

	(1)	(2)	(3)	(4)	(5)
Emp to 1980 pop by sector	federal civilian	manu	construct	retail	service
Δmil_c	0.251*** (0.057)	0.044 (0.085)	0.204*** (0.055)	0.246*** (0.087)	0.158 (0.164)
N	19787	19173	18656	19723	19223
First-stage F statistics	70.688	70.282	62.222	70.623	70.051
Mean dependent variable	0.025	0.075	0.031	0.097	0.143
percent change with $\Delta mil_c = -0.01$	10.4	0.57	6.58	2.54	1.10

Note: All columns are estimated using the weighted 2SLS estimator using the predicted military personnel contractions as the instrumental variable. The number of observations varies across columns because I drop counties that have a left-censored number of employees in each sector. Standard errors in parentheses are first clustered at the county level, then at the county-group level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.5: Dynamic Effects of Military Personnel Contractions

	(1)	(2)	(3)
	emp	earning	estab
$(Mil_{cs} - Mil_{cs-1})/Pop_{c1980}$			
$s = t - 2$	-0.038 (0.244) [23.749]	9.591 (14.323) [23.749]	0.041** (0.019) [23.749]
$s = t - 1$	0.644** (0.277) [28.407]	-0.062 (9.339) [28.407]	0.036** (0.017) [28.407]
$s = t$	0.714** (0.281) [15.009]	15.957 (11.159) [15.009]	0.022 (0.018) [15.009]
$s = t + 1$	0.017 (0.310) [12.253]	-2.427 (9.929) [12.253]	-0.005 (0.017) [12.253]
$s = t + 2$	-0.175 (0.302) [16.411]	10.185 (10.347) [16.411]	-0.010 (0.014) [16.411]
obs	257231	257231	257231

Note: Years in the sample are from 1988 to 2000. I use weighted 2SLS estimates in all specifications, as well as weights from the synthetic controls. I include changes in military personnel up to two years before and after the year in question. Outcome variables are the one-year change of the county attributes as indicated by the shorthand on top of each column. Specifically, the outcome variable is annual change in civilian sector employment per 1980 population in column 1; annual change in civilian sector labor income per 1980 population in column 2; annual change in private business establishments per 1980 population in column 3. Standard errors are first clustered at the county-group by year level, then clustered at the county level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The Angrist-Pischke first-stage partial F-statistic for each endogenous variable is reported in brackets.

Table 2.6: Military Personnel Contractions and Procurement Contractions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	civ emp	civ earning	civ pop	busi estab	log rent	occ houses	log wage	emp pop
Δmil_c	1.317*** (0.405) [72.380]	29.650 (18.818) [72.380]	3.217*** (0.804) [72.380]	0.101*** (0.032) [72.380]	1.374*** (0.189) [72.380]	0.937*** (0.185) [72.380]	0.465 (0.304) [72.380]	-0.240 (0.236) [72.380]
$\Delta proc_c$	0.075 (0.049) [15.557]	-3.295 (3.445) [15.557]	0.156** (0.070) [15.557]	0.005** (0.002) [15.557]	0.048 (0.046) [15.557]	0.042* (0.023) [15.557]	-0.004 (0.046) [15.557]	0.030 (0.023) [15.557]

Note: $\Delta Proc_c$ is the change in levels of military procurement between 1986-1988 and 1998-2000 in county i , scaled by 1980 population. Each column uses the outcome variable as indicated in the column headline. There are 19,787 observations in the regression. All columns are estimated using the weighted 2SLS estimator using the predicted military personnel contractions as the instrumental variable. The Angrist-Pischke partial F-statistic is reported in brackets. Standard errors are reported in parentheses, first clustered at the county level, then clustered at county-group-year level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.7: Simulation Results

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Panel A: Panel FE									
X_{it}	1.012	1.869	1.943	0.106	1.961	-0.380	2.725	-0.228	2.705
	(0.027)	(0.071)	(0.100)	(0.021)	(0.253)	(0.110)	(0.314)	(0.065)	(0.333)
N	4000	4000	4000	4000	4000	4000	4000	4000	4000
$\beta_1 = 1$ (p -value)	0.647	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Panel B: Synth reweighing									
X_{it}	0.984	1.034	1.086	0.539	1.141	0.375	1.090	0.352	1.089
	(0.010)	(0.202)	(0.160)	(0.092)	(0.146)	(0.241)	(0.171)	(0.220)	(0.171)
N	137900	137900	137900	137900	137900	137900	137900	137900	137900
Total weight	2000	2000	2000	2000	2000	2000	2000	2000	2000
$\beta_1 = 1$ (p -value)	0.112	0.868	0.589	0.000	0.334	0.010	0.600	0.004	0.601
γ_1	0	1	1	1	1	1	1	1	1
γ_2	0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
d_e	0	0	0	1	1	0	0	1	1
d_λ	0	0	0	0	0	1	1	1	1
d_ε	0	0	0	0	0	0	0	1	1

Note: Robust standard errors in parentheses, clustered at unit-group and unit levels in Panel B.

Table 2.8: Bias and Rejection Rate by Estimation Model

	emp	earning	pop	estab	occupied houses	emp/pop	ln(wage)	ln(rent)
emp	1							
earning	.821	1						
pop	.698	.476	1					
estab	.798	.682	.786	1				
occupied housing	.663	.516	.696	.635	1			
emp/pop	.164	.043	-.412	-.223	-.128	1		
ln(wage)	.129	.218	.013	.103	.112	0.068	1	
ln(rent)	.323	.189	.399	.285	.109	-.198	.05	1
bootstrap s.e.	.338	15.243	0.691	0.025	0.163	0.222	0.259	0.179

Table 2.9: Variance and Correlation Matrix from Bootstrapping

	emp	earning	pop	estab	occupied houses	emp/pop	ln(wage)	ln(rent)
emp	1							
earning	.821	1						
pop	.698	.476	1					
estab	.798	.682	.786	1				
occupied housing	.663	.516	.696	.635	1			
emp/pop	.164	.043	-.412	-.223	-.128	1		
ln(wage)	.129	.218	.013	.103	.112	0.068	1	
ln(rent)	.323	.189	.399	.285	.109	-.198	.05	1
bootstrap s.e.	.338	15.243	0.691	0.025	0.163	0.222	0.259	0.179

Table 2.10: Panel Regressions with County Fixed Effects, 1980-2000

	(1)	(2)	(3)	(4)	(5)	(6)
	emp	inc	pop	house price	estab	occ house
Δmil_{ct}	1.394	84.085	3.162**	1.900	0.014	1.159**
	(0.861)	(49.968)	(1.573)	(2.395)	(0.022)	(0.573)
County FE	X	X	X	X	X	X
Period FE	X	X	X	X	X	X
N	762	762	762	762	762	762
First Stage F	3.007	3.007	3.007	3.007	3.007	3.007

Note: Standard errors in parentheses, clustered at the county level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 2.11: Dynamic Effect of Military Personnel Contractions- Dynamic panel with county fixed effects

	(1)	(2)	(3)
	emp	earning	estab
$(Mil_{cs} - Mil_{cs-1})/Pop_{c1980}$			
$s = t - 2$	-0.059 (0.146) [23.036]	21.532 (14.461) [23.036]	0.004 (0.012) [23.036]
$s = t - 1$	0.161 (0.174) [36.739]	10.667 (9.616) [36.739]	0.003 (0.009) [36.739]
$s = t$	0.422** (0.207) [17.667]	9.978 (14.159) [17.667]	0.014 (0.011) [17.667]
$s = t + 1$	-0.137 (0.241) [10.241]	-17.431 (10.812) [10.241]	-0.009 (0.009) [10.241]
$s = t + 2$	-0.052 (0.188) [21.165]	-9.572 (13.837) [21.165]	0.001 (0.009) [21.165]
County FE	X	X	X
obs	4355	4355	4355

Note: Years in the sample are from 1988 to 2000. All columns control for county dummy variables and are estimated using the 2SLS estimator. I include changes in military personnel up to two years before and after the current year. Outcome variables are the one year change of the county attributes as indicated by the short-hand on top of each column. Specifically, the outcome variable is annual change of civilian sector employment per 1980 population in Column 1; annual change of civilian sector labor income per 1980 population in Column 2; annual change in private business establishments per 1980 population in Column 3. Standard errors are clustered at the county level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Angrist-Pischke first stage partial F-statistics for each endogenous variable is reported in brackets.

Table 2.12: Influential Counties

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	civ emp	civ earning	civ pop	busi estab	log rent	occ houses	log wage	emp pop
Panel A: excl. DC area								
Δmil_c	1.191*** (0.436)	33.092* (19.084)	3.346*** (0.882)	0.106*** (0.036)	1.366*** (0.186)	0.921*** (0.189)	0.491 (0.319)	-0.492** (0.213)
N	18579	18579	18579	18579	18579	18579	18579	18579
# of base counties	308	308	308	308	308	308	308	308
first stage F	49.8	49.800	49.8	49.8	49.8	49.8	49.8	49.8
Panel B: excl. largest bases								
Δmil_c	1.076*** (0.373)	29.663 (18.457)	2.662*** (0.671)	0.082*** (0.026)	1.485*** (0.206)	0.913*** (0.204)	0.632* (0.371)	-0.282 (0.224)
N	18885	18885	18885	18885	18885	18885	18885	18885
# of base counties	319	319	319	319	319	319	319	319
first stage F	76.973	76.973	76.973	76.973	76.973	76.973	76.973	76.973
Panel C: excl. politically connected								
Δmil_c	1.279*** (0.411)	32.286* (17.634)	3.112*** (0.820)	0.098*** (0.033)	1.323*** (0.193)	0.901*** (0.185)	0.469 (0.295)	-0.261 (0.230)
N	17912	17912	17912	17912	17912	17912	17912	17912
# of base counties	310	310	310	310	310	310	310	310
first stage F	69.272	69.272	69.272	69.272	69.272	69.272	69.272	69.272

Note: Each panel includes a sample as indicated. Each column uses the outcome variable as indicated in the panel headline. All columns are estimated using the weighted 2SLS estimator using the predicted military personnel contractions as the instrumental variable. Number of observations, number of county groups, first stage F- statistics are reported. Standard errors in parentheses are first clustered at the county level, and then clustered at the county group level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.13: Commuting Zones

	(1)	(2)	(3)	(4)
	civ emp	civ earning	civ pop	private busi per capita
Δmil_z	2.885*** (0.999)	50.942 (44.762)	3.886** (1.510)	0.235*** (0.070)
	(5)	(6)	(7)	(8)
	occupied housing units	civ emp pop	log wage	log median rent
Δmil_z	1.525*** (0.529)	0.104 (0.329)	0.949 (0.777)	3.660*** (0.983)

Note: Each column uses the outcome variable as indicated in the column headline. There are 16,721 observations and 208 commuting zones (indexed by z) with military bases. All columns are estimated using the weighted 2SLS estimator using the predicted military personnel contractions as the instrumental variable and weights constructed from the synthetic matching. First stage F statistics is 72.877. Standard errors in parentheses are first clustered at the commuting zone level, then clustered at the commuting zone group level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.14: Heterogeneous Effects by the Size of the Cut

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	civ emp	civ earning	civ pop	busi estab	log rent	occ houses	log wage	emp pop
Panel A: cuts $\geq 25^{th}$ percentile								
Δmil_c	1.326*** (0.403)	36.319** (17.483)	3.025*** (0.803)	0.100*** (0.032)	1.342*** (0.192)	0.922*** (0.182)	0.457 (0.296)	-0.240 (0.226)
N	16848	16848	16848	16848	16848	16848	16848	16848
# of base counties	270	270	270	270	270	270	270	270
first stage F	72.394	72.394	72.394	72.394	72.394	72.394	72.394	72.394
Panel B: cuts $\geq 50^{th}$ percentile								
Δmil_c	1.340*** (0.405)	38.090** (17.726)	3.002*** (0.802)	0.100*** (0.032)	1.344*** (0.192)	0.926*** (0.182)	0.459 (0.295)	-0.254 (0.226)
N	11695	11695	11695	11695	11695	11695	11695	11695
# of base counties	180	180	180	180	180	180	180	180
first stage F	72.044	72.044	72.044	72.044	72.044	72.044	72.044	72.044
Panel C: cuts $\geq 75^{th}$ percentile								
Δmil_c	1.300*** (0.390)	38.455** (16.872)	2.959*** (0.770)	0.094*** (0.031)	1.275*** (0.184)	0.850*** (0.162)	0.352 (0.282)	-0.276 (0.221)
N	5996	5996	5996	5996	5996	5996	5996	5996
# of base counties	91	91	91	91	91	91	91	91
first stage F	72.858	72.858	72.858	72.858	72.858	72.858	72.858	72.858

Note: Each panel includes a sample as indicated. Each column uses the outcome variable as indicated in the panel headline. All columns are estimated using the weighted 2SLS estimator using the predicted military personnel contractions as the instrumental variable. Number of observations, number of county groups, first stage F- statistics are reported. Standard errors in parentheses are first clustered at the county level, and then clustered at the county group level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.15: Heterogeneous Effects by Population Density

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	civ	civ	civ	busi	log	occ	log	emp
	emp	earning	pop	estab	rent	houses	wage	pop
Panel A: high population density								
Δmil_c	1.103*	2.570	2.793***	0.063*	0.917***	1.147***	0.370	0.009
	(0.598)	(27.972)	(1.032)	(0.036)	(0.293)	(0.380)	(0.424)	(0.471)
N	11031	11031	11031	11031	11031	11031	11031	11031
First Stage F	42.647	42.647	42.647	42.647	42.647	42.647	42.647	42.647
Panel B: low population density								
Δmil_c	1.338***	47.402**	3.249***	0.114***	1.550***	1.168***	0.518	-0.403*
	(0.498)	(21.479)	(1.065)	(0.040)	(0.213)	(0.317)	(0.376)	(0.237)
N	8756	8756	8756	8756	8756	11031	11031	11031
First Stage F	38.415	38.415	38.415	38.415	38.415	42.647	42.647	42.647

Note: Each column uses the outcome variable as indicated in the column headline. Panel A includes base counties that have 1986 population density above the median, panel B includes those with 1986 population density below the median. All columns are estimated using the weighted 2SLS estimator using the predicted military personnel contractions as the instrumental variable. First stage F statistics are reported. Standard errors are reported in parentheses, first clustered at the county level, then clustered at county-group-year level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3 CHAPTER 2: THE EFFECT OF AID ON GROWTH: EVIDENCE FROM A QUASI-EXPERIMENT

3.1 Introduction

Whether foreign aid causes economic growth in recipient countries is a highly debated research question. Following the influential studies by Boone (1996) and Burnside and Dollar (2000), many others have emerged, but a basic consensus is still absent. William Easterly and Roodman (2004) show that the key finding of Burnside and Dollar (2000) – namely that aid contributes to growth but only where economic policies are favorable – is not robust to the use of an updated and enlarged dataset. Rajan and Subramanian (2008) and Channing Arndt and Tarp (2010) are only two of many recent papers that review the bulk of the existing literature yet arrive at differing conclusions. Identification of the causal effect of aid on growth has been elusive so far due to foreign aid being endogenous in growth models. An instrumental variable is needed to address these problems. However, as Clemens et al. (2012) conclude in their recent assessment: “the aid-growth literature does not currently possess a strong and patently valid instrumental variable with which to reliably test the hypothesis that aid strictly causes growth.

In this paper we contribute to this literature by instrumenting, in an economic growth equation, the endogenous foreign aid variable exploiting a plausible quasi-experiment created by the income threshold set by IDA (International Development Association), the World Bank’s program of grants and concessionary loans to low-income countries. Exploiting this new instrument, we are able to plausibly investigate the causal effect of aid and growth.

This income threshold has been used as a key criterion in allocating scarce IDA resources since 1987, and is adjusted annually only to take into account inflation. Other major donors also appear to use the IDA threshold as an informative signal, and we show that total aid declines significantly once a recipient country crosses the IDA income threshold from below. The IDA threshold is nevertheless an arbitrary income level which does not necessarily represent any structural change in economic growth. Threshold crossing is thus a plausibly valid instrumental variable for variations in aid over years for a recipient country in a panel data model that controls for initial income levels and also includes country and period effects (we

group years into 8 three-year periods).

The main concern with the identification strategy exploited in this study is that some countries might cross the threshold by having a series of large positive shocks that are eventually reversed, making the exclusion restriction invalid.⁵⁷ This should not necessarily be the case since our estimated empirical growth model can account for the differential timing at which countries cross IDA threshold. In our analysis we exploit only the data for the countries that cross the threshold from below during the period studied while our growth model allows countries to grow at different rates over time (by allowing for country-specific effects on growth and by allowing conditional convergence). Additionally, countries start at the beginning of the sample period from different levels below the threshold. Hence, the differential timing at which countries cross the IDA threshold from below exploited for identification in our study does not have to be driven by unobservable shocks.⁵⁸ It could be accounted for by our empirical growth model. Nevertheless, we extensively investigate this threat to our identification strategy by implementing a battery of tests and robustness checks. Importantly, we use an alternative instrumental variable based on a smoothed income trajectory which is plausibly uncorrelated with the country specific idiosyncratic shocks and obtain similar results. All of the evidence gathered does not point toward rejecting our identification strategy.

Using a sample of 35 countries that crossed the IDA threshold from below between 1987 and 2010, we find that a one percent increase in the aid to GNI ratio raises the annual real per capita short term GDP growth rate by 0.0312 percentage points. The mean aid to GNI ratio at the crossing is 0.09, so a one percentage point increase in the aid to GNI ratio raises annual real per capita GDP growth by approximately 0.35 percentage points. Our effects are about 1.75 times as large as those reported by Clemens et al. (2012). Using OLS, they find that a one percentage point increase in aid/GDP (at aid levels similar to our sample mean) is followed by

⁵⁷On average, however, our sample of countries grew faster after crossing the IDA threshold than before, consistent with the fact that developing countries in general exhibited better performance in the latter part of our 1987-2010 sample period.

⁵⁸Although this is true for crossing from below, it is unlikely to be true for the smaller set of countries crossing from above the threshold; at least those countries were displaying systematic negative growth rates such that the country fixed effects in the growth model were negative. Moreover, IDA policies are premised on the expectation of growth and eventual graduation, and there is no corresponding set of formal policies for “de-graduation”: instances of crossing the threshold from above are dealt with in a more ad hoc fashion. Crossing from above and from below may therefore have highly asymmetric effects on aid, and in turn on subsequent growth. Thus, our study focuses only on the countries that cross the IDA threshold from below.

at most a 0.2 percentage-point increase in growth of real GDP per capita. We find similar effects of aid on growth to the ones reported by Clemens et al. (2012) without instrumenting foreign aid. We also present evidence consistent with the fact that OLS estimates suffer from attenuation bias due to measurement error in aid, which is exacerbated when the variability in the aid to GNI ratios is exploited to identify the effect of foreign aid on economic growth in a fixed effect or first-differenced growth equation commonly used in the literature. Thus, one should expect that 2SLS using a valid instrument produces larger estimates than OLS. This could also be the case if economic growth affects aid levels.

The sizable effect of foreign aid on economic growth we find may also be attributable in part to the fact that our sample consists of a group of similar low-income countries that are financially constrained, where aid could have the largest impact. Prior to crossing the IDA threshold, the countries we study tend to have relatively large amounts of aid and low capital levels. Aid is likely to be relatively effective in this context. Although we focus on a small group of countries, our results have strong policy relevance since the sample is composed of low-income countries. We also show suggestive evidence that our results might have meaningful external validity to the remaining poor countries as they grow closer to the IDA threshold.

Investment appears to be the channel through which aid affects growth. We show the investment rate drops following the reduction in aid. Increasing the aid to GNI ratio by one percentage point increases the investment to GDP ratio by 0.54 percentage points. The magnitude of the effects on growth and investment is consistent with the average capital stock to GDP ratio for the sample countries, which we estimate to be approximately 2.

As in most of the literature relying on panel data, we estimate the short-run effect of aid on economic growth, an effect that mainly operates through physical investment. In the long run, aid could affect growth through several other channels, but its identification requires exogenous changes in aid over a very long period of time. Our instrument does not provide such exogenous variability to estimate that parameter.⁵⁹

⁵⁹Regressing the average growth rate over a long period of time onto the average aid on that period does not identify the long-term effect of aid on economic growth, even if aid were exogenous in that equation.

3.2 Previous Aid-Growth Studies

Identifying the causal effect of foreign aid on economic growth is fraught with difficulties. First, aid relative to GNI is likely measured with error.⁶⁰ The problem of measurement error is exacerbated as the estimated model is often demeaned or first differenced in order to eliminate the country fixed effects (Griliches and Hausman, 1986). Second, identification might be confounded by unobserved factors that determine both economic growth and aid. Third, growth itself could also affect aid. In response to these potential problems, previous studies have introduced different instrumental variables to identify the causal effect of aid on growth, in some cases using cross-sectional data and in others panel data. In this section we briefly review the two major identification strategies previously exploited in the literature.

Studies that use cross-sectional data often rely on population size, economic policies and donor-recipient political connections as instruments for aid (e.g., Boone, 1996; Burnside and Dollar, 2000; Rajan and Subramanian, 2008). These cross-country instruments are likely to violate the exclusion restriction since they are correlated with observable and plausibly also unobserved country-level characteristics that also contribute to economic growth. For example, population size can affect economic growth through channels other than aid Bazzi and Clemens (2013). Donor-recipient ties (e.g. colonization, trade or migration) that are correlated with aid flows can also affect growth indirectly through the institutional environment Daron Acemoglu and Robinson (2001) or other channels.

Other studies relying on panel data rule out the impact of time-invariant determinants of economic growth by either first differencing the data or by conditioning on country fixed effects. These studies tend to focus more on the short term effect of aid on growth. Many studies in this category adopt a dynamic panel model and employ difference GMM or system GMM estimators. They instrument for current aid with lagged values of income and aid, as well as with other standard cross-country regressors Hansen and Tarp (e.g., 2001); Rajan and Subramanian (e.g., 2008). However, recent studies show that GMM estimators of dynamic panel models using all mechanical instruments are unstable and potentially biased in finite samples, due to the problem of many and weak instruments Roodman (2007,0,0); Bazzi and Clemens (2013); Bun and Windmeijer (2010). System GMM estimators, in addition,

⁶⁰Even if aid/GNI did not display measurement error, GNI itself could not be strictly exogenous in a growth equation.

might also suffer from the lack of valid exclusion restrictions.⁶¹

We contribute to the literature by proposing a plausibly valid instrument for foreign aid. The IDA income threshold is exogenously determined and applicable to all recipient countries, and hence uncorrelated with their current characteristics and historical backgrounds. Controlling for a continuous measure of per capita income, crossing the IDA threshold from below is found to significantly reduce aid levels.

3.3 Data and Sample

The data used in this study are primarily from two sources. Income, investment, economic growth, and other country characteristics are from the World Bank's World Development Indicators (WDI).⁶² Aid data are obtained from the Development Cooperation Directorate (DAC) of the OECD.⁶³ Following much of the previous literature, aid is measured by total net Official Development Assistance (ODA) disbursements as a share of GNI, in current US dollars.

We identify 35 countries that crossed the IDA income threshold from below between 1987 and 2010.⁶⁴ Table 3.1 shows the names and years of crossing for these countries. For countries that crossed the threshold more than once, in our baseline specification we consider only the first crossing in defining the instrumental variable. We then show, however, that our results are robust to changes in this criterion. Following the convention of the literature, we smooth out fluctuations in the annual data by using period averages. Due to the length of our panel dataset, and (more importantly) because IDA has a three-year replenishment cycle, we group years into 8 three-year periods that roughly coincide with the IDA replenishment

⁶¹A few other recent studies exploit donor-recipient connections interacted with over-time variations in total donor contributions. For example, Eric Werker and Cohen (2009) use the political connections between OPEC countries and other Islamic countries and changes in oil prices to identify how aid money is spent. They find a small and marginally significant effect of aid on growth. Nunn and Qian (2014) find that humanitarian aid extends civil war by using variation over time in crop harvests to instrument for the amount of humanitarian aid that a country receives.

⁶²<http://databank.worldbank.org/ddp/home.do?Step=12&id=4&CNO=2>, accessed and extracted in August, 2012. The WDI dataset is usually updated 4 times a year. It not only adds the most up-to-date data, but also revises historical data, some of them from many years back. Most of the revisions are minor.

⁶³Data are from DAC Table 2a, available at <http://stats.oecd.org/Index.aspx?DatasetCode=TABLE2A#>, accessed in August, 2012.

⁶⁴Sao Tome and Principe crossed the threshold in 2009. It has only 2 periods of data in the sample and is thus automatically dropped from the analysis and hence also from the sample.

periods.⁶⁵ The first period, with data from calendar years 1987-1989, corresponds roughly to IDA8, covering fiscal years 1988-1990 (July 1, 1987 to June 30, 1990). The final period, with data from 2008-2010, roughly corresponds to IDA15 (July 1, 2008 to June 30, 2011). Donors pledge contributions for each replenishment period, rather than annually. Moreover, policies for allocating IDA funds (e.g. the relative weights assigned to poverty, quality of economic policies, and quality of governance) among eligible recipients are often modified between IDA periods but never within an IDA period. For this reason country allocations should be more correlated from one year to the next within an IDA period than across two replenishment periods. The 3-year IDA periods are therefore a natural way of grouping the data. The timing of actual graduations from IDA also tends to coincide with the end of replenishment periods. The baseline sample contains 247 country-period observations.⁶⁶

Table 3.2 shows the summary statistics for the baseline sample. Real per capita GDP of the sample countries grew at an average annual rate of 2.9%. Investment on average accounted for 25% of GDP. ODA equaled about 8% of GNI for a typical country in a typical year in the sample. Of total ODA, about 9% is from IDA, 67% is from DAC countries, 2% is from non-DAC countries, and 23% is from multilateral agencies other than IDA.

3.4 The IDA Threshold and Foreign Aid

Beginning in 1987, a major criterion for IDA eligibility has been whether or not a country is below a certain threshold of per capita income, measured in current US dollars. This “operational threshold” was established for the purpose of rationing scarce IDA funds. Figure 3.1 shows the evolution of the IDA threshold converted in current US dollars between 1987 and 2010. It was originally set at \$580, and has been adjusted annually only for inflation. By 2010, the threshold had increased to \$1175.

Once a country has exceeded the IDA income threshold and is judged to be

⁶⁵Recent studies that use panel data often group years in 4- or 5-year periods. As Temple (1999, 132) observes, “The question of when to test for growth impacts plagues the entire growth literature, not just aid-growth research. Empirical research on the determinants of growth cannot escape the selection of a fixed observation period, but “selecting the time intervals over which to study growth ... is a question that remains largely unsettled.” (excerpted from Clemens et al., 2012). We nevertheless follow the convention of the literature by using non-overlapping periods.

⁶⁶Appendix Table 3.16 lists the definitions and sources of data for variables used in this paper.

creditworthy, it is considered on track for “graduation” from IDA. Allowance is made for the possibility of income fluctuations, so lending volumes typically are reduced (and repayments accelerated) only after a country has remained over the threshold for three consecutive years. Thus, in most cases threshold crossing will result in reductions of IDA flows beginning in the next replenishment period, not in the current one World Bank (2010,0). The decline in aid from IDA is amplified by similar behavior from other donors. Some agencies such as African Development Bank (AfDB) and Asian Development Bank (AsDB) explicitly use the IDA income threshold in their own aid eligibility criteria. Other donors often view crossing the IDA income threshold as a signal that countries are in less need of aid and cut their own aid, reinforcing the decline in aid from IDA Moss and Majerowicz (2012). As a result, although IDA usually contributes less than 10% of the total aid to a typical recipient, crossing the IDA threshold may have a sizable effect on total aid.

The relevance of IDA threshold crossing as an instrument for aid can be tested by looking at its effects on total aid and aid from different donors. We distinguish among four groups of donors: IDA, DAC (OECD Development Assistance Committee) bilateral donors, non-DAC bilateral donors, and other multilateral donors. Lagged aid is the main explanatory variable in our growth regressions, so our instrumental variable is a dummy indicating whether the country has crossed the IDA threshold at least two periods earlier. Throughout the paper we use t to represent a specific year and s to represent a specific period. We define $Crossing_{i,s-2}$ equal to 1 if a country’s first threshold crossing during the sample period took place at least two periods before period s , where s includes years $t - 2$, $t - 1$, and t . Otherwise, $Crossing_{i,s-2}$ equals 0. We estimate the following equation:

$$Aid_{j_{is-1}} = \beta_1 y_{is-1} + \beta_2 Crossing_{is-2} + \beta_3 Pop_{is-1} + \lambda_i + \tau_s + v_{j_{is}} \quad (3.1)$$

The dependent variable $Aid_{j_{is-1}}$ is the log of average ratio of aid from donor type j to GNI or the log ratio of average total aid (i.e., the sum of aid from all donor sources) to GNI for country i in period $s-1$, that is $Aid_{j_{is-1}} = \ln[\sum_{k=3}^5 (ODA_{j_{t-k}}/GNI_{it-k})/3]$.⁶⁷ y denotes log real per capita GDP measured in constant 2000 US dollars. y_{is-1} is measured as log real per capita GDP in the second year of the last period $s - 1$

⁶⁷We follow the convention of the majority of the literature and measure both GNI and ODA in current US dollars. A minority of studies, such as Boone (1996), use GNI in purchasing power parity terms, however.

and hence it is equal to y_{it-4} . Pop_{is-1} is the log average population of period $s - 1$. $Crossing_{is-2}$ is defined as earlier. This second lag is introduced because the IDA graduation process – including cuts in new lending and acceleration of repayments – typically begins only three years after a country crosses the threshold, i.e. in the next replenishment period. The crossing status lagged one period relative to aid also allows time for other donors to respond to threshold crossings.

Table 3.3 reports the results of estimating Equation 3.1. For Column 1 to Column 5, respectively, the dependent variables are the one-period lag of the logarithm of aid share of GNI from (1) IDA, (2) DAC countries, (3) non-DAC countries, (4) multilateral agencies except for IDA, and (5) all donors.

To be conservative, we use two alternative methods to conduct statistical inference throughout the paper. We first report robust standard errors clustered at the country level, which allow for arbitrary within-country correlation. There are 35 countries in our sample. Standard asymptotic tests might over-reject the null hypothesis under the presence of few clusters Bertrand et al. (2004). Although 35 clusters is not a small number, for robustness we also report the p-value from the wild bootstrap-t procedure following Cameron et al. (2008).⁶⁸ Either approach yields very similar statistical inferences.

We find that following IDA threshold-crossing, the IDA to GNI ratio dropped, on average, by about 92 percent (i.e., $1 - e^{-2.5}$). Other donors also cut their aid substantially. Estimates of the coefficients associated with threshold crossing are negative and substantial in magnitude. Except for aid from non-DAC donors, the estimated coefficients are also statistically significant at conventional levels. The total aid to GNI ratio dropped, on average, by 59 percent (i.e., $1 - e^{-0.88}$). Higher income levels are also a strong predictor of aid: a one percent increase in real per capita GDP is associated with reductions in aid of about 8.6 percent from IDA, 1.4 percent from DAC countries, 4.7 percent from non-DAC countries, 2.5 percent from other multilateral agencies, and 1.5 percent for the overall ODA to GNI ratio.

We test whether our results are robust to controlling for a quadratic relationship between aid and log initial income level, with results shown in Table 3.11. The

⁶⁸To calculate the wild bootstrap p-value, we first calculate the Wald statistic w_0 associated with the estimate $\hat{\beta}_0$ using the original sample. We then estimate $\hat{\beta}_b$ and its standard error $se(\hat{\beta})$ for each wild bootstrap sample b out of B ($= 1000$) bootstrapped samples: $w_b = (\hat{\beta}_b - \hat{\beta}_0)/se(\hat{\beta})$. The p-value is computed as follows: first we locate in the distribution of the w_b the percentile of w_0 , α . Then, if $\alpha \leq 0.5$, the p-value is equal to 2α while if $\alpha > 0.5$, then the p-value is equal to $2(1 - \alpha)$.

estimated coefficients on the IDA threshold crossing dummy all increase slightly (in absolute value), and the quadratic specification does not improve the fit between aid and income. Most notably, the coefficients for log initial income level and its square are not statistically significant for total aid/GNI (see Column 5, Table 3.11).

We conduct a placebo test to further ensure that these effects are not a statistical artifact. Specifically, we replace the true IDA threshold value with a fake threshold equal to 50% of the true value, and re-estimate equation 3.1 using a threshold-crossing dummy variable based on this false threshold.⁶⁹ Table 3.4 reports the results of this falsification test. In the analysis we retain only country-period observations prior to the period in which countries cross the actual threshold, so the regression sample is unaffected by the effect of actually crossing the true threshold.⁷⁰ Crossing the false threshold has no significant effect on aid, and its coefficient has an opposite sign of the coefficient associated with the true threshold value.

Another concern is that IDA threshold crossing may not be a good instrument to identify the direct or structural effect of aid on economic growth because country leaders may endogenously alter policies to take advantage of potential complementarity (or substitutability) between aid money and policies.⁷¹ *Ex ante* it is unclear whether and how aid might affect the quality of policymaking—one can conceive easily aid improves policymaking or instead make policymaking worse by stimulating rent seeking (e.g., Rodrik, 1996)—and whether country leaders can engineer quick policy changes to have an immediate effect on growth along desired directions.⁷² We investigate this potential threat to our identification strategy by estimating equation 3.1 after replacing aid as the dependent variable by a set of variables measuring policymaking and institutional quality. These variables include measures of civil liberty and political rights from Freedom House, the World Bank CPIA (Country Policy and Institutional Assessment), broad money (M2) as a percentage of GDP, inflation measured by the GDP deflator, and dummy variables

⁶⁹50% of the true value is an appropriate level for the placebo test. Few countries have income levels below alternative fake thresholds that are much smaller than 50%. Fake thresholds that are much closer to the true level predict the actual crossings too well, defeating the purpose of a placebo test.

⁷⁰For sample countries with per capita GNI always above the fake threshold, the crossing dummy is replaced with 0.

⁷¹This threat applies also to papers that take aid as exogenous.

⁷²As Rodrik (1996) notes, “. . . external resources reduce the costs both of reform and of doing nothing—that is, avoiding reform.” Thus, ultimately, it is an empirical question whether aid or, more generally, conditions attached to loans affect policymaking and/or policies, which might have a different answer in different samples.

indicating respectively bank, currency and debt crisis (See Table 3.16 for the definition of these variables).⁷³ Results reported in Table 3.12 show that crossing the threshold has no statistically significant effect on any of the variables considered. Thus, we document that crossing the IDA income threshold mainly affects the level of aid and not quality of policymaking and policies adopted in the countries in our sample. We also show below that our main results are robust to including these variables as economic controls.

3.5 Foreign Aid and Economic Growth

3.5.1 Econometric Models

We postulate the following model in order to test the null hypothesis that foreign aid does not affect economic growth:

$$g_{is} = \beta_1 y_{is-1} + \beta_2 Aid_{is-1} + X_{is} \cdot \beta_3 + \lambda_i + \tau_s + \varepsilon_{is}, \quad (3.2)$$

where, as before, s denotes non-overlapping 3-year periods. Period s includes years $t, t-1, t-2$. y denotes log real per capita GDP. g_{is} , constructed as $(y_{it} - y_{it-3})/3$, is the arithmetic average real GDP per capita growth rate of country i in period s . y_{is-1} is measured as the log real per capita GDP in the second year of the previous period (i.e., y_{it-4}).⁷⁴ We expect β_1 to be negative. The initial income level captures conditional convergence – as a country gets richer, it grows more slowly, *ceteris paribus*. Aid_{is-1} is the log of average aid received by country i as a share of GNI in the previous period.⁷⁵ We use the one-period lag of aid instead of contemporaneous

⁷³We do not include trade as a percentage of GDP in the analysis in this section, because our interpretation of the effect of aid in our sample is that it operates directly by allowing financially constrained recipient countries to increase investment, which likely affects imports, and hence total trade as a percentage of GDP.

⁷⁴Notice that, by construction, y_{is-1} is not mechanically correlated with the dependent variable. Some studies in the literature use per capita real GDP in purchasing power parity terms to measure income level and to calculate growth (e.g., Boone, 1996). Real per capita GDP based on current exchange rates (in constant dollar terms, and using the Atlas method) and real per capita GDP in PPP terms are highly correlated: in our sample, the correlation is over 0.95. Growth rates constructed from the two versions are essentially the same. We use per capita GDP based on current exchange rates (in constant dollars) because there are fewer missing observations in the WDI database than for the PPP measure. Using instead the PPP measure we obtain almost identical results for our basic specifications in Table 3.5.

⁷⁵The measure of aid is slightly different from most of the literature which often use the aid to GDP or GNI ratio as the main explanatory variable. We take the log of aid since previous

aid in order to allow time for aid to take effect, following Clemens et al. (2012). X_{is} is a vector of time-varying variables, which in our baseline specification includes the logarithm of population and it is assumed to be strictly exogenous (though we do not use its lagged values as instruments in any specification). We check below the robustness of our key results to controlling for other potential growth determinants. λ_i is the country i fixed effect. τ_s is the period s fixed effect.

The standard way to estimate equation 3.2 is to eliminate the unobservable country-specific effects, λ_i , by including a set of country dummy variables in the model, which is equivalent to demeaned equation 3.2 and estimate the transformed equation by OLS. This estimator, however, is likely to be inconsistent due to aid being also affected by economic growth, measurement error, and (perhaps also) time-varying unobservable variables. We therefore instrument aid in period $s - 1$ (i.e., Aid_{is-1}) with a dummy variable indicating whether the country has crossed the IDA threshold by the end of period $s - 2$, that is, $Crossing_{is-2}$, as defined in Section 3.4. To address the endogeneity of the initial income level y_{is-1} , we instrument the initial income level with further lags of the income level.

Our instrumental variable is based on per capita (nominal) GNI crossing the IDA threshold two periods earlier. Per capita (nominal) GNI level in period $s - 2$ is correlated with the idiosyncratic shock to (real) economic growth of that period, ε_{is-2} .⁷⁶ Thus, estimating equation 3.2 by means of the fixed effects estimator, which first de-means the equation, mechanically introduces a correlation between the instrumental variable and the demeaned error term, $\ddot{\varepsilon}_{is} = (\varepsilon_{is} - \bar{\varepsilon}_i)$. However, if ε_{is} is not serially correlated, the correlation between $\ddot{\varepsilon}_{is}$ and $\ddot{\varepsilon}_{is-2}$ will be small if the

evidence suggests that the marginal effect of aid on growth is decreasing. The logarithmic form is a parsimonious way to introduce concavity while preserving our ability to identify aid's causal impact with only one exogenous binary instrumental variable. However, note that instrumented aid still takes on a large number of values on its domain, since each country's aid is shifted by the instrument starting from different values (over time). The logarithmic specification is admittedly less flexible than a quadratic specification; in particular, it does not allow the marginal effect of aid to change its sign. Clemens et al. (2012), however, find that the effect of aid on growth does not turn negative until aid exceeds roughly 15% of GDP. In our sample, over 90% of the observations are below 11% of GDP. Thus a logarithmic specification provides a good approximation over the range of observed values on aid. Additionally, we report as a robustness test in Table 3.14 results from using $Aid_{is-1}^* = (\sum_{k=3}^5 (ODA_{it-k}/GNI_{it-k}))/3$ (i.e. not logged) as the measure for aid. Aid_{is-1}^* has a positive, quantitatively large, and marginally significant coefficient, despite a lower first stage F statistic. We find that a 1 percentage point increase in the aid to GNI ratio raises annual per capita GDP growth by 0.57 percentage point at the sample mean of the aid to GNI ratio. This estimate is even larger than our baseline result, reported later.

⁷⁶This correlation could be weak, however, since our instrumental variable is a dichotomous variable and the error term only affects the latent process behind the binary variable.

time dimension of the panel is large. Our sample has 8 periods, which is not considered short in the literature. We will also show below that we do not reject the null hypothesis of no serially correlation of the error terms in equation 3.2. Note, nevertheless, that since ε_{is-2} is negatively correlated with $\tilde{\varepsilon}_{is}$, $Crossing_{is-2}$ is also probably correlated with the demeaned error term and, hence, the 2SLS estimator of the demeaned equation will likely under-estimate, *ceteris paribus*, the true effect of aid on economic growth.

We propose two approaches to circumvent this potential statistical problem. The first approach relies on a smoothing method of the latent process that determines our instrumental variable. Specifically, we do the following. We take a panel of 130 developing countries that appear in the official DAC aid recipient country list at any time between 1987 and 2010, other than the 35 countries in our original sample.⁷⁷ We demean all the series in our extended panel (165 countries in total) by projecting the annual log of nominal per capita GNI onto a set of country fixed effects, denoted \hat{y}_i . We then take the residuals, \hat{e}_{it} . For each of the 35 countries in our working sample, we construct a set of weights $w_j \in \{w_1, w_2, \dots, w_J\}$ bounded between 0 and 1 for the 130 donor countries added to the dataset such that the following distance function is minimized:

$$D_i = \|\mathbf{e}_i - \sum_j w_j \cdot \mathbf{e}_j\|, \quad (3.3)$$

where $\|\cdot\|$ is the Euclidian distance operator. The vector \mathbf{e} includes the residuals \hat{e}_{ht} . For each country i , we use all the years in our sample because that minimizes the influence of a given observation around the period of crossing the IDA threshold. Denote the optimal weight assigned to country j as w_{ij}^* . We then define $\hat{e}_{it} = \sum_j w_{ij}^* \cdot \hat{e}_{jt}$, and construct the predicted log per capita nominal GNI as $\hat{y}_{it} = \hat{y}_i + \hat{e}_{it}$.⁷⁸ By the nature of construction, \hat{y}_{it} is plausibly uncorrelated with the error in equations 3.2.⁷⁹ Thus the predicted crossing of the IDA threshold based on $exp(\hat{y}_{it})$ is also

⁷⁷Since the both our sample and the extended dataset are unbalanced panels, for each of the 35 countries in our sample we only use a panel of balanced available donors.

⁷⁸This algorithm is in the spirit of the synthetic control approach discussed in Alberto Abadie and Hainmueller (2009), but its sole purpose here is to provide a set of weights to smooth income trajectories from country specific idiosyncratic shocks.

⁷⁹Exploiting the conditional expectation function of per capita GNI based on the cross country's own data to construct an alternative instrumental variable for aid would not work here because the factors of production, like capital, for example, could be subject to the similar country idiosyncratic shocks that we want to eliminate from the series of output in this robustness check.

likely to be uncorrelated with the error term, satisfying the exclusion restriction needed for identification. Accordingly, we also exploit in our analysis the predicted crossing two periods earlier, $Crossing_{is}^{pred}$, as the instrumental variable for Aid_{is-1} . There are 10 countries out of 35 for which we predict a different period of crossing than their respective actual period of crossing.⁸⁰

The second approach we adopt to circumvent the threat to our identification strategy is to first-difference equation 3.2 and estimating the following equation:

$$\Delta g_{is} = \beta_1 \Delta y_{is-1} + \beta_2 \Delta Aid_{is-1} + \Delta \mathbf{X}_{is} \cdot \beta_3 + \tau_s + \Delta \varepsilon_{is} \quad (3.4)$$

In using $\Delta Crossing_{is-2}$ to instrument ΔAid_{is-1} , our identification strategy exploits only the sharp variability in aid at the period after crossing the IDA threshold. Under treatment heterogeneity, both in terms of the effect of threshold-crossing on aid and of the latter on economic growth, this strategy will identify a particular local average causal effect. Instead, we can also use just $Crossing_{is-2}$ to instrument for ΔAid_{is-1} . Note that $\Delta \varepsilon_{is} = \varepsilon_{is} - \varepsilon_{is-1}$.

Thus, the validity of the exclusion restriction in this case basically requires that the error terms ε_{is} are serially uncorrelated. Therefore, the instrumental variable will be invalid if the unobservable idiosyncratic error term in the growth equation is serially correlated up to two periods. Below we investigate in depth the validity of this assumption. Finally, as another robustness check, we also use $Crossing_{is-2}^{pred}$ as the instrument for aid.

In the first differenced model, even if the error terms in equation 3.2 were *i.i.d.*, the transformed error terms will not be, and will exhibit first-order serial correlation. Standard GMM inference when using optimal weights takes into account this feature while 2SLS does not. Thus, through the rest of the paper (with the exception of models estimated by GMM), as in the previous section, we rely on two alternative methods to conduct statistical inference. We first report robust standard errors clustered at the country level, which allow for within-country correlation. For robustness, we also report the p -values from the wild bootstrap- t procedure. Both methods imply similar statistical conclusions for the main parameters of the growth equation.

⁸⁰These 10 countries are: Albania, Angola, Azerbaijan, Bosnia and Herzegovina, Guyana, Indonesia, Kiribati, Peru, Samoa, and Ukraine. Countries for which the predicted and actual periods of crossings differ are more likely to have experienced some abnormalities around the neighborhood of crossing.

3.5.2 Baseline Results

Column 1 of Table 3.5 reports the estimate of equation 3.2 without instrumenting aid while column 2 reports the estimate of equation 3.4 also without instrumenting aid. Columns 3 through 7 are all estimated by the 2SLS estimator where we instrument only for aid. Column 3 estimates equation 3.2 using $Crossing_{is-2}$ as the instrument for Aid_{is-1} . Column 4 uses the crossings based on predicted per capita GNI from the smoothing exercise, $Crossing_{is-2}^{pred}$, as the instrumental variable. Finally, columns 5 to 7 report estimates of equation 3.4 using respectively $\Delta Crossing_{is-2}$, $Crossing_{is-2}$ and $Crossing_{is-2}^{pred}$ as instrumental variables.

The fixed effect model in Column 1 and first difference estimate in Column 2 show that aid is positively correlated with real economic growth. The estimated coefficient is statistically significant at the 5% level but is small in magnitude. A one percent increase in the aid to GNI ratio increases annual real per capita GDP growth by 0.0105 and 0.013 percentage points.

Studies on aid and growth typically calculate the increases in annual real per capita GDP growth rate in percentage points implied by a one percentage point increase in ODA's share of GNI. Because we take logs of ODA/GNI, the implied effect of a percentage point increase in ODA/GNI depends on its level. We use the average aid to GNI ratio at the period of crossing (i.e., 0.09) because this is the most relevant value given our identification strategy. The OLS estimate in Column 1 suggests that a one percentage point increase in the aid to GNI ratio from the average level at the period of crossing is associated with a 0.12 percentage point increase in real per capita GDP growth. The result in Column 2 implies that a 1 percentage point increase in the aid to GNI ratio from the same level is associated with a 0.14 percentage point increase in growth. These findings are similar to those in Clemens et al. (2012). They address endogeneity of aid simply by lagging it one period, and find that a one percentage point increase in aid/GDP from the sample mean increases annual real per capita GDP growth by 0.1 to 0.2 percentage points in the next (4-year) period.

Columns 3 through 7 are estimated using the 2SLS method. The point estimates of the aid coefficient are more than twice as large as those estimated by OLS. In column 3 we report the estimate of equation 3.2 using $Crossing_{is-2}$ as the instrument for Aid_{is-1} . We find that a one percent increase in the aid to GNI ratio raises growth by 0.028 percentage points. The coefficient is statistically significant

at the 1% level according to cluster-robust standard errors or by the wild bootstrap procedure. The first stage is strong, with an F -statistic of about 16. Initial income is negatively correlated with growth, supporting the hypothesis of conditional convergence. In column 4 we use the predicted crossings based on the smoothed per capita GNI trajectory, $Crossing_{is-2}^{pred}$ as the instrument. The estimated coefficient associated with Aid_{is-1} increases slightly, as expected, to 0.352 and is statistically significant at the 5% level. The first stage is somewhat weaker, with an F -statistic of 7.4. Interestingly enough, the point estimates of the effect of aid on economic growth in columns 3 and 4 are not very different quantitatively and are not statistically different.

Columns 5 through 7 estimate the first differenced model in equation 3.4. The coefficients are all larger than those in Column 3 and Column 4. Column 5 uses $\Delta Crossing_{is-2}$ as the instrumental variable for ΔAid_{is-1} . In this specification, we are estimating the coefficient of aid using only the variability from the one period after crossing. The estimated coefficient associated with aid is 0.0475 and is statistically significant at the 5% level, and just misses the 5% level according to the wild bootstrap- t procedure. Higher initial income levels are associated with lower growth, and the coefficient is larger than those in the previous columns. Column 6 uses $Crossing_{is-2}$ as the instrumental variable for ΔAid_{is-1} . The first stage and the estimated coefficients are essentially unchanged from those in column 5. Column 7 uses the predicted crossing, $Crossing_{is-2}^{pred}$, as the instrument. The first stage is very strong, with an F -statistic of 24. The estimated coefficient associated with aid is 0.055 and is statistically significant at the 1% level. All in all, our instrumental variable estimates are robust and consistently larger than the OLS estimates.⁸¹

As we discussed in the previous subsection, for the first differenced model, our instrumental variable approach will be invalid if the unobservable idiosyncratic error term in the growth equation (equation 3.2) is serially correlated. We formally test for the presence of serial auto-correlation in the error terms in equation 3.2 following Arellano and Bond (1991). The Arellano-Bond test for serial correlation tests the n^{th} order of serial correlation of the first differenced error to infer the $(n - 1)^{th}$ order of serial correlation of the error terms in the original equation (see also Roodman, 2009b). We report the p -values of the Arellano-Bond tests for AR(2) after estimations in Columns 5, 6, and 7. None of the tests rejects the null

⁸¹Table 3.13 re-estimates the model in column 3 of Table 3.5 while including quadratic and cubic terms of y_{is-1} . The results remain similar.

hypothesis of no serial correlation in the errors in equation 3.2.

So far we have treated Aid_{is-1} as the only endogenous variable. However, the initial income level, y_{is-1} , is also endogenous. In columns 1 and 2 of Table 3.6 we re-estimate the models in columns 3 and 4 of Table 3.5, respectively, but also instrumenting y_{is-1} (i.e., y_{it-4}) with y_{it-5} . In column 3 we re-estimate the model in column 6 of Table 3.5 and uses y_{it-8} to instrument for Δy_{is-1} (i.e., $y_{it-4} - y_{it-7}$). The aid coefficient remains similar. Columns 4 and 5 are estimated by means of the difference GMM estimator, which is widely applied in this literature. In light of the potential problems of many instruments, we use a parsimonious set of instruments (Roodman, 2007,0,0; Bazzi and Clemens, 2013; Bun and Windmeijer, 2010). We use y_{it-8} , y_{it-9} , y_{it-10} as instruments. We use $Crossing_{is-2}$ as an instrument in Column 4 and $Crossing_{is-2}^{pred}$ in Column 5. With instruments outnumbering endogenous variables, we are able to test the validity of these over-identified restrictions, as a way to test for model specification. Both the Sargan and Hansen tests do not reject the null hypothesis of the validity of the over-identification restrictions in both columns.

Failing to reject the null hypothesis of the over-identification restriction, and failing to detect an $AR(1)$ structure in the error term in equation 3.2, both suggest that the error terms in the growth equation (equation 3.2) are serially uncorrelated. Here we provide a third piece of evidence. If the error terms were serially correlated, including lagged values of the dependent variable will likely alter the estimates of the aid coefficient. We re-estimate the model in Column 3 of Table 3.5 but include the once-lagged value of the dependent variable as a control variable (Column 6 of Table 3.6), or its twice-lagged value (Column 7 of Table 3.6), and then both the once- and twice-lagged values (Column 8 of Table 3.6). Columns 9, 10, and 11 of Table 3.6 repeat Columns 6, 7, and 8 but using $Crossing_{is-2}^{pred}$ to instrument aid. The estimates of the effect of aid remain similar.⁸²

We find a sizable effect of aid on economic growth. The estimated effects of aid on economic growth on Columns 1 to 5 in Table 3.6 are all very similar. Taking the point estimate in Column 2 we observe that a one percent increase in aid to GNI ratio increases real per capita GDP growth by 0.0312 percentage points. A one percentage point increase in the aid to GNI ratio from its average value at the

⁸²For the model in column 3 of Table 3.5, we find that the clustered standard errors and the robust standard errors are very similar (results not shown). The similarity between the two sets of standard errors is consistent with the evidence of lack of serial correlation of the error term in equation 3.2.

period of crossing (0.09) thus raises the growth rate by 0.35 percentage points.

In the short run, the main possible channel through which aid could cause growth is through fostering physical investment. Irrespective of the form aid takes, it constitutes a flow of funds to recipient countries, which (if they are financially and perhaps fiscally constrained) would release resources in the economy that could be invested. Because the countries we study were all financially constrained by definition, this is a relevant scenario. However, our results would not necessarily extrapolate to countries that are more developed and have better access to credit markets.

It is useful then to consider what would be the expected effect of an increase of 1 percentage point in aid if it were fully invested in physical capital. Assuming a linear technology at the aggregate level of the economy, the rate of economic growth would increase in the inverse of the capital-output ratio.⁸³ Using the standard perpetual inventory method, we estimate this ratio to be approximately equal to 3.2 in our sample.⁸⁴ Thus, one would expect that growth could increase by as much as 0.5 percentage points if all of the aid were invested.

How much aid fosters physical investment is therefore an empirical question, one we now investigate. In Table 3.15 we re-estimate the baseline specifications in Table 5 but replacing economic growth with the period average investment to GDP ratio as the dependent variable. The OLS estimates in Column 1 and Column 2 of the effect of aid on investment are statistically insignificant and have the “wrong” sign. Instrumenting investment in Columns 3 to 7, we find that the coefficient associated with aid becomes positive and ranges between 0.5 and 0.8. In one case it is significant at the 10% level but in general it is only marginally significant, with p-values around 0.15. In column 4 we see that a one percent increase in aid to GNI ratio increases the investment to GDP ratio by 0.049 percentage points. Evaluated at the average level of aid to GNI ratio at the period of crossing, an increase of 1 percentage point in the aid to GNI ratio increases the investment to GDP ratio by 0.54 percentage points. Using the estimated capital-output ratio of 2,

⁸³Note that even under a linear technology the effect of aid on economic growth could be strictly concave since, for instance, the effect of aid on physical investment might be decreasing in aid. Moreover, the aggregate technology could be non-linear.

⁸⁴Note also that this figure is also consistent with standard growth accounting assumptions. Assuming capital per capita depreciation rate of 10 percent per year, and an investment rate of 25 percent of GDP per year, a country with a capital-output ratio of 2 would grow, in per capita terms, at 2.5 percent per year, which is consistent with the figures in our sample.

a 0.54 percentage point increase in physical investment would raise economic growth by 0.27 percentage points, which is not far away from our back-of-the-envelope calculation of the effect of aid on growth.

3.5.3 Measurement Error in Aid

An issue that is overlooked in the aid-growth literature is that the amount of aid that a country receives as a ratio of its GNI is likely to be measured with error. Not all donors report their aid to the DAC in all years. For example, aid from the former Soviet Union and from China in the Mao era to other communist countries was not reported to the DAC. Aid from China and some other emerging donors has increased in recent years and is also not included in the DAC data. Additionally, but better understood, the denominator, GNI, is also measured with sizeable error for many less developed countries (Jerven, 2013). With classic measurement error, the OLS estimate of the effect of aid is biased towards zero. Demeaning or first differencing the model would exacerbate the bias if aid levels are persistent over time.

The natural experiment we exploit in this paper provides a unique opportunity to test the hypothesis that measurement error in foreign aid biases the OLS estimate towards zero. We have shown that the amount of aid a country receives tends to decline substantially following its crossing of the IDA threshold from below. Assuming that the measurement error is *i.i.d.*, it would contribute much less to the total variation in ΔAid_{is} in periods closer to threshold crossings. Thus the OLS estimates of equation 3.4 using only periods in the neighborhood of the crossings are likely to provide more accurate estimates of aid's effect on growth than the one exploiting all of the variability in aid from all periods. To test this, we re-estimate equation 3.4 by OLS and 2SLS (using $Crossing_{is-2}$ to instrument aid), successively narrowing the window of periods used in the analysis around the crossing point of each country.⁸⁵ We expect the OLS estimate of the effect of aid on growth to increase as we narrow the window of estimation. Naturally, given our identification strategy, the 2SLS estimate of the same parameter should remain stable independent of the window used for estimation. We find exactly that in Table 3.7, Panel B.

⁸⁵We rely on first differenced models in this exercise because changing the number of periods also affects the estimation of the country fixed effects, and we want to hold everything constant except for the signal to noise ratio in aid.

Panel A of Table 3.7 reports the OLS estimates. We start from the original sample, which has a maximum of 7 periods before crossing the threshold and 7 periods after crossing the threshold. Columns 2 through 6 gradually narrow to at most 2 periods before and 2 periods after the crossing point of each country. As we narrow the window, the estimated coefficient associated with foreign aid monotonously and gradually increases. The coefficient in Column 6 is 0.0201, more than 50% larger than that in Column 1. A generalized Hausman test of the null hypothesis that the coefficient associated with $Aid_{i,s-1}$ is the same in Column 1 and in Column 6 is rejected with a p-value of 0.07. These findings are consistent with the existence of significant measurement error in aid, reinforcing the supposition that the aid to GNI ratio is measured with substantial noise.

3.5.4 Bunching

Our identification strategy hinges on the large decline in the amount of aid received following the crossing of an arbitrarily given and pre-determined threshold. If countries can manipulate their income data to remain below the IDA threshold, then threshold crossing may not be a valid instrument for aid. Countries with lower expected growth and hence a greater “need” for aid would presumably be the most likely to understate their incomes to remain IDA eligible.

Endogenous manipulation of the income level is not likely to be prevalent for several reasons. First, the GNI estimates used by the World Bank are by no means entirely within a government’s control. The national accounts data produced by national statistical agencies are merely one of several inputs into the World Bank’s income estimates Jerven (2013). Governments cannot perfectly predict (1) the adjustments to those national accounts data often made by World Bank staff, (2) the exchange rates used, or (3) the population estimates used in constructing GNI per capita. Second, crossing also depends on the current IDA threshold, and its annual adjustments for global inflation rates cannot be predicted perfectly either. Finally, income level with respect to the threshold is not the only criterion for IDA eligibility; e.g. countries that cross the threshold from below can remain eligible if they are not judged to be creditworthy for IBRD or private lending.

Nevertheless, we tested for evidence of manipulation. A histogram on the left side of Figure 3.2 shows the distance between a country’s current GNI per capita

and the contemporaneous IDA threshold. All countries that were ever eligible for IDA between 1987 and 2010 are included, and each GNI per capita value in each country-year is treated as a separate observation. We group country-year observations in 100-dollar bins according to the distance between income level and the contemporaneous IDA threshold. If many governments understate GNI to stay below the IDA threshold, we should observe significant “bunching” of observations just below the threshold, relative to the number of observations just above it. Specifically, we should observe the bin just to the left of the threshold to be abnormally high relative to the neighboring bins. If there is no bunching, the numbers of observations in each bin should cross the threshold of zero smoothly. As shown in the histogram on the left in Figure 3.2, there is no visual evidence that countries bunch right below the IDA threshold. A formal test confirms this result. Using a density test proposed by McCrary (2008), we find no significant evidence of bunching. The graph on the right in Figure 3.2 shows the fitted kernel density functions at both sides of the threshold. Contrary to the pattern expected of endogenous manipulation, the density is actually lower to the left of the threshold. The estimated “jump” in density from left to right is the opposite of what we would expect if bunching exists (-0.000291, with standard error equal to 0.000818).

3.6 Robustness Checks

In this section we present various robustness tests. In all of these tests, we report two sets of results. The first uses $Crossing_{is-2}$ as the instrument for aid, and the second uses $Crossing_{is-2}$. All of the robustness checks are based on our preferred specification in Column 3 of Table 3.5 (or Column 4 of Table 3.5 when $Crossing_{is-2}^{pred}$ is used as the instrument). Results are largely consistent in both sets of results.⁸⁶

3.6.1 Omitted Variables

Throughout the study, we control for period fixed effects, log of initial income, and log of population. Period fixed effects take account of secular trends that affect all countries similarly. Initial income and population are among the key time-

⁸⁶Results are also robust to (for example) the model column 6 of Table 3.5.

varying factors that affect economic growth. We also control for country fixed effects which account for any time-invariant cross-country variability in economic growth. Other important determinants of growth such as the quality of economic policies and institutions exhibit much more cross-country variation than variation over time within a country, particularly over our relatively short 1987-2010 period. Their impact will thus mostly be captured by the country fixed effects.

We nevertheless test whether other time-varying factors could be confounding the effect of aid on economic growth, by adding to the baseline regression a host of economic and political variables, including the primary school enrollment rate, the Freedom House index of civil liberty and political rights, the World Bank's Country Policy and Institutional Assessment (CPIA) ratings, total trade as a percentage of GDP, broad money as a percentage of GDP, inflation as measured by the GDP deflator, and dummies for whether the country is experiencing a banking crisis, currency crisis, or debt crisis. Due to missing values, we add these variables in separate groups to maintain a reasonable sample size for each regression. Table 3.8 shows the results of these exercises. Columns 1 through 5 use $Crossing_{is-2}$ to instrument for aid while Columns 6 through 10 use $Crossing_{is-2}^{pred}$. In this small and homogeneous sample that controls for country fixed effects, few of the additional regressors have a statistically significant effect on growth in either set of estimations. In Columns 1 through 5, the first stage remains relatively strong throughout, while the estimated coefficients for aid remain robustly positive, large and significant. In Columns 6 through 10, the first stages are somewhat weaker. However, the aid coefficients remain statistically significant and similar in magnitude to their counterparts in the baseline regressions.

3.6.2 Timing of Crossing

Recall that our instrument is a dummy variable that switches from 0 to 1 two periods after the country crosses the income threshold from below. The countries in our sample all crossed the threshold at some point between 1987 and 2010, but for countries that crossed the threshold in the last and the next-to-the-last periods (i.e. in periods 7 or 8) the instrumental dummy variable is always zero. We keep these countries in the sample because they satisfy our simple rule for sample selection and they provide relevant information for estimating the effects of the control variables.

We now check whether our results still hold when we drop these countries.

In Column 1 of Panel A in Table 3.9, we drop the seven countries that crossed the threshold in the final period of our sample (2008-2010). The point estimate increases and remains statistically significant at conventional levels. In Column 2 we further drop the seven countries that crossed the threshold in the next-to-last period (2005-2007). The coefficient is slightly larger than that in Column 2 and remains statistically significant at the 5% level. The first stage remains strong in each case. The first stage is weaker when we use $Crossing_{is-2}^{pred}$ as the instrument in panel B, but the results are qualitatively similar.

As shown in Table 3.1, a few countries crossed the IDA threshold from below more than once during the sample period. These countries must have crossed the threshold from above after its first crossing from below, then crossed from below again. These cases might be a threat to our identification strategy. If income drops below the threshold again immediately after the first crossing for any reason other than a decline in aid, the estimated effect of aid might be confounded. In Column 3 of Panel A of Table 3.9, we drop countries with multiple crossings. The estimated effect of aid changes very little. In Column 4, we use the last threshold crossing (from below) instead of the first one to construct the instrumental variable. The estimated effect of aid on growth again changes very little, relative to the baseline specification. It remains statistically significant in panel A, but in panel B the standard errors increase and we cannot reject the null hypothesis of no effect at conventional levels.

3.6.3 Other Robustness Checks

There are two exceptional groups of countries in the sample. First, a few countries were never classified as IDA eligible in the 1987-2010 period due to various reasons, despite having income levels below the threshold for one or more years. We include them in the sample because the IDA income threshold potentially serves as a useful benchmark for donors other than IDA. Second, several countries in the sample benefit from the “small island country exception,” which permits island nations with populations below 1.5 million to remain IDA eligible even after surpassing the income threshold. In Column 5 we drop the three small island countries of Table 3.9. In Column 6, we drop the four countries that were never eligible for IDA throughout

the sample period. The estimated effects of aid are robust to these sample changes.

3.7 External Validity

Our sample consists of a special group of aid-recipient countries that all successfully crossed the IDA income threshold from below between 1987 and 2010. A natural question is how our results would apply to other aid-recipient countries, particularly those that are still below the IDA cutoff. On the one hand, if the sample countries crossed the IDA threshold level because they have fundamentally different attributes, the results of this paper may have little relevance for those remaining very poor. On the other hand, if the difference between crossing and non-crossing countries is mainly due to being in different development stages, and the two samples have similar growth patterns conditional on the initial income, then it is conceivable that our results may apply to these countries as well when their income level approaches the IDA threshold.

In this section we investigate whether the crossing countries have systematically higher growth rates (conditional on the initial income level) than the non-crossing countries using a simple regression. We include all country-year observations with per capita GNI level below the threshold, and linearly project the annual real per capita GDP growth onto a dummy variable indicating whether the observation belongs to a country in the crossing sample, controlling for a one-year lag of log per capita real GDP and its quadratic form, as well as for year dummies. When their income levels were still below the IDA threshold, real per capita GDP in countries that eventually crossed were growing on average 1.98% per year. During the same period, real per capita GDP in countries that had not crossed the threshold by 2010 were growing on average 1.11% per year. In Table 3.10 we show that, after we control for the initial real GDP per capita level (Column 1) as well as its quadratic form (Column 2, the difference in average annual GDP per capita growth between the crossing countries and the non-crossing countries is approximately 0.7 percentage points and not statistically significant. In Column 3, we compare the average annual GDP per capita growth of the two groups of countries within each quartile of the distribution of the lagged income levels. When the income level is in the lowest quartile, countries that eventually crossed the threshold were growing at a much higher rate than those that have not crossed. In other higher quartiles

of income level, the differences in annual growth rates between the two groups are small in magnitude and not statistically significant. Thus, for most of the income distribution below the IDA threshold, we do not find significant differences in pre-crossing growth rates between the sample of countries studied in this paper and those still below the IDA threshold.

3.8 Conclusions

This paper presents new evidence on the effect of foreign aid on the recipient country's economic growth, exploiting the substantial drop in aid after a country crosses an exogenous income threshold set by the World Bank for IDA eligibility. We use a group of countries that have crossed the IDA threshold since 1987 and find that foreign aid has a sizable positive effect on economic growth. Increasing the aid to GNI ratio by one percentage point from average aid to GNI ratio at the period of crossing raises the real per capita GDP growth rate by approximately 0.35 percentage points.

We address various identification concerns, and our results remain robust throughout. Despite focusing on a special group of countries, this group is particularly interesting because it is composed of poor countries for which aid is particularly important since they are financially constrained. Rendering further support to the positive causal effect of aid, we provide evidence that aid also increases the investment rate significantly in our sample. Indeed, a back of the envelope calculation suggests that the increase in physical investment is the main channel through which aid operates in the short-run. Finally, we provide some evidence that suggests that our results may generalize to countries that are still under the IDA threshold, as they grow closer to the threshold.

We estimate a model similar to those in the recent aid-growth literature, and address the identification challenge using a novel instrument for aid, derived from a plausible exogenously-determined aid allocation policy. We also provide a new way of constructing predicted income trajectory using a smoothing technique. Our finding of a positive impact of aid on growth is consistent with several other recent studies, most notably Clemens et al. (2012). Our estimated effect, however, is somewhat larger than in most other studies.

Identification of causal effects is a daunting task (See Samuelson, 1948), especially at the macroeconomic level. All causal estimates of country level parameters

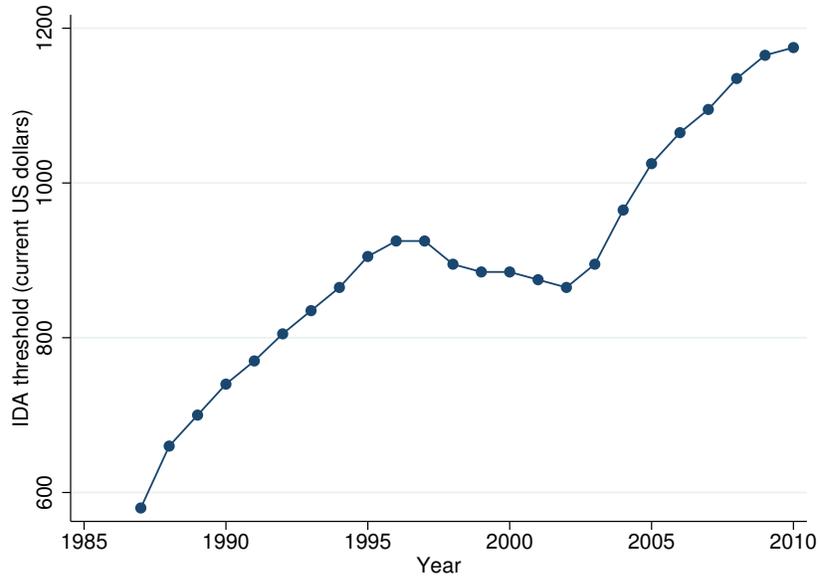
should be taken cautiously. Needless to say, at the micro level, researchers still need to evaluate on a case by case basis which aid projects work better, if at all. Our evidence only shows that overall foreign aid increases economic growth among poor countries where foreign aid is a large source of funding. Moreover, even at the macro level aid may have heterogeneous effects depending on recipient characteristics, aid modalities, and donor motives (Mekasha and Tarp, 2013).⁸⁷ Our relatively small and homogeneous sample is not ideal for testing heterogeneous effects of aid. Moreover, because we identify only the effect of aid on growth in the short term, our evidence does not contradict any view of the effect aid has on long-term economic development. Despite these caveats, we believe our evidence contributes to understand the effect of aid on economic growth in the short-term for poor countries that are financially constrained.

Our results also contribute to the empirical literature on donors' aid allocation decisions across recipient countries (See Alesina and Dollar, 2000; Chong and Gradstein, 2008). They support the conjecture by Moss and Majerowicz (2012) that bilateral donors use IDA policies – and specifically its income eligibility threshold – as an informative signal of recipient need. Patterns of donor “herding” measured by Frot and Santiso (2011) may be partially due to donors' common responses to recipient countries' crossing the IDA income threshold.

Our findings also have implications for IDA graduation policies. The threshold-crossing sample likely consists of growing countries facing financial constraints, perhaps explaining why aid has large positive effects in our study. If aid reductions are followed by declines in growth for this set of countries, arguably the IDA graduation process should be made lengthier. For example, declines in aid and changes in lending terms following threshold crossing could be made more gradual. Consideration should also be given to increasing the income threshold in real terms. Those changes could lead to a more effective use of IDA funds, compared to the alternative of concentrating more aid among the shrinking number of countries projected to remain under the current threshold (Moss and Majerowicz, 2012).

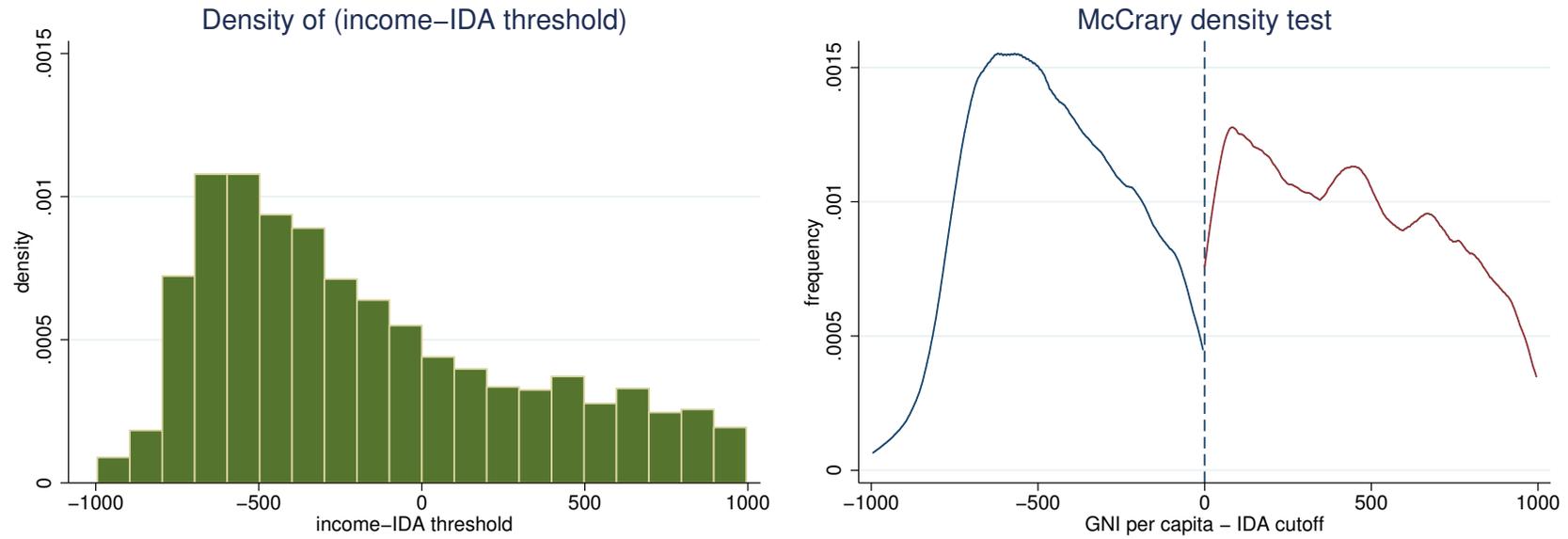
⁸⁷Bilateral aid in particular may be more effective following the end of the Cold War, if more of it is now motivated by developmental rather than geopolitical concerns Headey (2008). Our data are almost entirely post-Cold War.

Figure 3.1: IDA Threshold 1987-2010



Note: IDA threshold is originally nominated in current international dollars. We convert it in current US dollars.

Figure 3.2: Histogram of Income and McCrary Test of Bunching



Note: There are 1,920 country-year observations from 112 countries that were ever on the DAC list between 1987 and 2010. For each country-year observation, we calculate the distance of the current per capita GNI (y_{it}) from the current IDA threshold (\bar{y}_t). We restrict the distance ($y_{it} - \bar{y}_t$) between -1000 and 1000. The graph on the left is a histogram of country-year observations against $(y_{it} - \bar{y}_t)$, grouped in 100-dollar bins. The graph on the right shows the McCrary density test. Epanechnikov kernel density function with bandwidth of 100 US dollars. The difference in density between observations just to the left and those just to the right of the threshold is estimated by fitting two separate kernel density functions. The jump from the just left of the threshold to the just right of the threshold is 2.909×10^{-4} . Bootstrapped standard error is 8.18×10^{-5} . Choice of the kernel functional form is not important, the results are similar to a reasonable range of bandwidths.

Table 3.1: Sample Countries and Years of Crossing the IDA Threshold

Country Name	Year of Crossing	Country Name	Year of Crossing
Albania	1999	India	2010
Angola	2005	Indonesia	1994
Armenia	2003		2004
Azerbaijan	2005	Kiribati	1988
Bhutan	2004		1992
Bolivia	1997	Moldova	2007
	2005	Mongolia	2006
Bosnia and Herzegovina	1997	Nigeria	2008
Cameroon	2008	Papua New Guinea	2009
China	2000	Peru	1990
Congo, Rep.	2006	Philippines	1994
Djibouti	2007	Samoa	1995
Egypt	1995	Solomon Islands	1997
Equatorial Guinea	1998	Sri Lanka	2003
	2000	Sudan	2008
Georgia	2003	Syrian Arab Republic	1998
Ghana	2009	Timor-Leste	2006
Guyana	1999	Turkmenistan	2002
	2005	Ukraine	2003
Honduras	2000	Uzbekistan	2010

Note: Countries that crossed the IDA threshold from below between 1987 and 2010.

Table 3.2: Summary Statistics

Variable	N	mean	<i>s.d.</i>	25 th	50 th	75 th
real GDP per capita growth	247	.029	.054	.007	.028	.048
log real GDP per capita 4 years earlire	247	6.635	.551	6.274	6.666	6.988
lag of log(ODA/GNI)	247	-3.282	1.472	-4.473	-2.910	-2.184
lag of log(IDA/GNI)	247	-3.282	1.472	-4.473	-2.91	-2.184
lag of log(DAC/GNI)	247	-10.618	8.022	-21.043	-6.256	-4.689
lag of log(NDAC/GNI)	247	-3.774	1.574	-4.941	-3.356	-2.697
lag of log(Other MLA/GNI)	247	-11.281	6.073	-13.267	-9.246	-6.641
lag of ODA/GNI	247	.081	.094	.011	.054	.113
lag of IDA/GNI	247	.007	.012	0	.002	.009
lag of DAC/GNI	247	.054	.068	.007	.035	.067
lag of NDAC/GNI	247	.002	.005	0	0	.001
lag of other MLA/GNI	247	.019	.027	.002	.009	.024
crossed IDA threshold 2 periods earlier	247	.2308	.4222	0	0	0
lag of Investment/GDP	231	0.253	0.119	0.187	0.235	0.297
log population	247	15.950	2.259	14.765	15.874	17.108
lag of terms of trade (year 2000=100)	167	100.548	15.894	92.328	99.983	104.495
CPIA z-score	238	-.214	.934	-.876	.0192	.439
civil liberty	247	4.306	1.604	3	4	6
political rights	247	4.273	1.981	2	4	6
primary school enrollment	224	97.508	19.290	91.358	101.563	110.337
merchandized trade as % of GDP	247	64.469	31.215	41.355	60.568	83.855
broad money as % of GDP	225	37.383	25.520	18.216	31.825	50.164
inflation (%)	247	94.608	537.018	5.107	8.726	18.506
bank crisis (dummy)	247	.052	.224	0	0	0
currency crisis (dummy)	247	.109	0.313	0	0	0
debt crisis (dummy)	247	.202	.141	0	0	0

Note: Each observation is a country-period. For each variable, the mean, standard deviation, median, 25th percentile, and 75th percentile are reported. Missing values for ODA and ODA by donor are treated as zeros, following the precedent of Arndt, Jones and Tarp (2010) (Page 14). In this sample there are no missing values in total ODA. Zero values in aid by donor are replaced with 1 dollar before taking logarithm. See Appendix Table 3.16 for more details in construction and sources of these variables.

Table 3.3: IDA Threshold and Aid

	(1)	(2)	(3)	(4)	(5)
	IDA	DAC	NDAC	MLA	ODA
$Crossing_{is-2}$	-2.485 (1.371)*	-0.961 (0.238)***	-2.222 (1.776)	-0.750 (0.302)**	-0.876 (0.216)***
<i>Wild cluster bootstrap-t p-value</i>	0.058	0.000	0.192	0.004	0.000
y_{is-1}	-8.587 (1.691)***	-1.443 (0.420)***	-4.739 (1.964)**	-2.508 (0.865)***	-1.535 (0.324)***
<i>Wild cluster bootstrap-t p-value</i>	0.000	0.004	0.018	0.024	0.000
Country FE	X	X	X	X	X
Period FE	X	X	X	X	X
N	247	247	247	247	247
N countries	35	35	35	35	35

Note: Each observation is a country-period. Dependent variables are the log average share of aid in GNI by donor in the last period and share of total aid in GNI in the last period. There are 35 countries in the sample. Country fixed effects, period fixed effects, and log population in the last period are controlled in all columns. $Crossing_{is-2}$ is a dummy variable indicating whether the country crossed the IDA cutoff at least two periods earlier. y_{is-1} is the log real GDP per capita in the second year of the last period, y_{it-4} . Cluster-robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported.

Table 3.4: IDA Threshold and Aid - Placebo Threshold at 50% of the True Level

	(1)	(2)	(3)	(4)	(5)
	IDA	DAC	NDAC	MLA	ODA
$Crossing_{is-2}$	3.068	0.404	2.565	-0.118	0.299
	(1.867)	(0.661)	(2.396)	(0.383)	(0.514)
<i>Wild cluster bootstrap-t p-value</i>	0.006	0.672	0.042	0.712	0.708
y_{is-1}	-12.29	-2.251	-8.808	-4.622	-2.245
	(3.778)***	(0.859)**	(4.367)*	(1.507)***	(0.657)***
<i>Wild cluster bootstrap-t p-value</i>	0.000	0.032	0.000	0.000	0.004
Period FE	X	X	X	X	X
Country FE	X	X	X	X	X
N	162	162	162	162	162
N countries	34	34	34	34	34
F-statistics on $Crossing_{is-2}$					0.57

Note: Each observation is a country-period from countries that crossed the IDA threshold between 1987 and 2010. Country-period observations included in the sample are prior to the period of crossing the real threshold. There are 34 countries and 162 observations in the regression. Dependent variables are one period lag of average shares of ODA in GNI by donor. Country and period fixed effects, and log population in the last period are controlled in each column. F-statistic on $Crossing_{is-2}$ is reported in Column 5. y_{is-1} is log per capita real GDP in the second year of the last period, y_{it-4} . Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported.

Table 3.5: Baseline Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Main Specification</i>	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Aid_{is-1}	0.0105 (0.00455)**	0.0133 (0.00615)**	0.0281 (0.0100)***	0.0352 (0.0147)**	0.0475 (0.0239)**	0.0485 (0.0177)***	0.0552 (0.0190)***
<i>Wild cluster bootstrap-t p-value</i>	0.020	0.032	0.006	0.008	0.052	0.000	0.000
y_{is-1}	-0.0675 (0.0246)***	-0.161 (0.0231)***	-0.0371 (0.0256)	-0.0249 (0.0322)	-0.0976 (0.0516)*	-0.0957 (0.0387)**	-0.0835 (0.0415)**
<i>Wild cluster bootstrap-t p-value</i>	0.090	0.000	0.298	0.496	0.082	0.000	0.140
Period FE	X	X	X	X	X	X	X
Country FE	X		X	X			
First differenced		X			X	X	X
IV			X	X	X	X	X
IV from predicted income				X			X
IV first differenced					X		
N	247	212	247	247	212	212	212
Number of countries	35	35	35	35	35	35	35
First atage F statistic (Kleibergen-Paap Wald)			16.50	7.385	19.52	16.16	24.06
AR(2) p-value					0.729	0.830	0.824

Note: Each observation is a country-period. The dependent variable is the period average real per capita GDP growth rate. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported. See text for more details.

Table 3.6: Alternative Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Main Specification	2SLS	2SLS	2SLS	GMM	GMM	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Aid_{is-1}	0.0258 (0.00966)***	0.0312 (0.0138)**	0.0427 (0.0182)**	0.0298 (0.0122)**	0.0308 (0.0133)**	0.0198 (0.00984)**	0.0229 (0.0101)**	0.0205 (0.00909)**	0.0331 (0.0149)**	0.0359 (0.0152)**	0.0377 (0.0155)**
Wild cluster bootstrap-t p-val	0.008	0.010	0.018			0.042	0.020	0.012	0.000	0.000	0.000
y_{is-1}	-0.0525 (0.0226)**	-0.0431 (0.0290)	-0.137 (0.0679)**	-0.128 (0.0554)**	-0.136 (0.0529)***	-0.0540 (0.0184)***	-0.0543 (0.0217)**	-0.0518 (0.0201)**	-0.0326 (0.0282)	-0.0334 (0.0281)	-0.0245 (0.0261)
Wild cluster bootstrap-t p-val	0.054	0.158	0.026			0.012	0.020	0.020	0.874	0.844	0.916
Period FE	X	X	X	X	X	X	X	X	X	X	X
Country FE	X	X				X	X	X	X	X	X
First differenced			X	X	X						
IV for y_{is-1}	X	X	X	X	X	X	X	X	X	X	X
Predicted crossing		X			X				X	X	X
lagged dependent variables						1	2	1,2	1	2	1,2
N	247	247	212	212	212	245	229	229	245	229	229
Number of countries	35	35	35	35	35	35	35	35	35	35	35
First stage F stat (K-P Wald)	8.098	3.601	11.46	4.453	6.164	5.951	6.520	6.448	3.263	4.083	4.103
Number of IVs				12	12						
Hansen test for over-id (p -val)				0.330	0.174						
Sargan test for over-id (p -val)				0.275	0.106						
AR(2) p -value				0.950	0.849						

Note: Each observation is a country-period. The dependent variable is the period average real per capita GDP growth rate. Instrumental variable for y_{is-1} is y_{it-5} all columns except for Columns 3, 4, 5. y_{is-1} is instrumented by y_{it-8} in Column 3, and is instrumented by y_{it-8} , y_{it-9} , and y_{it-10} in Columns 5 and 6. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported. See text for more details.

Table 3.7: Narrowing Periods

Panel A: OLS - first differenced	(1)	(2)	(3)	(4)	(5)	(6)
# of maximal periods around the crossings	7	6	5	4	3	2
Aid_{is-1}	0.0133 (0.00615)**	0.0133 (0.00615)**	0.0137 (0.00640)**	0.0142 (0.00631)**	0.0154 (0.00814)*	0.0201 (0.00946)**
<i>Wild cluster bootstrap-t p-value</i>	0.032	0.032	0.040	0.026	0.142	0.098
y_{is-1}	-0.161 (0.0231)***	-0.161 (0.0232)***	-0.162 (0.0232)***	-0.163 (0.0235)***	-0.161 (0.0256)***	-0.157 (0.0298)***
<i>Wild cluster bootstrap-t p-value</i>	0.000	0.000	0.000	0.000	0.000	0.000
Period FE	X	X	X	X	X	X
N	212	211	203	188	165	133
Number of countries	35	35	35	35	35	35
Test for Aid_{is-1} ((6)-(1), p -value)						0.069
Panel B: 2SLS - first differenced	(1)	(2)	(3)	(4)	(5)	(6)
# of maximal periods around the crossings	7	6	5	4	3	2
Aid_{is-1}	0.0485 (0.0177)***	0.0481 (0.0177)***	0.0460 (0.0181)**	0.0442 (0.0157)***	0.0427 (0.0192)**	0.0527 (0.0222)**
<i>Wild cluster bootstrap-t p-value</i>	0.002	0.002	0.000	0.000	0.014	0.006
y_{is-1}	-0.0957 (0.0387)**	-0.0969 (0.0389)**	-0.103 (0.0397)***	-0.107 (0.0369)***	-0.113 (0.0441)**	-0.104 (0.0488)**
<i>Wild cluster bootstrap-t p-value</i>	0.008	0.008	0.006	0.004	0.012	0.040
Period FE	X	X	X	X	X	X
N	212	211	203	188	165	133
Number of countries	35	35	35	35	35	35
First stage F statistics	16.159	16.353	15.741	19.946	16.577	19.971
Test for Aid_{is-1} ((6)-(1), p -value)						1.000

Note: Each observation is a country-period. The dependent variable is the period average real per capita GDP growth rate. The growth equation is first differenced before estimation. IV in the first differenced equation is $Crossing_{is-2}$. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported.

Table 3.8: Adding Covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	IV is $Crossing_{is-2}$					IV is $Crossing_{is-2}^{synth}$				
	baseline	schooling	political	CPIA	econ cond	baseline	schooling	political	CPIA	econ cond
Aid_{is-1}	0.0281*** (0.0100)	0.0307*** (0.0110)	0.0287*** (0.0102)	0.0336*** (0.0105)	0.0308*** (0.0110)	0.0352** (0.0147)	0.0393*** (0.0152)	0.0368** (0.0158)	0.0383*** (0.0148)	0.0456** (0.0205)
<i>Wild cluster bootstrap-t p-value</i>	0.006	0.004	0.002	0.000	0.008	0.008	0.008	0.002	0.022	0.050
y_{is-1}	-0.0371 (0.0256)	-0.0402 (0.0273)	-0.0361 (0.0262)	-0.0232 (0.0241)	-0.0119 (0.0269)	-0.0249 (0.0322)	-0.0245 (0.0332)	-0.0222 (0.0343)	-0.0153 (0.0321)	0.0101 (0.0377)
<i>Wild cluster bootstrap-t p-value</i>	0.298	0.234	0.314	0.438	0.496	0.486	0.548	0.636	0.840	0.552
log population	-0.0086 (0.0738)	-0.0149 (0.0753)	-0.0114 (0.0754)	0.0423 (0.0767)	0.0133 (0.0556)	0.0161 (0.0859)	0.0203 (0.0869)	0.0149 (0.0879)	0.0584 (0.0903)	0.0571 (0.0858)
primary school enrolment rate		-0.0003 (0.0004)					-0.0003 (0.0005)			
Freedom House civil liberty index			0.0023 (0.0073)	-0.0036 (0.0062)	0.0038 (0.0063)			0.0021 (0.0078)	-0.0045 (0.0061)	0.0021 (0.0071)
Freedom House political rights index			0.0013 (0.0042)	0.0004 (0.0040)	-0.0018 (0.0039)			0.0026 (0.0044)	0.0011 (0.0039)	-0.0001 (0.0041)
World Bank CPIA-Z score				0.0151* (0.0082)					0.0156* (0.0082)	
total trade as percentage of GDP					0.0007 (0.0005)					0.0006 (0.0005)
broad money					-0.0003 (0.0006)					-0.0002 (0.0006)
inflation (GDP deflator)					-0.0000** (0.0000)					-0.0000** (0.0000)
bank crisis during the period					0.0056 (0.0117)					0.0045 (0.0134)
currency crisis during the period					-0.0023 (0.0097)					0.0023 (0.0110)
debt crisis during the period					-0.0095 (0.0162)					-0.0080 (0.0154)
Period FE	X	X	X	X	X	X	X	X	X	X
Country FE	X	X	X	X	X	X	X	X	X	X
N	247	224	247	238	225	247	224	247	238	225
Number of countries	35	34	35	35	33	35	34	35	35	33
First stage F stats	16.495	14.469	16.389	16.459	13.647	7.385	7.909	6.590	7.732	5.259

Note: Each observation is a country-period. The dependent variable is the period average real per capita GDP growth rate. All columns are estimated using 2SLS.

Columns 1 through 5 use actual crossings as instrument. Columns 6 through 10 use predicted crossings as instrument. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported.

Table 3.9: Model Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: IV is $Crossing_{is-2}$	excl. last xing	excl. last 2 xing	multiple crossings	last crossings	small islands	non IDA
Aid_{is-1}	0.0413** (0.0172)	0.0451** (0.0186)	0.0257*** (0.00991)	0.0241** (0.0112)	0.0273*** (0.00929)	0.0299** (0.0120)
<i>Wild cluster bootstrap-t p-value</i>	0.050	0.022	0.008	0.024	0.006	0.012
y_{is-1}	-0.0194 (0.0337)	-0.0200 (0.0366)	-0.0717*** (0.0263)	-0.0440 (0.0320)	-0.0396 (0.0251)	-0.0194 (0.0259)
<i>Wild cluster bootstrap-t p-value</i>	0.642	0.712	0.016	0.350	0.250	0.510
Period FE	X	X	X	X	X	X
Country FE	X	X	X	X	X	X
N	193	151	208	247	225	220
Number of countries	28	21	30	35	32	31
First stage F stat	8.635	20.06	18.48	19.04	14.39	11.97
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV is $Crossing_{is-2}^{synth}$	excl. last xing	excl. last 2 xing	multiple crossings	last crossings	small islands	non IDA
Aid_{is-1}	0.0580** (0.0294)	0.0883 (0.0570)	0.0248* (0.0138)	0.0258 (0.0159)	0.0324*** (0.0124)	0.0382** (0.0172)
<i>Wild cluster bootstrap-t p-value</i>	0.044	0.000	0.022	0.008	0.008	0.020
y_{is-1}	0.00745 (0.0493)	0.0387 (0.0861)	-0.0736** (0.0332)	-0.0411 (0.0406)	-0.0308 (0.0289)	-0.00346 (0.0345)
<i>Wild cluster bootstrap-t p-value</i>	0.930	0.632	0.020	0.508	0.358	0.904
Period FE	X	X	X	X	X	X
Country FE	X	X	X	X	X	X
N	193	151	208	247	225	220
Number of countries	28	21	30	35	32	31
First stage F stat	3.401	1.905	6.438	8.158	7.958	5.946

Note: Each observation is a country-period. The dependent variable is the period average real per capita GDP growth rate. Sample restrictions are marked in the short handle in each column. All columns are estimated using 2SLS. Panel A uses actual crossings as the instrumental variable. Panel B uses predicted crossings as the instrumental variable. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported.

Table 3.10: External Validity

	(1)	(2)	(3)
Crossing sample = 1	0.00669 (0.00669)	0.00672 (0.00698)	
Crossing sample = 1 ×			
1 st quartile of y_{it-1}			0.03078 (0.00606)***
2 nd quartile of y_{it-1}			0.01327 (0.00868)
3 rd quartile of y_{it-1}			-0.00049 (0.01047)
4 th quartile of y_{it-1}			0.00739 (0.01018)
y_{it-1}	0.00799 (0.00656)	0.00940 (0.13174)	0.01848 (0.14883)
y_{it-1}^2		-0.00012 (0.01133)	-0.00085 (0.01294)
Year FE	X	X	X
N	1303	1303	1303

Note: Each observation is a country-year. The dependent variable is annual log per capita real GDP growth. There are 78 countries that were ever eligible for IDA between 1987 and 2010 as well as the 35 countries in our baseline sample. The key variable of interest is a dummy variable indicating whether the country belongs to the crossing sample. The sample consists of country-year observations between 1987 and 2010 that have per capita GNI level below the IDA threshold. Year fixed effects are controlled. Log real GDP per capita in the last year and its quadratic terms are included in the regressions. Standard errors are reported in parentheses, clustered at the country level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.11: IDA Threshold and Aid with Quadratic Log Initial Income Levels

	(1)	(2)	(3)	(4)	(5)
	IDA	DAC	NDAC	MLA	ODA
$Crossing_{is-2}$	-2.960	-1.020	-2.747	-1.027	-0.940
	(1.264)**	(0.272)***	(1.725)	(0.420)**	(0.243)***
<i>Wild cluster bootstrap-t p-value</i>	0.012	0.000	0.096	0.002	0.000
y_{is-1}	-30.28	-4.133	-28.67	-15.12	-4.462
	(21.24)	(3.995)	(17.17)	(8.215)*	(3.120)
y_{is-1}^2	1.562	0.194	1.724	0.909	0.211
	(1.464)	(0.277)	(1.147)	(0.555)	(0.217)
Country FE	X	X	X	X	X
Period FE	X	X	X	X	X
N	247	247	247	247	247

Table 3.12: First Stage - Policies as Outcome Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Freedom house		Word Bank		inflation		Crisis	
	civil liberty	political rights	CPIA z-score	broad money	(GDP deflator)	bank	currency	debt
<i>Crossing_{is-2}</i>	0.1255	0.2955	0.1032	0.2007	-21.2578	0.0028	0.0482	0.0427
	(0.2527)	(0.4039)	(0.1867)	(2.6788)	(100.4773)	(0.0404)	(0.0489)	(0.0292)
<i>Wild cluster bootstrap-t p-val</i>	0.606	0.466	0.548	0.904	0.856	0.918	0.306	0.06
<i>y_{is-1}</i>	-0.0573	-0.1061	-0.2203	7.7313	483.0089	0.0668	0.0229	0.0617
	(0.1675)	(0.2603)	(0.1624)	(8.6056)	(385.9596)	(0.0642)	(0.0744)	(0.0511)
<i>Wild cluster bootstrap-t p-val</i>	0.706	0.622	0.132	0.544	0.462	0.260	0.774	0.242
Country FE	X	X	X	X	X	X	X	X
Period FE	X	X	X	X	X	X	X	X
N	255	255	244	232	255	255	255	255
# of countries	35	35	35	33	35	35	35	35

Note: Each observation is a country-period. The dependent variable for each column and its source are indicated on top of the column. See Appendix Table 3.16 for description of these variables. All columns are estimated using 2SLS. Log population one period lag is included as a covariate. Standard errors in parentheses, clustered at the country level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported.

Table 3.13: First Stage - Policies as Outcome Variables

	(1)	(2)	(3)
Aid_{is-1}	0.0281***	0.0297***	0.0342**
	(0.0100)	(0.0105)	(0.0148)
<i>Wild cluster bootstrap-t p-value</i>	0.006	0.008	0.028
Polynomials of y_{is-1} included	1	1,2	1,2,3
Period FE	X	X	X
Country FE	X	X	X
Observations	247	247	247
# of Countries	35	35	35
First stage F statistic	16.495	14.935	10.560

Note: Column 1 replicates Column 3 of Table 3.5. Column 4 and Column 5 adds quadratic and cubic terms of y_{is-1} . Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported.

Table 3.14: Functional Form of Aid

	(1)	(2)
Aid_{is-1}^*	0.573** (0.264)	0.525** (0.251)
<i>Wild cluster bootstrap-t p-value</i>	0.044	0.060
y_{is-1}	-0.0211 (0.0475)	-0.0381 (0.0384)
<i>Wild cluster bootstrap-t p-value</i>	0.804	0.454
Period FE	X	X
Country FE	X	X
Observations	247	247
# of Countries	35	35
First Stage F statistic	8.062	4.111

Note: $Aid_{is-1}^* = \sum_{k=3}^5 (ODA_{it-k}/GNI_{it-k})/3$ is the period average aid to GNI ratio in the last period. There are 212 country-period observations from 35 countries. 2SLS estimator is used in both columns. In both columns, ΔAid_{is-1}^* is instrumented with $\Delta Crossing_{is-2}$. Δy_{is-1} is also treated as endogenous in Column 2 and y_{it-8} is used as an additional instrumental variable. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported.

Table 3.15: Effects of Aid on Investment

<i>Dep var (Inv/GDP)_{is-1}</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Main Specification</i>	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
<i>Aid_{is-1}</i>	-0.0166 (0.0150)	-0.0103 (0.0158)	0.0580* (0.0350)	0.0490 (0.0333)	0.0542 (0.0511)	0.0831 (0.0624)	0.0719 (0.0465)
<i>Wild cluster bootstrap-t p-value</i>	0.374	0.560	0.086	0.120	0.308	0.164	0.116
<i>y_{is-2}</i>	-0.0638 (0.0392)	-0.0962*** (0.0339)	0.0308 (0.0514)	0.0194 (0.0523)	-0.0503 (0.0409)	-0.0298 (0.0461)	-0.0378 (0.0399)
<i>Wild cluster bootstrap-t p-value</i>	0.172	0.006	0.588	0.728	0.244	0.486	0.512
Period FE	X	X	X	X	X	X	X
Country FE	X		X	X			
First differenced		X			X	X	X
IV			X	X	X	X	X
IV predicted				X			X
IV first differenced					X		
Equation first differenced		X			X	X	X
N	206	171	206	206	171	171	171
Number of countries	34	34	34	34	34	34	34
First atage F statistic (Kleibergen-Paap Wald)			9.223	9.714	10.94	4.724	10.51

Note: Each observation is a country-period. The dependent variable is the period average investment to GDP ratio. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. p -values from the wild cluster bootstrap-t procedure are reported. See text for more details.

Table 3.16: Construction and Sources of Key Variables

Variable	Notation	Source*	Description
Period	τ_s		Each period consists of 3 consecutive years. The first period is from 1987 to 1989.
IDA threshold		WB	Denoted in current US dollars, available since 1987.
Crossing 2 periods earlier	$Crossing_{i,s-2}$		Country i crossed the IDA cutoff for the first time in the sample at least two periods earlier.
Foreign aid	$Aid_{i,s}$	WDI/DAC	$Aid_{i,s} = \sum_s (ODA/GNI)/3$. Total net Official Development Aid (ODA) in current US dollars is from the DAC. GNI in current US dollars is from the WDI.
Initial income level	$y_{i,s-1}$	WDI	Real per capita GDP in 2000 constant US dollars in the second year of the last period, y_{it-4} .
Real per capita GDP growth	$g_{i,s}$	WDI	Denote real per capita GDP for country i in year t as y_{it} , annual real per capita GDP growth is $\ln(y_{it}) - \ln(y_{it-1})$. Period real pre capita GDP growth is the mathematical average of annual real per capita GDP for years in the period.
Aid by donor	$Aid_{j,i,s}$	DAC	Donor groups (j) include IDA, DAC countries, non- DAC countries, and multilateral agencies (MLA) except for IDA.
Investment		WDI	Gross capital formation as ratio of GDP. Investment in a period arithmetic average of annual gross capital formation as ratio of GDP.
Population ⁺		WDI	Population
Primary school enrollment ⁺		WDI	Gross primary school enrollment ratio. It is the total enrollment in primary education, regardless of age, expressed as a percentage of the population of official primary education age.
Trade ⁺		WDI	Measured as merchandise trade as percentage of GDP.
Money supply ⁺		WDI	Broad money as percentage of GDP
Inflation ⁺		WDI	GDP deflator (percentage annual)
Crisis ⁻		WB	Dummy variables indicating whether there is any bank, currency, or debt crisis during the years within the period.
Bank crisis ⁻		WB	Whether the country experiences a bank crisis.
Currency crisis ⁻		WB	Whether the country experiences a currency crisis.
Debt crisis ⁻		WB	Whether the country experiences a debt crisis.
Political rights ⁺		FH	Freedom House political rights indicator. It ranges from 0 to 7, with a higher number indicating less political rights.
Civil liberties ⁺		FH	Civil liberties indicator. It ranges from 0 to 7, with a higher number indicating less civil liberty.
Bureaucratic quality ⁺		ICRG	0-4, with a higher number indicating less risk in bureaucratic quality.
Corruption ⁺		ICRG	0-6, with a higher number indicating less risk in corruption.
Rule of law ⁺		ICRG	0-6, with a higher number indicating less risk in rule of law.
Ethnic tension ⁺		ICRG	0-6, with a higher number indicating less risk in ethnic tension.
CPIA z-score ⁺		WB	Public sector management and institutions cluster average. 1=low to 6=high.

* WB is short for the World Bank; WDI is short for the the World Development Indicators from the World Bank; DAC represents the OECD Development Assistance Committee. FH is short for the Freedom House. ICRG is short for International Country Risk Guide.

⁺ averaged within each period. ⁻ summed over each period.

4 CHAPTER 3: INDUSTRIALIZATION FROM SCRATCH: THE PERSISTENT EFFECTS OF CHINA'S "THIRD FRONT" MOVEMENT

4.1 Introduction

Industrializing from an agricultural economy is an inevitable process of modern economic development. How industrialization takes place and whether there is a role for active policy interventions are thus important questions facing many developing countries.

While some theories argue that industrialization is a natural process that follows increasing productivity in the agricultural sector (Rostow, 1990; Schultz, 1953; Lewis, 1954), others advocate the role of active development policies that aim at fostering the nascent industrial sector. They argue that development in the agricultural sector per se does not automatically lead to industrialization: an economy can be trapped in a low-income equilibrium and fail to embark on a self-sustaining path of industrialization which ends in a high-income equilibrium. A poverty trap may stem from coordination failures among different sectors (Rosenstein-Rodan, 1943; Murphy et al., 1989), the lack of supply and demand linkages that support an industrial system (Hirschman, 1958), or low levels of knowledge accumulation (Hausmann and Rodrik, 2003). Subsidies to the nascent industrial sector can generate positive externalities that benefit other economic activities. If the spillover effects are strong enough, temporary subsidies can kick start a virtuous cycle and permanently push an economy out of the low-income equilibrium.

The idea that temporary subsidies to the industrial sector have long-run impacts on economic development has motivated various subsidies to industrialization, in forms of investment in physical and human capital, and support in technological adoption. While theoretically appealing, empirical tests of the efficacy of these policies are difficult, because they are often implemented in selected places and industries with particular economic returns in mind. The selection problem makes it difficult to identify the true causal effects of those policies.

This paper estimates the causal effects of temporary subsidies to the nascent industrial sector on long-run economic development by evaluating a unique policy experiment in China, known as the "Third Front Movement". The Movement was launched in 1964 in preparation for potential military confrontations with the United

States and the Soviet Union. It aimed at creating self-sustaining industrial clusters in China's remote and mountainous hinterland, such that in the event of war, China could still produce and thus fight.⁸⁸ The scale was massive: in the heydays of the Movement, over half of the national industrial investment, in forms of large plants and infrastructure, was dedicated to the Third Front Region. In 1964, provinces containing the Third Front Region accounted for about 10 percent of the nation's total industrial output. This number rose to 37 percent in 1978.

For national defense purposes, where a Third Front project was located had little to do with the location's current economic conditions or future economic potentials. As airstrikes and nuclear attacks were the primary concerns, the sites for major projects were required to follow the criteria: "Dispersed from existing industrial clusters, hidden, and close to the mountains" (*fensan, yinbi, kaoshan*).⁸⁹ Consistent with the criteria, we show that, conditional on a set of initial conditions, the location choices of industrial investment during the Third Front period are uncorrelated with unobservable factors that affect long-run economic performance. The disconnect between industrial investment and economic development potential lends us credible evaluation of the causal effects of subsidized industrialization.

The Third Front Movement lasted for a decade before it gradually shrank in magnitude since in the mid-1970s, and eventually faded away with China's market reform in the urban sector in the 1980s. Since then, industrial plants established during the Third Front Movement, along with other State-Owned Enterprises (SOEs), were required to be responsible for their own profitability. Many of these plants failed due to fundamental deficiencies in location choices, management, and operations.⁹⁰

Yet the Movement fundamentally changed the economic landscape of some local economies. By the time China started its market reform, some previously agricultural local economies possessed a significant industrial presence with some of the most sophisticated plants in the country, while some other local economies with otherwise similar initial characteristics before the Movement remained pre-

⁸⁸The area, known as the "Third Front Region", was originally a national defense concept. It refers to the mountain-locked areas that are far from potential war fronts.

⁸⁹"Hidden" means that plants should be located in terrains such that they are hard to be found by the enemies. The criterion was taken seriously as in some extreme cases, entire factories were built in caves.

⁹⁰Of course, this is a simplification of China's market reform in the urban sector. Subsidies to some SOEs continued after the reform. In the empirical section, we provide robustness checks controlling for the post-reform growth of the SOE.

dominantly agricultural. This gives us quasi-exogenous variation in levels of initial industrial presence on the eve of the market reform. We compare prefectures in the Third Front Region that had initial high levels of industrial presences with those that had initial low levels in the two decades since the market reform. We evaluate whether, and through what channels, the initial industrial presence established during the Third Front Movement helped localities gain advantages in economic development that persisted in the post-reform era.

The Third Front Movement provides a unique natural experiment for evaluating the effects of temporary subsidies in the industrial sector on long-run economic development from a very low level of industrialization. First, it is a rare event in history when large amounts of investment were made to places without particular economic returns in mind. This enables us to disentangle the role of the investment from the role of unobservable fundamental characteristics of the locality. Second, the movement covers a wide range of industries, ranging from basic manufacturing such as textile to sophisticated industries such as machinery and automobiles. This gives us rich variation in types of industries to be subsidized, and opportunities to disentangle different mechanisms.⁹¹ Finally, we look at a within-country setting, which avoids omitted variable biases due to differences in institutional backgrounds or other nationwide policies, which are prevalent in earlier empirical work based on cross-country evidence.

We focus on a group of prefectures in the Third Front Region that were predominantly rural before the Movement. Our empirical approach boils down to comparing similar prefectures within the Third Front Region that received large amounts of industrial investment with those that received little. We measure the quasi-exogenous initial industrial presence by the per capita employment from the key industrial plants in 1985.⁹² The primary challenge for identifying a causal effect is that, in spite of a rich set of initial conditions we control for, there are still

⁹¹For example, one recent strand of literature has been particularly interested in the effects of transportation infrastructure on economic development (Banerjee et al., 2012; Faber, 2014; Donaldson, 2012).

⁹²The key industrial plants are the Large and Medium Sized (LMS) industrial plants listed in the 1985 Second Industrial Census. Conditional on the urbanization rate in 1964, the size of employment from these plants are good proxies for the magnitude of Third Front investment for two reasons. First, Third Front industrial projects were overwhelmingly large and were considered “key” to the national economy. So they should be included in the directory. Second, 1985 was the year before the market reform in the urban sector. Waves of bankruptcies of SOEs and mass layoffs did not take place until the 1990s. So the size of employment still reflects the magnitude of investment.

unobservable factors about the locality that affect both the magnitude of the Third Front investment it received and its potential economic performance in a market economy. In particular, the concern is that the government might have additional information about the location's economic potential and select places to host large industrial projects based on the information unobservable to the econometrician.

We test the validity of this concern by conducting a case study in which additional information about the location choice regarding one particular Third Front plant is available. By comparing historical documents on the plans for Third Front projects and the actual locations of these projects, we find that this plant was originally planned in another location. We compare three locations to illustrate the effect of receiving this plant and test the importance of potential unobservable characteristics. First, we have a treated prefecture where the plant was actually located. Second, we have a natural comparison prefecture where the plant was originally planned. Third, we construct a comparison prefecture based on a rich set of observable initial conditions using the synthetic control method (Abadie et al., 2010). The synthetic comparison prefecture and the natural comparison prefecture contain different sets of information. The synthetic comparison incorporates information from all the observable initial characteristics, while the natural comparison contains all the information available to the decision maker at the time of the Movement, some of which may be unavailable to the econometrician. If the unobservable characteristics play an important role in the location choice and these characteristics affect long-run economic development, we would expect the synthetic comparison to have different economic performance in the post-reform period from those of the natural comparison. On the contrary, we find that both initial characteristics and subsequent economic outcomes of the two comparison prefectures are very similar. Using a permutation-like method as in Abadie et al. (2010), we find that the differences are not statistically significant either. This suggests that the rich set of initial conditions we control for covers a sufficient portion of information available to the decision makers. The unobservable characteristics that played a role in government's location-choice decisions, if any, do not matter for long-run economic development.⁹³

The case study lends confidence in the validity of the assumption that location choices of Third Front projects are independent of unobservable factors once we control for initial conditions. Yet the case study may fall short on coverage and

⁹³Unfortunately, we could find only one case such that the eventual location of the project was different from the originally planned location.

representativeness of the overall effect. We therefore estimate the average treatment effect by pooling all the prefectures in the sample. To ensure the right functional form of the covariates, we try variations of the specification. All these approaches yield quantitatively similar results.

We find that places that had larger industrial presence on the eve market reform persisted to have higher levels of industrialization. The effect is mainly accounted for by a larger non-state-owned sector, which was essentially non-existent during the command economy era but had been growing rapidly after the reform. The effect is economically sizable and statistically significant. Our baseline result shows that having one more worker in key industrial plants in 1985 increases 1.4 more workers in the non-state sector in 2004. Put another way, the difference between the 75th percentile and the 25th percentile in initial per capita industrial employment explains about 80% of the 75th-to-25th difference in employment from the non-state industrial plants twenty years later. A prefecture with higher initial industrial presence also had significantly higher per capita employment in the service sector, resulting in a much higher urbanization rate. It also had a better educated workforce, as measured by the percentage of adults with a middle school diploma or a college degree.

We further investigate the mechanisms through which temporary subsidies to industrialization sustained further urbanization and industrialization. Initial industrial subsidies brought in by the Third Front plants might have promoted local economic development in the post-reform era in several ways. The first is a linkage channel: industrial firms may generate demand for goods and services from the local economy, and provide inputs to other related industries. The second is the human capital channel: knowledge and human capital brought in with the industrial investment facilitates local entrepreneurs to start new businesses, and stimulate startups in the related sectors. Third, roads and other infrastructure are needed to be built to facilitate an industrial plant. The infrastructure in turn benefits other economic activities in the local economy.

We test these hypotheses. Prefecture-industry level analysis shows that input-output linkages account for about a quarter of the total effect. To check whether the effects of initial industrial presence take place through some intermediate factors, we control for various economic conditions in the 1980s along with the size of initial industrial employment. We show that human capital and infrastructure can explain

little of the effect.

Finally, we test whether the higher urbanization rates in subsidized prefectures came from “pulling” local agricultural workers out of farmland, or simply from “stealing” urban workers from other local economies. Unsurprisingly given the restrictive cross-region migration, we find that increased urbanization rate is mainly accounted for by workers transitioning from the rural sector to the urban sector within a prefecture.

This chapter contributes to the literature on evaluating the long-run effects of industrial investment in less developed economies. The paper most closely related to ours is Kline and Moretti (2014a), which studies the long-run effects of the Tennessee Valley Authority (TVA) in the United States. The TVA program is a typical case of the Big Push strategy (Rosenstein-Rodan, 1943; Murphy et al., 1989), which advocates simultaneous investment in a wide range of sectors to solve the coordination failure. In contrast, the Third Front Movement involved investment in some particular industries in scattered locations. The Third Front Movement was never intended to promote balanced development of the local economy. Indeed, before the market reform, many Third Front plants were isolated little societies and had little economic interaction with the local economies around them.

This chapter is also related to papers that study the determinants of geographic distribution of economic activities. This strand of literature intends to test whether the geographic distribution of economic activities is unique. Exploiting the quasi-random nature of bombings during wars, Davis and Weinstein (2002), Davis and Weinstein (2008) and Miguel and Roland (2011) suggest that locations of economic concentrations are deterministic: even serious damages during a war could not alter a region’s long-run growth trajectory. In contrast, Rauch and Michaels (2013), Severnini (2012), Bleakley and Lin (2012) and Redding et al. (2011) find evidence supporting multiple equilibria: loci of economic activities can change due to historical events or policy interventions. By showing that the Third Front Movement had long-run impacts on the spatial distribution of industries, this chapter adds a piece of evidence supporting the existence of multiple equilibria.

To the best of our knowledge, this chapter is also the first to quantitatively evaluate the long-run impacts of the Third Front Movement, adding to a rapidly

growing literature on China’s regional development.^{94,95} Despite a caution on its overall efficiency, we find that the Movement had long-run positive impacts on local economies that received large amounts of industrial investment. Given the cost of mobility and heterogeneous preference in locations, our findings imply potential positive welfare gains for local residents, suggesting the relevance of the tradeoff between equity and efficiency in regional development policies (Ziliak, 2012).

The rest of the chapter is organized as follows. Section 4.2 introduces the background of the Third Front Movement. Section 4.3 describes our data and sample. Section 4.4 presents a case study. Section 4.5 shows the baseline results. Section 4.6 investigates some possible mechanisms. Section 4.7 concludes.

4.2 Background

In 1964 China was a very poor country with a per capita annual GDP only one fortieth of that of the United States. Its economy was predominantly agricultural: agriculture accounted for 40% of the GDP while over 80% of the population lived in rural areas. The country’s limited industrial capacity was concentrated in a handful of cities in the eastern part of the county, particularly the Northeast and the East Coast. In a command economy system, the majority of the limited industrial investment went to existing industrial centers.

The Third Front Movement was a large-scale industrialization campaign in the remote western China. It was motivated by national defense considerations. National security conditions facing China deteriorated in 1964: The Gulf of Tonkin incident in August signaled an escalation of the Vietnam War, with America’s massive bombing in North Vietnam and increasing military presence in the Taiwan Strait threatening China’s East Coast. Since in the late 1950s, the Soviet Union had been threatening China’s northern border. China’s leaders worried that the concentrations of the country’s industrial capacity in the Northeast and the East

⁹⁴Naughton (1988) is the first to systematically document the Third Front Movement. However, his approach is qualitative. He concludes that the Movement was ill-designed and poorly executed, and had resulted in great amounts of waste. While noting that the Movement is likely to be highly inefficient, this chapter is agnostic about its overall cost-benefit analysis. Instead, we treat the large industrial investment during the Third Front Movement as given and investigate the long-run development paths of places receiving various amounts of investment.

⁹⁵Alder et al. (2013) and Wang (2012) study the regional economic growth of the special economic zones. This paper differs in two important aspects: first, we study the long-run consequences of temporary industrial investments initiated in the Mao Era; second, we are able to track the exact treatment, instead of a bundle of policies. See Yao (2014) for a recent review on this literature.

Coast were vulnerable in possible wars with the United States or the Soviet Union. The Third Front Movement was launched later that year as a response to these imminent threats. The name “Third Front” was first used as a national defense concept, which refers to the vast area far away from possible war fronts. The purpose of the Movement was to establish self-sustaining industrial clusters in the Third Front Region, such that in the event of war, China could still have the industrial capacity to support the fight even if it loses its industrial clusters in the east. The top map in Figure 4.1 shows the ruggedness of terrains by county.⁹⁶ The bottom map in Figure 4.1 shows the boundary of the Third Front Region. The Region includes the rugged areas to the west of the Beijing-Hong Kong line encircled by mountains but avoids Xinjiang, the Tibetan Plateau and other ethnically minority areas.

The scale of the investment was massive. In the heydays of the movement between 1965 and 1971, it is estimated that about 140 billion *yuan* worth of investment went to the Third Front Region, accounting for over a half of the total national investment during that period (Naughton, 1988). Figure 4.2 shows the investment per capita in provinces containing the Third Front provinces between 1955 and 2008, with the national average standardized at 1 in each year. Per capita investment in the Third Front Region was low relative to the national average before the early 1960s. It increased dramatically after 1964 and peaked in 1970, with a level of per capita capital formation about 1.5 times that of the national average. Investment then gradually declined after the mid-1970s and never achieved the same relative level to date. A decade of intensive investment changed the economic landscape in the Third Front Region. Before the Movement, this region was among the poorest. In 1960, the provinces in that area, with one fifth of the nation’s population, accounted for only 10 percent of its industrial output. This share rose to about 37 percent by 1977.⁹⁷

The Movement was organized in a command system, led by a central committee in Beijing. Investment focused on the industries that produce basic industrial materials and sophisticated manufacturing industries that produce machinery, electrical equipments, and electronic devices. Different places were designated to host

⁹⁶Ruggedness, or slope, measures the first gradient of the elevation. To calculate the average ruggedness in a county, we first divide counties into small equal sized squares, and then calculate the slope in each square as the ratio between the point with the highest elevation and the point with the lowest elevation. Finally, we take the average of the slopes in all squares. See Nunn and Puga (2012) and Dell (2010) for applications of slope.

⁹⁷Source: National Bureau of Statistics of China.

different industries. For example, Panzhihua in Sichuan Province was designated to become a large steel and non-ferrous metal production center. Counties around Chongqing were assigned to host conventional weaponry industry. Mianyang was made a base for nuclear industries. In haste and short of funding, many new plants were spin-offs from existing plants in other parts of the country. For example, the Second Auto Works in Shiyang had its initial machines and workers moved from the First Auto Works in Changchun, in the Northeast. To minimize the damage due to potential air strikes and nuclear attacks, the location choices of large industrial projects followed a guiding principle which stated that the plants should be “dispersed, hidden, close to the mountains, and when necessary, they should be in caves” (*fensan, yinbi, kaoshan, biyaoshi jindong*).

Table 4.1 summarizes the structural shift of industrial investment in the Third Front Movement. The table shows the characteristics of the key industrial plants by year of opening.⁹⁸ Columns 1 through 3 compare industrial plants that started operating between 1949 and 1964 with those started operating between 1964 and 1978, the Third Front period. The characteristics of the counties where these plants were located differ greatly. Only 16% of the industrial plants established during the first period were in the Third Front Region, compared with 41% of those that were established during the second period. Plants that started operating during the Third Front period are less likely to be located in a provincial capital city; they are more likely to be located in less densely populated counties with higher elevation and more rugged terrains. They are also less connected to the existing or planned railroads. These differences are all statistically significant. The characteristics of the plants that started operating in the two periods differ as well. Plants built in the Third Front period are less likely to be in light industries, and more likely to be in sophisticated manufacturing sectors. They also have higher capital per worker but lower output per worker. This is probably due in part to the shift of new plants to the machinery, electric, and electronic industries. Columns 4 through 6 show that even within the Third Front region, the patterns still hold: plants established during the Third Front period are located in geographically more difficult places.

Wars against the United States or the Soviet Union did not happen after all.

⁹⁸These plants are from a directory included in China’s Second Industrial Census, conducted in 1985. These plants in sum accounted for about half of the total industrial output in 1985. Other industrial plants at that time included smaller SOEs and Township and Village Enterprises (TVEs), which started to boom since in the late 1970s.

Investment tilted towards the Third Front Region gradually faded away since in the 1970s as relations between China and the United States improved. Market reforms of China's SOEs started in the mid-1980s when the firms were made responsible for their own profitability and, if they continued losing money, were allowed to go bankrupt. Many Third Front plants did particularly badly due to their flawed designs and irrational location choices. Nevertheless, the Third Front Movement was the first and only wave of industrialization before the market reform in the region. Third Front plants left some local economies with valuable stocks of physical capital, advanced technology, and skilled workers. These were scarce factors that entrepreneurs could make good use of in early stages of the market economy, and could help the local economy embark on a trajectory of economic development different from the one without the Movement.

4.3 Data and Sample

4.3.1 Data Sources

We compile a unique panel dataset of Chinese prefectures over 70 years with comprehensive information on geography, demography, and economic conditions, especially the industrial sector.⁹⁹ These variables are collected from various sources. Whenever possible, we construct our variables from first-hand micro level data from population and industrial censuses, or other large-scale nationally representative surveys.¹⁰⁰ We extract consistently defined variables from various sources. To our knowledge, the dataset we have compiled is among the most comprehensive at the prefecture level. In this subsection we describe the sources of main variables. Appendix A lists the details of the data sources used in the chapter.

County geographic characteristics, including elevation, slope, and access to water ways, are from GIS files available at the China Historical GIS database.¹⁰¹ Access to roads are from GIS maps for roads and railroads in 1962, 1980, 1990, 1995, 2005 and 2010 compiled by Baum-Snow et al. (2012). County boundaries are available for census years in 1953, 1964, 1982, 1990, 2000, and 2010. County

⁹⁹A prefecture is a jurisdiction that is smaller than a province but larger than a county. China was divided into 345 prefectures in 1985. Prefecture is a natural choice to represent a local economy. Each prefecture usually consists of one urban center and a few rural counties.

¹⁰⁰We are aware of the concern that macro level data in China are sometimes subject to manipulation of the politically motivated local officials. Manipulation of macro data may be correlated with local economic conditions. Using censuses and nationally representative micro-level surveys mitigates this concern.

¹⁰¹<http://www.fas.harvard.edu/~chgis/>

boundaries change from time to time. In order to define variables consistently, we convert county characteristics into 1982 county boundaries. Appendix B.1 describes how consistent counties are defined. These characteristics are then aggregated to the prefecture level.

Characteristics of the local industrial sector are from the National Industrial Survey of 1936, the Second Industrial Census in 1985, and the First Economic Census in 2004. Additional information is also extracted from county-tabulation of population censuses in 1982 and onwards, which include counts of employment by sector.

We construct demographic characteristics for consistently defined prefectures from county-tabulations of all six population censuses. These characteristics include population density, share of urban population, share of adult population with a middle school diploma, and share of adult population with a college degree.

4.3.2 Sample and variables

The geographic boundary of this study includes prefectures in the Third Front Region as shown in the bottom map of figure 4.1. The area encompasses Sichuan, Guizhou, Yunnan, and parts of Gansu, Hubei, Hunan, Guangxi, and Shaanxi. There are 89 prefectures in the Third Front Region.

Some Third Front investment went to places in the Region that were already industrial centers. Chengdu and Chongqing, two major cities in the Third Front Region, received large investment during the Third Front Movement. We exclude from the sample prefectures that were already urban centers before the start of the Third Front Movement. The reasons are twofold. First, the existing industrial centers may have some unobservable characteristics such that they have different development trajectories to begin with. Third Front investment in these existing industrial centers is more likely to be confounded with other factors. Second, this chapter focuses on the long-run effects of subsidies at early stages of industrialization. Additional industrialization in existing urban centers may have different effects.

We do not have direct measures for levels of industrialization at the prefecture level on the eve of the Third Front Movement. Instead, we use the 1964 urban population rate as a proxy.¹⁰² We drop prefectures that had urban population rate

¹⁰²Urban rate is a good measure of non-agricultural economic activity. In 1958 China implemented the *hukou* system, a residential registry system which essentially banned migration across regions. The only way to transit from a rural *hukou* to an urban *hukou* was through recruitment of new

greater than 10% in 1964.¹⁰³ We are left with 63 prefectures with average urban population rate in 1964 of about 6.3%.

We proxy the initial industrial presence due to the Third Front Movement by the size of employment from the key industrial plants as listed in the Second Industrial Census in 1985. Admittedly imperfect, this variable captures the quasi-exogenous part of initial industrial presence brought in by the Third Front Movement. First, the Third Front projects were overwhelmingly large-scale and were regarded as crucial for the nation's economy. Therefore, they are likely to be included in the directory.¹⁰⁴ Second, since we are using a sample of local economies that were predominantly rural in 1964, the industrial presence from large plants could only be added during the Third Front period. Third, 1985 provides a last snapshot of the urban economy before the comprehensive market reform took place, after which subsidies shrank and firms were required to be responsible for their own profitability. Before that, independent decisions of hiring and layoffs were almost impossible. Therefore, although large parts of the Movement ended in the 1970s, the size of the employment in these plants gives a reasonable measure of the magnitude of the Third Front investment.

The main outcome variables of interest include industrial employment as a share of population, and the share of urban population in the post-reform years. We are particularly interested in the development of the non-state industrial sector. We also look at population density, percent of adults with a middle school diploma and the percent of adults with a college degree.

industrial and public sector workers conducted by the government, or through marriage to a person with urban *hukou*. Only industrial, government and service workers were given urban *hukou*. The employment of government services (including the employees of the government and the Communist Party organizations, employees of the public service such teachers, doctors, retailers, etc) is largely proportional to the size of population.

¹⁰³The 10% cutoff is at the 75th percentile in the distribution of urban population rate in 1964 in the Third Front Region. We show later that our results are robust to a wide range of cutoff values in the initial urbanization rate.

¹⁰⁴One important limitation of the 1985 directory is that it does not include plants that were directly established by the military or plants that were considered to be sensitive to national security. Bachman (2013) compares the list from the 1985 industrial census with plants from industrial census publications at the provincial level, as well as lists of industrial plants from Military Industrial Census, and he finds that about 10 percent of key plants are not included in the 1985 list.

4.3.3 Summary Statistics

Table 4.9 shows the summary statistics. Panel A shows the employment from the key industrial plants in 1985 by sector. For every 100 residents, on average there were 0.57 industrial workers in the non-mining sector employed by the key industrial plants in 1985. The distribution of investment to different locations had a fat tail. The prefecture that received medium investment had 0.22 workers in the non-mining industrial sector per 100 residents, while the prefecture at the 90th percentile has per capita employment of 1.3. About a third of these workers were employed in sectors that produce basic industrial materials such as iron, steel, petroleum, chemical products, and electricity. About two fifths were employed in plants that produce machinery, transportation equipments, electric and electronic devices, instrument equipment and apparatus. These were the sophisticated manufacturing by the standard of China's industrial development at that time. The remaining were employed in light industries such as those produce textile and food. One thing worth pointing out is that since industrial plants were purposefully chosen to be dispersed across locations in the Third Front Movement, the correlations in employment across these sectors are low. The correlation between employment in the sophisticated manufacturing industries and that in the materials industries is -0.04, and the correlation between employment in the sophisticated industries and light industries, highest among all three pairwise correlations, is 0.3.

Panel B shows the initial conditions of sample prefectures before the Third Front Movement. The median prefecture in the sample had a population density of 90 people per square kilometer (4.5 in log value). After censoring the urban rate at 10%, the urban rates among the sample prefectures are rather evenly distributed. The minimum is 2.8 and the maximum is 9.9, with the mean and median is 6.3. These prefectures are in remote and rugged areas. The median distance to the existing or planned railroads was about 55 kilometers (4 in log value), the median distance to the nearest provincial capital is 220 kilometers (5.4 in log value).¹⁰⁵ The median elevation is over 1 kilometer, and the average slope is 4 degrees (the national average is less than 1 degree). Except for a few cases, these places had no industrial presence in 1936.

These prefectures were significantly more developed in the early 2000s. As

¹⁰⁵For these distances we first calculate the distance between the centroid of the county to the railway line or the centroid of the provincial capital. We then calculate the average prefecture-level distance by using county population as weight.

Panel C shows, population grew by an average of 60 percent over the 36 years and the average urban rate rose to 20 percent. In 2004, there were about 2.1 non-mining industrial workers who were employed in the non-state sector, and over 90% of them worked in firms that did not exist in 1985. 11 percent of the adult population worked in the tertiary sector. Middle school education was common (34 per 100 adults), especially among the younger generations, although college education was still rare (1.6 per 100 adults).

4.4 A Case Study of Shiyan

Ideally, we would like to compare post-reform economic development between prefectures that received large amounts of industrial investment during the Third Front to prefectures that had similar initial conditions, but received little investment. We observe and control for a rich set of initial prefecture conditions before the Third Front Movement, including urban rate, population density, historical industrial presence, transportation infrastructure, and geographic characteristics. In addition, the historical background and the principles of location choices during the Third Front Movement suggest that these observable characteristics indeed cover the factors that were taken into consideration when location choices were made. Nevertheless, we cannot rule out the possibility that other things may matter. In particular, government might have access to pieces of information about the localities that are not observable to the econometrician, and these pieces of information were used in deciding locations and matter for long-run economic outcomes. In that case, the conditional independence condition does not hold and the Ordinary Least Squares (OLS) regression will fail to yield causal estimates.

One possible identification strategy is to take advantage of the decision process itself. If a couple of locations were shortlisted for each Third Front project and one was finally chosen, the runner-ups serve as good comparisons as they arguably have similar characteristics that matter for the consideration of the government, no matter whether they were observable to the econometrician. The final choice of the winner is arguably determined by less important factors, and is likely not to affect the outcomes of interest. The average differences between the treated and the runner-ups within a case is likely to be as good as causal. This approach is adopted by Greenstone et al. (2010) in estimating the agglomeration effect of large

manufacturing plants in the United States.

However, by comparing the planned locations for large industrial plants in the Third Front Movement and the locations where these plants were actually located, we are only able to find one case where the actual location is different from the original plan.¹⁰⁶ So our approach is different from Greenstone et al. (2010), who pool many cases together and estimate the average treatment effect. Instead, we conduct a case study and use it to show that conditional on the observable characteristics we can control for, additional local characteristics that are unobservable to the econometrician either did not matter in government’s decision of location choices, or did not affect outcomes of interest. After establishing the conditional independence, in the next section, we run pooled OLS regressions with the full sample controlling for the same set of initial conditions and estimate the average treatment effects.¹⁰⁷

4.4.1 A Natural Comparison

The case we study here is Shiyan in Hubei Province, home to the Second Auto Works of China. The plant was a prominent example of the scale of the Third Front Movement and its decision process. It was the only large investment Shiyan received during the entire Third Front Movement (and naturally, the only large investment in the entire command economy period). In 1985, the plant employed more than 65,000 workers and was the second largest automobile manufacturer in China. It was fair to say that Shiyan was a city built around the plant. According to the original plan, the site was chosen in Xiangxi, a similarly remote and mountainous prefecture in western Hunan Province. After some debate and a few inspection trips to the region, Shiyan was finally chosen for its proximity to a planned rail road.¹⁰⁸

Xiangxi and Shiyan had very similar geographic characteristics and economic conditions before the Movement. They are equally far away from the coast. Shiyan is about 500 kilometers to the north of Xiangxi. They have similar terrains, as is shown in panel B of table 4.2, average slope is 4.7 degrees in Shiyan and 3.2 degrees in Xiangxi. Log elevation is 6.6 versus 6.3. They had almost the same distance

¹⁰⁶A list of planned large projects is from “The Outlines for the Third Five-Year Plan” (*guanyu disange wunian jihua anpai qingkuang de tigang*), published in 1965.

¹⁰⁷We make an important assumption here that the potential selection mechanism on the unobservable characteristics is the same for all cases.

¹⁰⁸In fact, Xiangxi is also connect to a planned rail road. By 1980, both prefectures were connected by railroads.

to the nearest provincial capital and rail line. The population density was almost the same (4.48 versus 4.50 in log values). Most importantly, Shiyang and Xiangxi both had very low levels of industrialization and urbanization before the Movement: both places had no industrial plants in 1936, and in 1964, only 5 percent of the population in Shiyang was urban, compared with 7.9 percent in Xiangxi.¹⁰⁹

The Third Front Movement dramatically changed this parity. In 1985, the Second Auto Works employed thousands of industrial workers, while during the same period Xiangxi received little investment and had remained largely agricultural. By 1985 Shiyang was significantly more industrialized than Xiangxi. According to panel A of table 4.2, there were 2.6 workers in the key industrial plants per 100 residents in Shiyang, compared with only 0.06 in Xiangxi. 9.2 percent of the adult population were employed in the industrial sector in Shiyang, compared with 5 percent in Xiangxi.

The advantage of Shiyang in industrialization and urbanization persisted in the post-reform era. Panel C compares the economic outcomes in early 2000s between the two locations. Table 4.10 shows that Shiyang was more urbanized and had larger shares of employment in non-agricultural sectors. Besides having a much larger state-owned sector, Shiyang's advantage over Xiangxi mainly came from the non-state sector. Non-state plants accounted for about 60 percent of the total industrial employment in Shiyang. The non-state sector in Shiyang, relative to population, was more than twice as large as that in Xiangxi. Shiyang's workforce was much more skilled: its share of college-educated workers was almost twice as large as Xiangxi.

4.4.2 A Synthetic Comparison

In this subsection we construct a statistical comparison based on observed initial conditions. The statistical comparison serves two purposes. First, by comparing Shiyang to this statistical comparison, we can test whether the effects of having the plant are statistically significant. Second, by comparing Xiangxi to this statistical comparison, we can test whether there are other important unobservable conditions that confound our results.

We use a statistical case study approach known as the synthetic control (Abadie et al., 2010). The synthetic control of Shiyang, the treated prefecture denoted as i , is constructed as the weighted average of potential comparison prefecture (j 's) such

¹⁰⁹Note that we cannot evaluate the statistical difference between Shiyang and Xiangxi.

that the distance of a vector of initial conditions (\mathbf{Z}) between the treated prefecture and the synthetic prefecture is minimized. The weights are restricted to be bounded between 0 and 1 and sum up to 1. Formally, we find the weights according to

$$\begin{aligned}
 w_j &= \underset{w_j}{\operatorname{argmin}} \left\| \mathbf{Z}_i - \sum_{j \in J} w_j \mathbf{Z}_j \right\|, & (4.1) \\
 \text{s.t.}, & w_j \in [0, 1], \forall j \\
 & \sum_i w_j = 1
 \end{aligned}$$

We pick potential comparison prefectures whose employment to population ratios from key industrial plants are below the 75th percentile in the distribution. There are 40 such potential comparison prefectures.¹¹⁰ \mathbf{Z} includes log distance to existing and planned rail in 1962, log distance to the nearest provincial capital, log average elevation, average slope, urban population rate in 1964, log population density, industrial employment in 1936 per 1964 population, and per capita employment in the mining sector.

The effect of the treatment is therefore the difference between that of the treated prefecture, y_i , and that of the synthetic prefecture, $y_i^0 = \sum_j w_j y_j^0$, we denote the treated effect as $d = y_i - y_i^0$.

We test the importance of the unobservable factors by comparing the natural comparison and the synthetic comparison. Since Xiangxi, the actual alternative site for the plant, incorporates all the information relevant for the location choice, both observable and unobservable to the econometrician, if the synthetic control and Xiangxi exhibit similar economic performance both before and after the Movement, it will be a compelling evidence that whatever confounding unobservable factors there might be, they are unlikely to bias our results. More formally, denote the economic condition in Xiangxi as y_s , we test whether the difference between Xiangxi and synthetic Shiyang, denoted as $d_0 = y_s - y_i^0$, is statistically different from zero.

Statistical inference for both d and d_0 are obtained by permutation-like tests (Abadie et al., 2010). Specifically, among the group of potential comparison prefec-

¹¹⁰These potential comparison prefectures are called the “donor pool”. We also exclude Xiangxi from the donor pool. If Xiangxi is included, it only contributes a small fraction to synthetic Shiyang, and the results are not changed. The reason that Xiangxi only contributes a small fraction to synthetic Shiyang, we speculate, is that there are potentially other prefectures that serve as better comparison to Shiyang than Xiangxi. Indeed, since all the potential comparison prefectures were underdeveloped in 1964, many had very similar attributes. Consistent with this argument, the results are also robust to a wide range of selections of the donor pool.

tures, we assign a placebo treatment to each of the potential comparison prefecture, and construct the synthetic control from the remaining comparison units in the same way as we construct synthetic Shiyang. Denote the synthetic control for each prefecture b from the donor pool as y_i^b , we calculate differences parallel to d and d_0 as $d^b = y_i - y_i^b$ (and $d_0^b = y_s - y_i^b$). With all the d^b 's (d_0^b 's), we rank them along with the original estimate, d (d_0), and use the percentile of d (d_0) in the distribution as the p -value of d (d_0).

4.4.3 Comparing Shiyang and Xiangxi with Synthetic Shiyang

Column 3 of table 4.2 shows the characteristics of synthetic Shiyang before the Third Front Movement (panel B), after the Third Front Movement but before the market reform (panel A), and after the reform (panel C). Shiyang and synthetic Shiyang had very similar characteristics before the Third Front Movement, but differences were obvious in the two later periods. Column 4 reports the p -values of these differences. For characteristics in 1964, the p -values are often close to 1, and never below 0.05. Synthetic Shiyang had little industrial presence in 1985. According to panel C of table 4.2, the difference in non-state non-mining industry employment per capita between Shiyang and Synthetic Shiyang is large and statistically significantly at 5% level. Table 4.10 shows that Shiyang and its synthetic control exhibit substantial differences in most economic conditions in early 2000s, except for population density. The differences are economically large as well. For example, the urban population rate in Shiyang is 10 percentage points higher than that in its synthetic control.

In contrast, the economic conditions for synthetic Shiyang are very close to those of Xiangxi throughout the periods. Column 5 of table 4.2 and Column 5 of table 4.10 report the p -values of the differences between Xiangxi and synthetic Shiyang. These differences are small and statistically insignificant.¹¹¹

For a set of consistently defined outcomes, we investigate how Shiyang developed over time compared with Xiangxi and synthetic Shiyang. Figure 4.3 shows these evolutionary paths for rate of urbanization, log population density, and share of adult population with a college degree. For each outcome variable, the top graph shows the evolution of the values for Shiyang (in solid blue lines), Xiangxi (in solid red

¹¹¹The only exception is that Xiangxi is a little farther away from the nearest provincial capital city than synthetic Shiyang.

lines), and synthetic Shiyan (in blue dashes). The lower graph shows the associated p -values of the differences between Shiyan and synthetic Shiyan (in blue dots), and the associated p -values of the differences between Xiangxi and synthetic Shiyan (in red diamonds). Starting in 1982, Shiyan began to have higher urban rate and college rate than synthetic Shiyan. The difference is statistically significant, and remains at around the same magnitude over time. We find no statistically significant effect on log population density. In contrast, the differences between Xiangxi and synthetic Shiyan in essentially every dimension remained small and statistically insignificant throughout the whole period.

4.5 Long-Run Effects of Temporary Subsidies to Industrialization

4.5.1 Baseline Specification

In this section we estimate the effects using the full sample of 63 prefectures in the Third Front Region that had small urban sectors before the Movement. The econometric approach we adopt is a simple OLS regression controlling for prefectures' initial conditions before the Movement:

$$y_{it} = \beta_{0t} + \mathbf{X}_i \cdot \gamma_t + \beta_{1t} emp_i + \varepsilon_{it}, \quad (4.2)$$

where y_{it} is some outcome of interest of prefecture i in year t . \mathbf{X} is a vector of initial geographic and economic conditions. emp_i is the size of employment from key industrial plants in non-mining sectors in 1985, divided by 1982 population. \mathbf{X}_i includes the same set of initial conditions used in constructing the synthetic control. We estimate equation 4.2 separately for each outcome variable y_{it} , so each regression is cross-sectional and each observation is a prefecture. The years of the outcomes include 1982, 1990, 2000, and 2010, all of which are years a population census was conducted, and 2004, the year when the first economic census was conducted. To facilitate our discussion, we focus on two key outcome variables: the prefecture's urban population in 2000 and the size of non-mining industrial employment from the non-state sector in 2004, both variables are divided by 1982 population. We loosely call the two outcome variables "urban rate" and "non-state employment per capita". Other outcomes of interest include log population density, the share of adult population with a college degree, and the share of adult population with a middle

school degree.

The identification assumption is that, conditional on the observable initial characteristics, \mathbf{X}_i , the assignment of the Third Front investment across locations is uncorrelated with unobservable factors, ε_{it} . Although the identification assumption is not directly testable, there is evidence that it is likely to hold. First, in the case study in the previous section, we show that the synthetic comparison prefecture based on the observable initial conditions exhibits very similar outcomes in the post-reform period with those of the comparison prefecture chosen by the government back in the 1960s, which presumably incorporates information both observable and unobservable to the econometrician. This means that our control variables cover most important factors that determined the location choices of the Third Front industrial projects. Second, the documentation of the Movement indicates that location choices of investment were largely based on national defense considerations. We are able to control for characteristics that were key determinants according to the criteria for location choices. Third, we focus on a sample of prefectures that were poor and rural at the time of the Movement so they were homogeneous in terms of the stage of economic development.

We estimate three versions of equation 4.2. The purpose of trying various versions is to make sure that we specify the functional form of \mathbf{X} correctly such that the conditional independence assumption is satisfied, yet not controlling for so many covariates as to lose power given our fairly small sample. In the first approach, we include \mathbf{X} in linear forms. In the second model we also include the quadratic forms of these covariates to control for residual variation. The second approach involves more than a dozen covariates, a sizable number compared to our sample size of 63.

Another concern is that even with many covariates, there is still no guarantee that the covariates will be balanced between those that received large amounts of investment versus those received little. The third model uses a reweighing approach. We divide the sample into two groups. The first group includes 15 prefectures that had the highest industrial presence as measured by per capita non-mining employment in key industrial plants in 1985, and the second group includes the remaining prefectures in the sample. We assign those in the first group each has a weight equal to 1 and reweight the second group following a decomposition estimator as in Kline (2011). The weights are constructed such that the two groups have the same distribution of covariates. We then estimate equation 4.2 using Weighted Least

Squares without any controls.

4.5.2 Local Economic Outcomes in the Post-Reform Years

Industrial development in the non-state sector

Table 4.3 reports the results on per capita non-state non-mining industrial employment in 2004. The first three columns in panel A report the results of estimating equation 4.2 using three different functional forms of \mathbf{X} . All yield very similar results. In general, the estimates show that one additional industrial worker from the key industrial plants in 1985 leads to 1.3-1.4 more industrial workers in the non-state non-mining sector in 2004. The coefficient is statistically significant at the 1% level. The effect seems to be linear: in column 4, where we include the quadratic form of the per capita industrial employment in 1985, we find a small and statistically insignificant coefficient associated with the quadratic term while the coefficient associated with the linear term is unchanged.¹¹²

The effect is meaningfully large. The 25th and the 75th percentiles in the distribution of per capita employment from key industrial plants are 0.06 and 0.67, respectively. So the 25th-to-75th percentile difference in initial industrial presence in 1985 predicts a difference in per capita non-state non-mining employment twenty years later by $(0.67 - 0.06) \times 1.3 = 0.79$, which is about a 50% increase from the mean of the dependent variable (1.57), or 70% of the 25th-to-75th percentile difference in the outcome variable (0.67 to 1.81). Since the non-state sector was essentially non-existent in 1985, this result can be interpreted as the causal effect of having a larger industrial presence on the eve of the market reform on the growth of the non-state sector. Although the state sector had been in relative decline since the market reform, prefectures that received large industrial investment during the Third Front remained their advantage in industrialization more than two decades since the end of the Movement.

Over the years of reform, some SOEs were privatized. One concern is that the higher levels of non-state sector industrial employment in prefectures with historically high SOE concentrations are a mechanical result of privatization. To alleviate this concern, in panel B we focus our attention on non-state firms that were estab-

¹¹²If an economy is mired in a low-income equilibrium and only large-scale investment can push it out of the poverty trap, we would expect the effect to be convex in the size of the investment. The largely linear effect suggest the absence of poverty traps.

lished after 1985. Presumably, when a firm changes ownership it does not change its first year of operation, thus focusing on firms that started after 1985 rules out privatized SOEs. The effects become smaller as the mean dependent variable becomes smaller, but the key message remains strong: higher initial industrial presence leads to higher levels of industrial development in the non-state sector twenty years later.¹¹³

Other post-reform economic outcomes

Table 4.4 reports other economic outcomes in 2000. Table 4.4 only reports one specification, that with initial conditions entering linearly. Other specifications yield similar results, as reported in table 4.11.

We find that higher industrial presence on the eve of market reform resulted in higher urbanization, higher human capital stock, and higher population density. One percentage increase in per capita industrial employment from key industrial plants in 1985 results in a 5.4 percentage point increase in urbanization rate (with a mean of 20% in 2000). This increase in urbanization rate can be roughly decomposed into 2.3 percentage points higher in industrial employment and 2 percentage points higher in the service sector employment. Population density is about 6 percent higher. The share of adult population with a middle school diploma was 5 percentage point higher (15 percent higher than the mean), and the share that had a college diploma was 0.6 percentage higher (39 percent higher than the mean).

Robustness

The baseline sample includes prefectures in the Third Front Region with a 1964 urban population rate lower than 10 percent. The cutoff rate, which excludes about a quarter of the prefectures in the region, is arbitrary. There is a tradeoff in deciding the cutoff level of the urban population rate. On the one hand, we worry that investment that went to the existing industrial centers are more likely to be endogenous, so we want to focus on prefectures that had a small urban sector before the Movement. On the other hand, restricting our attention to a sample of

¹¹³One may still be concerned that firms may report a new first-year-of-operation once they are privatized or restructured. We alternatively focus on the non-state firms that were established after 2001, because by then the vast majority of the SOE privatization had completed. The results (not reported) remain economically meaningful and statistically significant.

prefectures with very low urban population ratio reduces sample size and weakens statistical power.

We test the robustness of our results with regard to samples using different cutoff values of initial urban rate. We use cutoff values in urban population ratio ranging from 6 percent to 30 percent with a step of 1 percentage point, and repeat the baseline estimation for each cutoff value. If the endogeneity problem is correlated with initial urban population ratio, we would see estimates change with the cutoff values. Robust results to the different cutoff values within a reasonable range will be evidence that endogeneity is not a serious concern.

Figure 4.4 summarizes the results for selected outcomes reported in table 4.4 and table 4.11. The blue solid line shows the coefficients associated with initial industrial employment as a function of the cutoff value. The blue dotted lines show the 95% confidence intervals. The red dotted line (with scale on the right) shows the number of observations in the sample for each cutoff. There were 45 prefectures in the Third Front Region with urban population rate below 8 percent in 1964, and only 8 prefectures had urban rate above 30 percent. The point estimate becomes less precise as the cutoff value gets lower and the sample size gets smaller, but the magnitude is relatively stable throughout the space of the cutoff values. This pattern is universal for all eight outcome variables we investigate. There is also no discernible changes in the size of confidence intervals as we increase the cutoff value, probably because the sample size grows more slowly with higher cutoff values. These findings lend us more confidence that bias in estimation due to potential omitted variables is not a serious concern.¹¹⁴

Effects over time

So far we have shown that investment in the Third Front period had long-run effects on the local economy's performance more than two decades after the end of the subsidy. But those results do not indicate whether the differences between the local economies receiving different levels of subsidized industrialization were growing or converging over time.

¹¹⁴We did a host of other robustness checks. First, we repeat the analysis using a sample of counties. With many more observations in the sample and many received no investment, we are able to conduct a matching estimator. The disadvantage with the county sample is that a county is too small to be regarded as a self-sustaining local economy. Second, we drop prefectures that have large reservations of natural resources. Third, we drop prefectures where large military plants were located. Results are robust throughout.

To look at the effects over time, we extract from population censuses consistently defined outcomes in population density, urban population rate, and share of adult population with a college degree. We estimate the baseline regression for each outcome variable in each census year, report the coefficient associated with initial industrial employment in 1985, and connect these dots over time to see the effects over time.

Figure 4.5 shows the results. In each graph, the solid line shows the coefficients. The dotted lines show the 95 percent confidence intervals. The advantages of local economies that had higher initial industrial presences remained strong.¹¹⁵ China's economy grew more than fourfold between 1980 and 2010, and the persistent effects suggest that that temporary industrialization subsidies helped set local economies onto a self-sustaining development path.

The continuing role of the SOEs

We have interpreted the long-run effects of the employment in key industrial plants in 1985 as the effect of the “initial condition” at the start of the market reform. However, as these plants did not immediately disappear right after the market reform, it is not clear whether the effect is due to the initial condition on the eve of market reform, or is simply driven by the fact that these plants continued to receive subsidies from the government, and some of the subsidies trickled down to the non-state sector in the local economy.¹¹⁶

To distinguish between these two alternative explanations, we divide our sample prefectures into two groups, one group includes prefectures that had their state-owned sector growing relatively fast (or declining relatively slowly) relative to the other group. If the second explanation is correct, we would expect the effect of initial industrial presence to be greater in the group of prefectures where the state-owned sector grew relatively fast.

In the baseline specification we include a dummy indicating whether the prefecture has its SOE sector growth between 1985 and 2004 above sample median,

¹¹⁵The trend, if any, shows that the effects became stronger. Although since the average levels of these outcome variables increased substantially over this period, the advantage in relative terms had been stable or in slight decline.

¹¹⁶In fact, a small number of the Third Front plants remained sizable, state-owned, and relevant for the local economies to date.

and an interactive term of this dummy and the per capita initial industrial employment in 1985. For brevity, we only report the effects on per capita non-state sector employment in 2004 (and that from new plants that opened after 1985). Table 4.5 reports the results. We find that the interactive term is small and statistically insignificant, and the coefficient associated with initial employment is unchanged. This finding is inconsistent with the explanation that the effects were driven by the government's continued support of the remaining SOE sector.

4.6 Mechanisms

Temporary industrial subsidies may promote long-run local economic development in several ways. First, knowledge and human capital brought in with the industrial investment facilitates local entrepreneurs to start new businesses in related sectors. Second, the Third Front investment may have also brought in investment that improved local infrastructure, which benefits other economic activities in the local economy. Third, the industrial firms may generate demand for goods and services from the local economy, and provide inputs to other related industries.

We investigate these channels here. First, we test whether there are some intermediate factors that drive the results. In other words, we control for higher initial industrial presence itself does not affect long-run economic development. It takes effect by improving some other conditions that in turn have significant impacts. Second, we test the idea of forward and backward inductions of industrial production (Hirschman, 1958). Third, we investigate where the newly urbanized people came from. If the increasing urban population was simply due to migration of urbanized workers from other prefectures, the Third Front investment simply affected the distribution of urban population. As Gottlieb and Glaeser (2008) has argued, if the positive effects of a location based policy is obtained at the cost of other locations, the net effect at the national level may well be negative, as government inventions distort the efficient allocation of businesses.¹¹⁷

¹¹⁷That does not mean, however, if the positive effects of the Third Front investment were all captured by the local economy, the Movement is efficient at the national level. National welfare might well have been higher if the investment went to places with better economic development potentials. However, the discussion of efficiency at the national level is beyond the scope of the current version of this paper.

Intermediate factors

If the effect of initial industrial presence is mediated through some intermediate factors, we would expect that the effect of initial industrial presence to shrink once we control for potential intermediate factors in the 1980s along with the size of initial industrial employment. In table 4.6 we report results from specifications where include as control various economic conditions in the 1980s. The outcome variable in all columns is the per capita employment in 2004 from non-mining non-state industrial plants that were established after 1985. Column 1 replicates the result in column 2 of table 4.4. Column 2 includes province dummies and the estimate does not budge. It suggests that province-level policies are not driving the results. Column 3 shows that the estimate remains the same after accounting for log distance to the railway. It suggests that improved infrastructure does not explain much of the effect, either. Column 4 includes variables capturing the human capital of the local economy, including the share of adult population with a college degree and the share with a middle school diploma. The coefficient associated with initial industrial presence dropped by about 10%. Column 5 includes log population density in 1982 and the estimate is the same as that in column 1. Column 6 adds all the covariates from column 1 to column 5, the coefficient still remains similar.¹¹⁸

Finally, in column 7, we divide employment in the key industrial plants in 1985 into three broad industries: sophisticated manufacturing industries, material industries, and light industries. All three sectors have positive and significant effects on the post-reform employment levels, and it seems that employment in the material industries generate the largest spillovers.

Industrial linkages

The original idea of Hirschman (1958) on unbalanced economic growth is that firms generate demand for upstream industries and provide input for downstream industries. Certain industries generate more such input-output linkages than others, so temporary subsidies on these industries have positive effects on the overall economy. Fostering an advanced industry generates spillovers of knowhow in the

¹¹⁸Some intermediate factors may only take effect in complement with initial industrial presence. For example, both transportation infrastructure and spillovers from existing firms are important. To test potential complementary effects, we interact initial industrial presence with these intermediate factors, we find small and statistically insignificant coefficients.

economy, and prepares specially trained labor. At early stages of development, the positive spillover effects from the tradable sector are likely to dominate the local congestion effects. We quantify how much effects we find can be explained by industrial linkages, both within the same industry and through input-output linkages.

Specifically, input linkages refer to spillover effects in which industry j' uses the products of industry j as inputs. Similarly, an output linkages refer to the spillover effect in which the products of industry j' are used as inputs for industry j . Let matrix IN be a $J \times J$ lower matrix with element $\rho_{jj'}^{input}$, where $\rho_{jj'}^{input}$ is the share of industry j' 's purchase in industry j 's total sale. Let matrix OUT be a $J \times J$ lower matrix with element $\rho_{jj'}^{output}$. $\rho_{jj'}^{output}$ is the share of industry j' 's sales that is bought by industry j . The pairwise input share, $\rho_{jj'}^{input}$, and the pairwise output share, $\rho_{jj'}^{output}$ are both from 1995 input-output tables published by the National Bureau of Statistics (NBS).¹¹⁹ Let R be a $I \times J$ matrix with element $R_{ij'}$. $R_{ij'}$ is the 1985 output from key industrial firms in industry j' in prefecture i , divided by 1982 population. Let I be an $J \times J$ identity matrix. Define the $I \times J$ input-linkage matrix $INLINK = (I - IN)^{-1} \cdot R$ with element $inlink_{ij}$ capturing the input linkages to industry j in prefecture i . Similarly, define the $I \times J$ output-linkage matrix $OUTLINK = (I - OUT)^{-1} \cdot R$ with element $outlink_{ij}$ capturing the output linkages to industry j in prefecture i .

To estimate the effect due to input-output linkages, we estimate the following equation. Each observation is a prefecture-industry. We estimate the effects of input-output linkages on non-state sector employment using the following specification:

$$y_{ijt} = \beta_{0t} + \mathbf{X} \cdot \gamma_t + \beta_{1t}inlink_{ij} + \beta_{2t}outlink_{ij} + \varepsilon_{ijt}, \quad (4.3)$$

Estimating the effects on non-mining employment from non-state firms established after 1985, columns 1 of table 4.7 replicates the baseline results using the prefecture level sample. Column 2 shows the result using the prefecture-industry level sample. Adding one industrial worker from the key industrial plants in 1985 raises 0.142 more workers in 2004 in the non-mining non-state sector from newly established firms. This effect is statistically significant at 10% level, and the size of the effect is about 15 percent of that in the first column. Column 3 shows that input

¹¹⁹The NBS input-output tables provide output flows between any pair of 24 2-digit industries. We group them to 19 industries to match the 1985 data. The results are similar when using any input-output table (available for 1987, 1990, 1992, 1995, 1997, 2002).

and output linkages also explain variations in employment by industry. A simple back-of-the-envelope calculation shows that input-output linkages collectively explain about 25 percent of the total effect.

Urbanization or migration

A natural question given our results so far is where did the newly urbanized workers come from. One explanation is that the initial industrial presence stimulated other non-agricultural businesses, generating labor demand in the urban sector, such that local peasants, lured by higher wages and better living standard in the city, acquired human capital and started working in the local urban center. Another explanation is simply that the Third Front industrial plants attracted higher quality workers from nearby regions. As we have discussed earlier, distinguishing the two channels has important policy implications.

To gauge the contribution of cross-prefecture migration in accounting for the results, we use micro-level data of the 2000 census, which has some information on migration. Specifically, the 2000 census asked respondents whether they were born in the current county of residence or the current province of residence. It also asked respondents' residence 5 years ago (in 1995). If the former residence was not in this community, a followup question specifies whether it was outside the current county of residence, or outside the current province of residence. Unfortunately, in both situations, we do not know whether a person had moved from a different prefecture. Nevertheless, since a prefecture is a jurisdiction between a county and a province, we can construct cross-county migration rates and cross-province migration rates that jointly give a bound for cross-prefecture migration rates.

We construct two sets of variables for urban residents in 2000: the first set indicates whether an urban resident was originally from a different county (or a different province); the second set indicates whether an urban resident migrated from a different county (or a different province) within the past 5 years. Therefore, the percentage of urban residents who were originally from a different prefecture is bounded between the percentage of urban residents who were originally from a different county (upper bound), and the percentage who were originally from a different province (lower bound). Similarly, the percentage of urban residents who migrated from a different prefecture within the past 5 years is between the percentage of urban residents who migrated from a different county within 5 years

(upper bound), and the percentage who migrated from a different province within 5 years (lower bound).

We focus on urban population growth between 1982 and 2000. The percentage of urban residents who migrated from a different prefecture between 1982 and 2000 are bounded between the percentage of urban residents who were originally from a different prefecture (which includes cross-prefecture migration in all years up to 2000, thus the upper bound), and the percentage of urban residents who migrated from a different prefecture in the past 5 years (between 1995 and 2000, a subset of years between 1982 and 2000, thus the lower bound).

We first calculate the urban population growth rate between 1982 and 2000. We divide the changes in total urban population between the two years by 1982 population ($\times 100$). Column 1 of table 4.8 shows that increasing per capita key industrial employment in 1985 by one percentage point increases the urban population growth by 4.3 percentage points. Columns 2 to 5 report the effect of per capita industrial employment from key industrial plants in 1985 on migration. With a one percentage point increase in per capita industrial employment in 1985, the percentage of urban residents who were originally from a different prefecture is bounded between 0.53 and 0.96 (denote $m_1 \in [0.53, 0.96]$). The percentage of urban residents who migrated from a different prefecture in the past 5 years is bounded between 0.1 percent and 0.18 percent (denote $m_2 \in [0.1, 0.18]$). The percentage of urban residents that migrated from a different prefecture is thus $m_3 \in [\min\{m_2\}, \max\{m_1\}]$, which in turn is a subset of $[0.1, 0.96]$. Notice that the maximal possible value m_3 can take is less than a quarter of the total urban growth rate estimated in column 1. Therefore, we can conclude that increased urban population is mainly driven by the urbanization of local rural workers.¹²⁰ This conclusion, after all, is not surprising due to the strict migration restrictions across jurisdictions.

¹²⁰It is possible that an urban resident who migrated from other prefectures earlier has a larger effect on urban rate due to natural population growth: the offsprings of urban residents automatically have urban *hukou*. National population grew 26% between 1982 and 2000. The natural growth rate of the urban population is also lower than that of the rural population. Even multiplying the national average population with m_1 and m_2 gives a bound for the value of m_3 that is still a small fraction of the urban growth rate found in Column 1. We are working on an alternative accounting of migration rate based on a formal demographic model taking into consideration of birth rate and death rate.

4.7 Conclusions

How to transform an agricultural economy into an urbanized and industrialized economy has been an important question facing development economists and policymakers. Previous studies on the effect of industrial investment on long-run economic growth suffer from the endogeneity of investment. This chapter studies the long-run effects of temporary subsidies to industrialization in predominantly agricultural local economies by exploiting a unique policy design known as the Third Front Movement. The design of the Movement allows us to separate the effects of locational fundamentals from the effects of temporary subsidies. Combining evidence from a case study, the synthetic control approach, pooled OLS and Weighted Least Squares estimates, we find consistent evidence that more than two decades after the subsidies had ended, prefectures that had received large-scale industrial investment persisted to have higher rates of urbanization and industrialization, and better educated workforce. In particular, these results are mainly driven by a fast growing non-state sector economy.

We explore a few possible channels through which an initial advantage in industrialization reinforces itself over time. We first show that the results cannot be explained by the lingering effects of the state-owned sector: neither the privatization of some SOEs nor the continued subsidies to the remaining SOEs could explain a faster industrial development in the local non-state sector. The results cannot be explained solely by factors such as human capital, infrastructure, or population density. We find some evidence of industrial linkages, which are able to explain about a quarter of the effect. Finally, we provide evidence that increased urban rate mainly comes from workers transitioning from the rural sector to the urban sector within the local economy, not from people migrating from other jurisdictions.

Our results have strong policy implications. On the one hand, we show that a government can permanently transform a rural local economy into an industrialized one by providing temporary subsidies to the nascent industrial sector. On the other hand, these policy interventions are likely to be inefficient since the local fundamentals may not be suited for particular industries. In fact, we find in our data that firms in the Third Front Region are consistently less efficient than those in other parts of the country.¹²¹ However, as mobility is costly, place-based subsidies to industrialization in remote rural areas can transform the local economy

¹²¹Our next step is to evaluate the effect on the total factor productivity (TFP) of the firms.

and promote living standards. This helps reduce regional inequality. When trade costs and migration costs are high, these subsidized local economies, though inefficient at the national level, serve as local magnets for production and consumption. Thus the effect on overall welfare compared to a scenario in which firms are located in places with the highest productivity will be unclear if we factor in the tradeoff between efficiency and equity. Evaluating the overall welfare impact of the Third Front Movement in the long-run is part of our ongoing work.

4.8 Appendices

4.8.1 Appendix A. Sources of Data

Data sources that we use in this chapter include:

1. *1936 Industry Survey*. The 1936 industry survey is a firm level dataset conducted by the the Nationals Party (*Kuomingdang*) and is the only existing national industry survey prior to the Third Front Movement.¹²² The 1936 industry survey includes summary characteristics of capital, worker, product, machinery and output for each firm. The survey includes data on 146 cities (counties) that have significant industry presence, which covers all of China.¹²³ We aggregate the firm level data to get city (county) level industrial output and employment.
2. *1985 Industry Census*. We use the statistics from the 1985 Industry Census published by the National Bureau Statistics. Industry is defined in all surveys here to include mining, manufacturing and public utilities. We use mainly two volumes of the publication. The first volume lists all key (large and medium sized, LMS) industrial firms in 1985. Virtually all of them are state-owned. For each LMS firm, we observe its name, address, 2-digit industry, employment, output, capital, year of opening and main products. We use the year of opening and the location to determine whether a plant belongs to the Third Front Movement. The second volume we use lists county level number of establishment, industrial output, and employment for years 1980, 1984, and 1985. We use this second volume to construct the overall conditions of the

¹²²There is a census of industrial plants in 1950. However, disaggregated data at the county or industry level are not available to the public.

¹²³Except for the 3 provinces in the Northeast, which were controlled by the Japanese. These provinces are not in our sample either.

industrial sector right before the market reform.

3. *County tabulation of population censuses in 1953, 1964, 1982, 1990, 2000 and 2010.* These include all population censuses conducted in China after the establishment of the People's Republic in 1949. All county level tabulations include population. The availability of other information varies by censuses. Size of urban population is available in 1964, 1990, 2000, and 2010. Employment in the industrial sector is available since 1982. Population by education level is available since 1982. We also use the 1/1000 micro sample of the 2000 census for information on migration.
4. *Firm level panel data for "above scale" manufacturing firms from National Bureau of Statistics, 1998-2007.* "Above scale" is defined as annual sales above 5 million *yuan*. The data have comprehensive information on firms' ownership, labor and capital input, industry, output, and financial conditions.
5. *2004 Economic Census.* It includes the firm-level data on all 1.4 million industrial firms in China. For each firm, we observe its ownership, labor, capital, output, industry, and location.
6. *GIS maps* of county boundaries from China Historical GIS project. The GIS maps include boundaries of county-level jurisdictions in each year when population census is conducted. Area covered by each county is calculated from the polygons. Overlaying these polygons, we define counties with consistently defined boundaries.
7. Digital Elevation Model (DEM) data for China's geographic characteristics. The DEM map is overlaid with GIS maps with county boundaries to calculate a county's average elevation and ruggedness.
8. GIS maps for roads, railroads, and highways in 1962, 1980, 1990, 1995, 2005 and 2010. This dataset is collected in Baum-Snow et al. (2012) and is generously provided by the University of Toronto library.

A list of key variables used in the study and their sources are listed in Table A.

4.8.2 Appendix B. Details of Variable Construction

B.1 Consistently Defined Counties

County boundaries have changed over the decades. We create a geographically consistently defined counties defined by 1982 county boundary by overlapping GIS maps of county boundaries accompanied in population censuses in 1953, 1964, 1982, 1990, and 2000. We currently do not have GIS map for county boundaries in 2010. We manually match 2010 county characteristics to 1982 county boundaries.

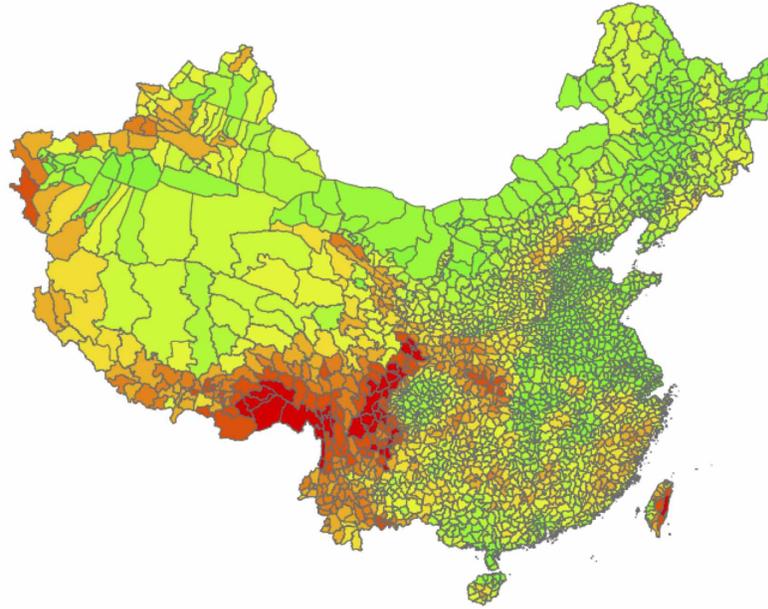
County geographic characteristics are directly calculated for the consistently defined county boundaries since geographic characteristics, such as elevation, ruggedness, distance to river, etc, are all continuous across locations. County characteristics from industry and population censuses are based on current county boundaries. We match counties to the nearest year when a GIS map of county boundaries is available, and then convert it to the 1982 boundary using a weighted average of intersecting counties. For example, say we have a 1982 county A whose territory belongs to county A_1 and county A_2 in 1990. If α share of the territory of county A is in A_1 and $(1 - \alpha)$ share of the territory of county A is in A_2 , then the characteristics of county A in 1990 are calculated as the weighted average of the characteristics of A_1 and A_2 , with the share of overlapping territory serving as weights.

B.2 Geographic Characteristics

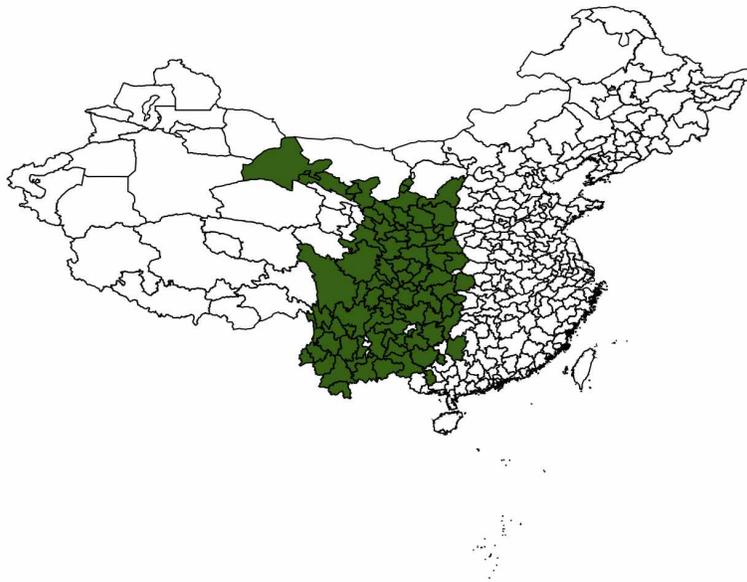
We construct two main variables for county geographic characteristics. The first variable is county average elevation. The second variable is county average slope. Both variables are constructed from a DEM (Digital Elevation Model) file. A DEM file contains elevation information of a point in space with fine resolutions. We first overlap the DEM file with the shape file containing county boundaries. The average elevation is calculated as the average elevation all points in the DEM data that fall within the boundaries of the county. To calculate the county average slope, the county polygon is first divided by 1 km by 1 km grids. For each grid, the slope is calculated as the maximal elevation divided by the minimal elevation from the points in the DEM file that fall in the grid. Slope for each grid is then averaged across all grids that fall in the polygon of the county. If a grid only partly falls within the county polygon, the slope of that grid is given the weight proportional to the percent of its area that intersects with the county polygon. Slope thus measures the ruggedness of the terrain. See Dell (2010) and Nunn and Puga (2012) for discussion

and application of the slope.

Figure 4.1: Third Front Region

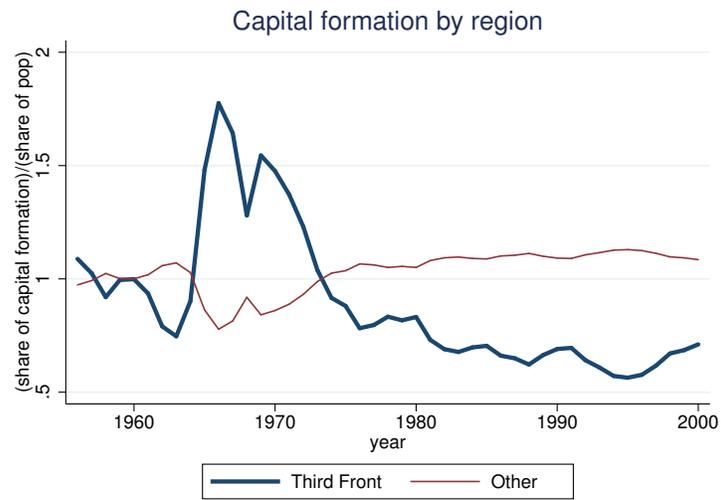


Third Front Region



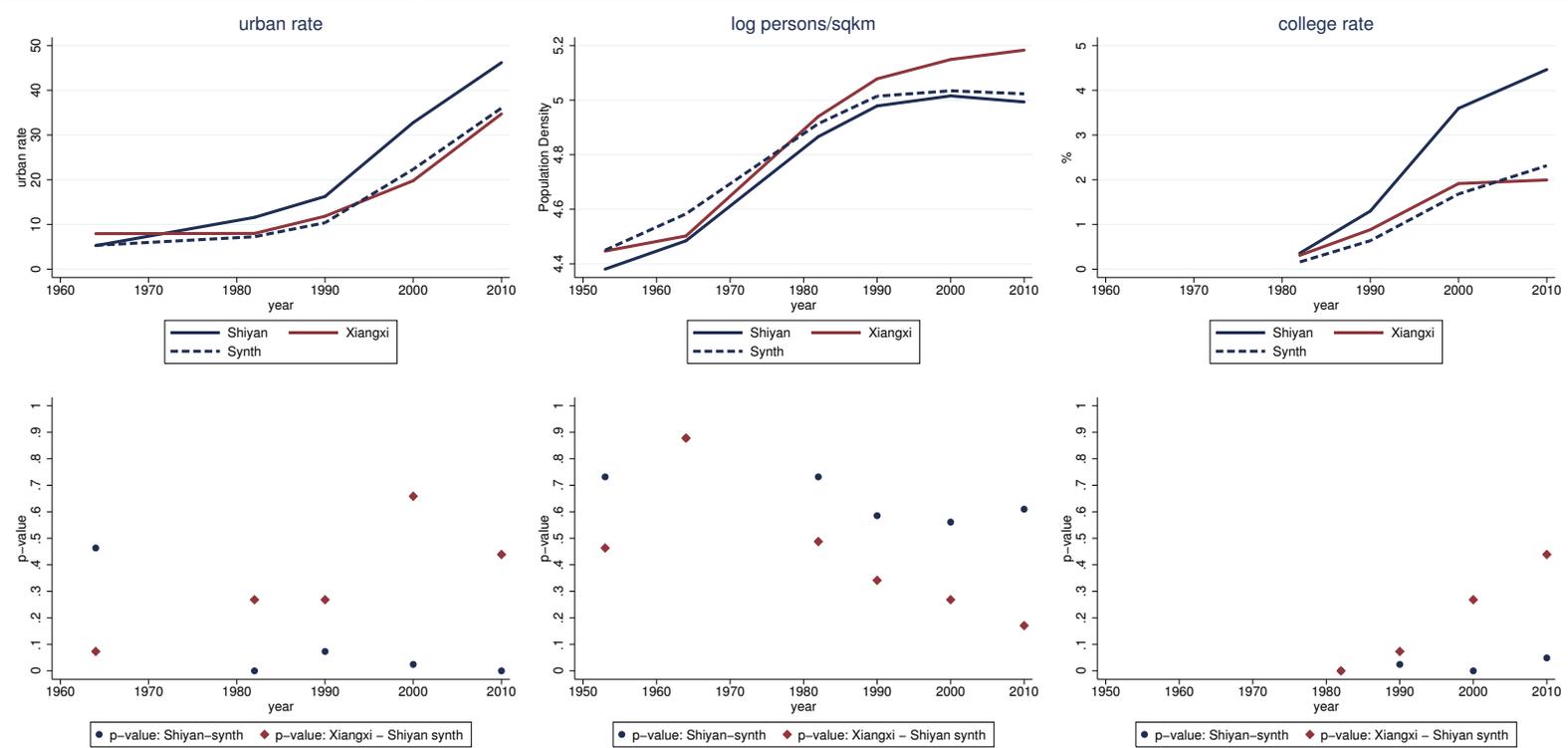
Note: The top map shows the average slope by county. Ruggedness of terrain increases from green to red. The source data for this map is from the China Historical GIS Database. The bottom map shows the boundaries of the Third Front Region, as is shown in Naughton (1988).

Figure 4.2: Intensity of Investment by Region



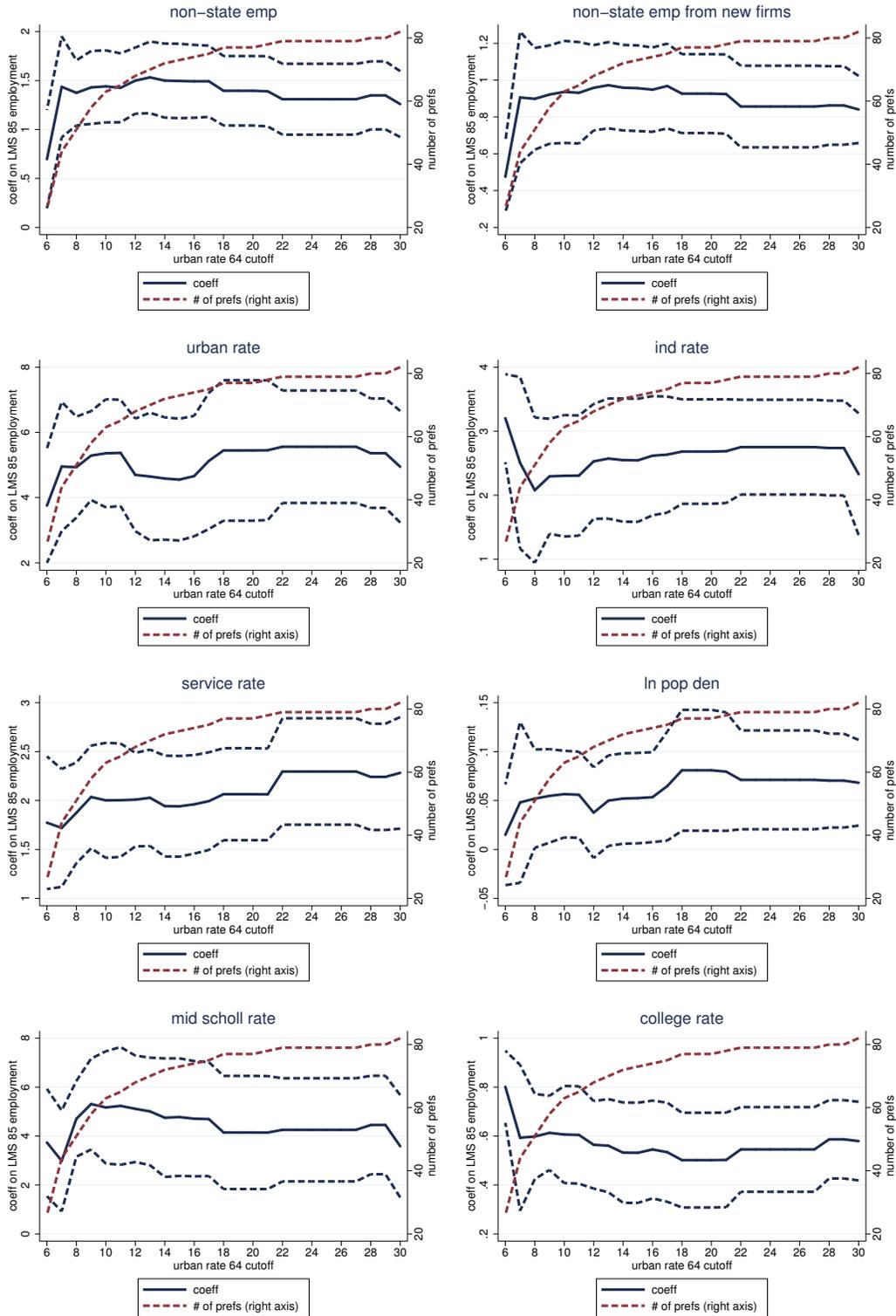
Note: Provincial investment and population data are from 60-year Statistical Summary constructed by the National Bureau of Statistics of China. The measure of investment intensity is the region's share of national investment divided by its share of national population.

Figure 4.3: Effects over Time - Shiyan and Its Comparisons



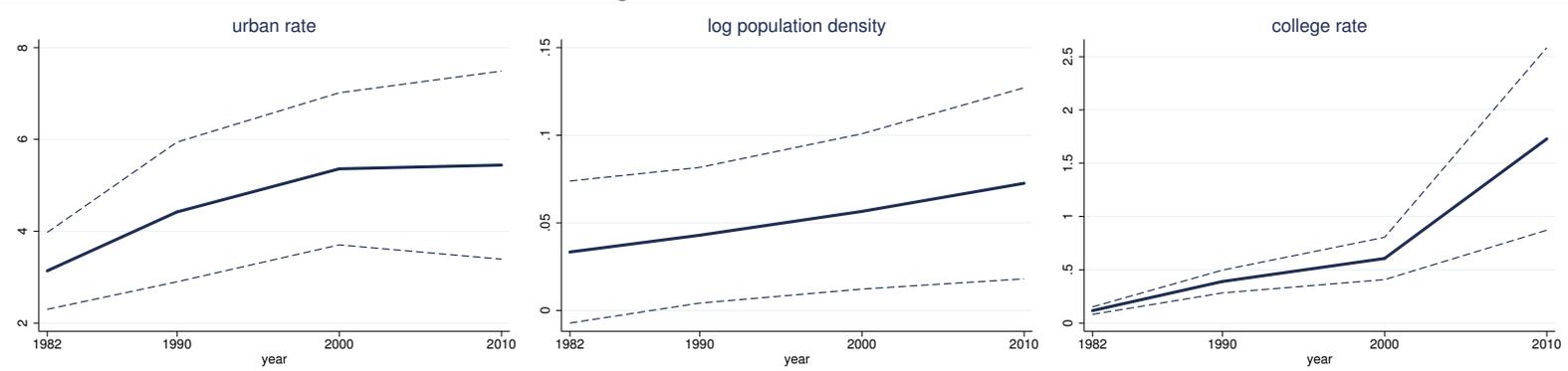
Note: Graphs in the top panel depict trajectories of outcomes in Shiyan (solid blue), Xiangxi (solid red), and Synthetic Shiyan (dotted blue) from 1964 to 2010. Graphs in the bottom panel show the p -values for the differences between Shiyan and the synthetic Shiyan (blue dots), and the differences between Xiangxi and the synthetic Shiyan (red diamonds). The outcomes are urban rate (Column 1), log population density (Column 2), and rate of population with college degrees (Column 3)

Figure 4.4: Varying Cutoff Urban Rate in 1964



Note: The baseline regressions are repeated with varying cutoff values of 1964 urban rate. Graphs show the coefficient with each cutoff value and its 95% confidence interval for indicated outcome variables.

Figure 4.5: Effects over Time



Note: Coefficients associated with non-mining LMS employment in 1985 per 1982 population ($\times 100$) over the years.

Table 4.1: Characteristics of Firms by Year of Opening

year of opening	nationwide			Third Front Region		
	1949-1964 (1)	1965-1978 (2)	diff. (3)	1949-1964 (4)	1965-1978 (5)	diff. (6)
<i>County characteristics</i>						
in Third-Front region	.164 (.37)	.411 (.492)	.247 (.012)			
in provincial capital city	.442 (.497)	.208 (.406)	-.235 (.011)	.336 (.473)	.154 (.362)	-.182 (.021)
average slope (ruggedness)	1.229 (1.418)	1.917 (1.76)	.689 (.044)	2.399 (1.966)	2.824 (1.976)	.425 (.097)
mean elevation (meters)	276.489 (464.138)	515.985 (629.93)	239.496 (15.406)	856.154 (737.519)	933.513 (715.589)	77.359 (35.784)
density (persons per km ² , '64)	2317.995 (4326.287)	810.062 (2029.033)	-1507.932 (76.21)	603.113 (525.61)	327.713 (334.777)	-275.4 (21.785)
distance provincial capital (km)	91.293 (114.448)	138.886 (122.196)	47.593 (3.154)	125.172 (114.271)	156.042 (110.561)	30.87 (5.537)
distance to 1962 rail (km)	8.325 (35.506)	23.628 (57.762)	15.303 (1.373)	17.945 (49.031)	32.723 (66.079)	14.778 (2.856)
log industrial output in '36	8.677 (8.542)	4.28 (7.027)	-4.397 (.197)	5.9 (7.038)	2.18 (5.035)	-3.72 (.302)
<i>Firm characteristics</i>						
capital per worker (10k '80 yuan)	1.854 (3.107)	2.571 (4.442)	.718 (.108)	2.446 (5.056)	2.747 (5.083)	.301 (.25)
output per worker (10k '80 yuan)	2.344 (3.834)	2.061 (2.583)	-.283 (.079)	2.097 (4.499)	1.736 (2.404)	-.361 (.179)
mining	.057 (.233)	.041 (.198)	-.016 (.005)	.087 (.281)	.034 (.182)	-.052 (.012)
light industries	.299 (.458)	.222 (.416)	-.077 (.011)	.278 (.448)	.216 (.412)	-.062 (.021)
power/Water	.068 (.251)	.061 (.239)	-.007 (.006)	.103 (.304)	.05 (.218)	-.053 (.013)
chemical	.138 (.345)	.155 (.362)	.016 (.009)	.106 (.308)	.106 (.308)	-.001 (.015)
ferrous and non-ferrous metal	.099 (.298)	.064 (.245)	-.035 (.007)	.094 (.292)	.064 (.245)	-.03 (.013)
machinery	.251 (.434)	.293 (.455)	.042 (.012)	.252 (.435)	.343 (.475)	.091 (.022)
electric and electronic	.088 (.283)	.164 (.371)	.077 (.009)	.08 (.272)	.186 (.39)	.106 (.016)
number of plants	4927	2049		809	842	

Note: Data from the large and medium-sized manufacturing plants from the Manufacturing Census of 1985. Year range indicates years in which the plant started operating. County characteristics are from various population censuses and manufacturing censuses. In column 1 and column 2, standard deviations are in parentheses. In column 3, standard errors are in parentheses.

Table 4.2: Shiyan and Its Comparisons

	(1)	(2)	(3)	(1)-(3)	(2)-(3)
	Shiyan	Xiangxi	Synth	p -value	p -value
Panel A: Emp from key industrial plants per 100 1982 residents					
non-mining industries	2.586	0.057	0.132	0.000	0.512
Panel B: Initial conditions					
log population density 1964 (count per sqkm)	4.484	4.502	4.583	0.878	0.878
urban rate 1964 (%)	5.261	7.934	5.328	0.463	0.073
log distance to existing and planned railway 1962 (km)	5.230	4.296	4.981	0.268	0.854
log distance to nearest provincial capital (km)	5.560	5.785	5.578	0.780	0.000
log average elevation (meters)	6.599	6.304	6.640	0.780	0.951
average slope (degrees)	4.669	3.216	4.678	0.512	0.976
industrial employment in 1936 per 10k 1964 resident	0.000	0.000	0.185	0.854	0.854
Panel C: Post-reform industrial outcomes (2000s)					
non-state non-mining industry emp per 100 1982 residents	3.052	1.291	0.971	0.024	0.220
from firms established after 1985	1.925	0.805	0.659	0.049	0.293

Note: Conditions for Shiyan are reported in Column 1. Xiangxi is the natural comparison for Shiyan, the prefecture that received the Second Auto Works during the Third Front Movement. Its conditions are reported in Column 2. Conditions for the synthetic control of Shiyan are reported in Column 3. Column 4 reports the p -values of the differences between Shiyan and synthetic Shiyan. Column 5 reports the p -values of the differences between Xiangxi and synthetic Shiyan. The p -values are obtained using the permutation-like method. Panels A, B, C report conditions in 1964, 2000, and industrial composition in 2004.

Table 4.3: Post-Reform Industrial Employment in the Non-State Sector (2004)

	non-state employ			
Panel A	(1)	(2)	(3)	(4)
non-mining emp	1.441***	1.307***	1.315***	1.426***
	(0.188)	(0.277)	(0.217)	(0.384)
non-mining emp squared				0.003
				(0.053)
initial conditions	X			X
initial conditions polynomial		X		
weights			X	
mean dep var	1.568	1.568	1.568	1.568
	from firms established after 1985			
Panel B	(1)	(2)	(3)	(4)
non-mining emp	0.936***	0.728***	0.881***	0.671***
	(0.141)	(0.169)	(0.169)	(0.238)
non-mining emp squared				0.058*
				(0.029)
initial conditions	X			X
initial conditions polynomial		X		
weights			X	
mean dep var	0.916	0.916	0.916	0.916
N	63	63	63	63

Note: The sample includes prefectures in the Third Front Region with an urban population rate lower than 10 percent in 1964. The dependent variable in Panel A is the non-mining industrial employment in the non-state sector in 2004 per 1982 population ($\times 100$); the dependent variable in Panel B is the part of employment from firms that started operating after 1985. The key explanatory variable is non-mining employment from the key industrial plants in 1985 in the prefecture divided by its 1982 population ($\times 100$). Column 1 and Column 4 include prefecture initial conditions. Column 2 includes the linear and quadratic forms of these initial conditions. Column 3 uses a weighted least squares estimation with the weights constructed such that the initial conditions balances a group of prefectures that had the largest non-mining industrial employment per capita, and the remaining prefectures in the sample. See the text for details. Initial conditions include: (1) urban rate in 1964, (2) log population density in 1964, (3) industrial population ratio in 1936 divided by 1964 population, (4) log distance to existing and planned railways in 1962, (5) log distance the nearest provincial capital, (6) log mean elevation, (7) log average slope, (8) per capita mining employment. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.4: Prefecture-Level Outcomes in 2000

	urban	ind emp	ser emp	ln pop den	mid sch	college
	%	%	%	# per sqkm	%	%
	(1)	(2)	(3)	(4)	(5)	(6)
non-mining emp	5.360***	2.303***	2.002***	0.057**	5.169***	0.606***
	(0.845)	(0.484)	(0.300)	(0.023)	(1.172)	(0.101)
initial conditions	X	X	X	X	X	X
mean dep var	19.824	6.134	11.466	5.117	33.650	1.557
N	63	63	63	63	63	63

Note: Outcome variables are indicated in the shorthand on top of each column. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.5: Initial Conditions versus Continuing Spillovers

	(1)	(2)	(3)	(4)
dep var: non-state ind emp from	all firms	new firms	all firms	new firms
non-mining emp	1.441***	0.936***	1.506***	0.973***
	(0.188)	(0.141)	(0.196)	(0.135)
non-mining emp × high SOE sector growth			-0.217	0.088
			(0.594)	(0.356)
high SOE sector growth = 1			0.468	0.279*
			(0.315)	(0.161)
initial conditions	X	X	X	X
N	63	63	63	63

Note: Unit of observation is a prefecture. The dependent variables are non-state sector industrial employment from all firms (Columns 1 and 3) or from firms that were established after 1985 (Columns 2 and 4), divided by 1982 population ($\times 100$), measured in 2004 Economic Census. State-Owned Sector (SOE) growth is defined as changes in levels of SOE employment between 1985 and 2004, divided by 1982 population (Columns 3 and 4). “High SOE sector growth” is a dummy which is equal to one if the SOE sector growth is above median. Robust standard errors are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.6: Mechanism - Conditional on Early 1980s Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(8)
non-mining emp	0.936*** (0.141)	0.952*** (0.157)	0.929*** (0.135)	0.784*** (0.173)	0.980*** (0.149)	0.941*** (0.282)	
sophisticated manu emp							0.714*** (0.233)
material manu emp							1.136*** (0.134)
light manu emp							0.962* (0.524)
conditions in early 1980s	none	prov	transport	human capital	pop den	all	none
initial conditions	X	X	X	X	X	X	X
N	63	63	63	63	63	63	63

Note: Outcome variable is the non-mining non-state industrial employment in 2004 from firms established after 1985, divided by 1982 population ($\times 100$). All columns control for prefecture initial conditions. Each column controls for additional prefecture conditions in the 1980s. Column 1 includes no additional covariates, thus it replicates the baseline result. Column 2 includes dummy variables for provinces. Column 3 adds transportation conditions in early 1980s, which is measured by log distance to railroad in 1980. Column 4 controls for human capital characteristics, which includes in percent of adult population with middle school diploma, and percent of adult population with college diploma in 1982. Column 5 includes log population density in 1982. Column 6 includes all the covariates from Column 2 to Column 5. Column 7 uses the same set of baseline controls as in Column 1, but breaks the 1985 industrial employment into sophisticated manufacturing (including machinery, transportation equipment, electric devices, electronic products, and instruments), raw material manufacturing (including ferrous and non-ferrous metal products, oil refinery, electricity, etc), and light manufacturing (including textile, paper, etc). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.7: Within-Sector and Input-Output Spillovers

	(1)	(2)	(3)
	emp	emp	emp
non-mining emp	0.936*** (0.141)	0.142* (0.076)	
input linkage			4.267* (2.361)
output linkage			0.138 (2.261)
industry dummies		X	X
'64 conditions	X	X	X
unit of obs	pref	pref-ind	pref-ind
standard error	robust	cluster	cluster
N	63	945	945

Note: The dependent variable is non-state non-mining industrial employment in 2004 from firms established after 1985 in the prefecture (Column 1), and in each industry (Column 2 and Column 3). 3-digit manufacturing firms except for mining industries are included in the sample. Prefecture-industry pairs with zero output in 2004 are dropped from prefecture-industry level regressions. Robust standard errors are reported for Column 1. For Columns 2 and 3, the standard errors are two-way clustered at the industry and prefecture level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.8: Migration and Urbanization

	(1)	(2)	(3)	(4)	(5)
	urban pop grth	urban pop due to migration			
	82-00	other county	95-00	other province	
		before 00		before 00	95-00
non-mining employment	4.273*** (1.288)	0.962*** (0.335)	0.179** (0.078)	0.529*** (0.157)	0.103*** (0.024)
initial conditions	X	X	X	X	X
mean dep var	15.871	3.490	0.939	1.000	0.206
N	63	63	63	63	63

Note: Unit of observation is a prefecture. The dependent variable in Column 1 is the increase in urban population between 1982 and 2000, divided by 1982 population ($\times 100$). The dependent variable in Column 2 is the number of urban residents in 2000 that were born in a county that is different from the current county of residence, divided by 1982 population ($\times 100$). The dependent variable in Column 3 is the number of urban residents in 2000 that migrated from a county that is different from the current county of residence between 1995 and 2000, divided by 1982 population ($\times 100$). The dependent variable in Column 4 is the number of urban residents in 2000 that were born in a province that is different from the current province of residence, divided by 1982 population ($\times 100$). The dependent variable in Column 5 is the number of urban residents in 2000 that migrated from a province that is different from the current province of residence between 1995 and 2000, divided by 1982 population ($\times 100$). Sample mean of the dependent variables are reported. Robust standard errors are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.9: Summary Statistics

	mean	s.d.	p10	median	p90
Panel A: Emp from 1985 key industrial plants per 100 residents					
non-mining sector	.567	1.015	0	.222	1.282
sophisticated-manufacturing sector	.223	.388	0	.051	.581
material-manufacturing sector	.196	.879	0	0	.245
light-industries sector	.149	.206	0	.082	.355
Panel B: Initial conditions					
log population density 1964 (count per sqkm)	4.487	.849	3.418	4.502	5.687
urban rate 1964 (%)	6.303	1.762	3.955	6.361	8.693
log distance to existing and planned railway 1962 (km)	3.808	1.26	2.351	4.023	5.299
log distance to nearest provincial capital (km)	5.345	.466	4.636	5.378	5.894
log average elevation (meters)	7.007	.556	6.284	7.05	7.704
average slope (degrees)	4.308	2.381	1.713	4.019	7.009
industrial employment in 1936 per 10k 1964 resident	.182	1.224	0	0	0
Panel C: Post-reform outcomes (2000s)					
urban rate (%)	19.824	7.437	10.632	18.836	30.152
non-state non-mining industry emp per 100 residents	2.104	1.86	.791	1.535	4.023
from firms established after 1985	1.944	1.764	.759	1.362	3.635
service emp per 100 1982 resident	11.466	3.202	7.778	10.956	15.659
log population density (# per sqkm)	5.117	.771	4.099	5.116	6.182
percent middle school and above (%)	33.65	8.823	21.657	33.623	43.889
percent college (%)	1.557	.621	.956	1.461	2.362

Note: There are 63 prefectures in the sample. These are the prefectures in the Third Front Region with 1964 urban rate less than 10 percent. Key industrial plants are the large and medium sized plants that appear in the 1985 Industrial Census. Industries here do not include the mining sector.

Table 4.10: Shiyan and Its Comparisons (More Results)

	(1)	(2)	(3)	(1)-(3)	(2)-(3)
	Shiyan	Xiangxi	Synth	<i>p</i> -value	<i>p</i> -value
Panel A: Employment from 1985 key industrial plants per 100 resident by industry					
sophisticated manufacturing industries	2.219	0.000	0.016	0.000	0.463
material manufacturing industries	0.195	0.000	0.004	0.000	0.488
light industries	0.172	0.057	0.113	0.195	0.610
Panel B: Post-reform Outcomes (2000s)					
non-state non-mining industry emp per 100 1982 residents					
from sophisticated manufacturing industries	1.534	0.020	0.031	0.000	0.439
from material manufacturing industries	0.285	0.624	0.283	0.341	0.098
from light industries	1.232	0.646	0.657	0.098	0.439
urban rate (%)	32.780	19.776	22.334	0.024	0.659
service emp per 100 1982 resident	19.927	17.184	15.016	0.049	0.122
log population density (# per sqkm)	5.016	5.149	5.034	0.561	0.268
percent middle school and above (%)	44.002	36.337	36.394	0.098	0.341
percent college (%)	3.597	1.914	1.681	0.000	0.268

Note: Conditions for Shiyan are reported in Column 1. Xiangxi is the natural comparison for Shiyan, the prefecture that received the Second Auto Works during the Third Front Movement. Its conditions are reported in Column 2. Conditions for synthetic Shiyan are reported in Column 3. Column 4 reports the *p*-values of the differences between Shiyan and synthetic Shiyan. Column 5 reports the *p*-values of the differences between Xiangxi and the synthetic Shiyan. The *p*-values are obtained using the permutation-like method.

Table 4.11: Prefecture-Level Outcomes in 2000 - Robustness

	urban %	ind emp %	ser emp %	ln pop den # per sqkm	mid sch %	college %
Panel A	(1)	(2)	(3)	(7)	(4)	(6)
non-mining emp	5.990*** (1.122)	2.907*** (0.469)	2.033*** (0.427)	0.019 (0.027)	7.034*** (2.031)	0.799*** (0.096)
initial conditions	X	X	X	X	X	X
initial conditions sqr	X	X	X	X	X	X
Panel B	(1)	(2)	(3)	(7)	(4)	(6)
non-mining emp	5.223*** (0.943)	2.038*** (0.550)	1.692*** (0.626)	0.134 (0.136)	5.268*** (1.286)	0.572*** (0.117)
reweighing	X	X	X	X	X	X
mean dep var	19.824	6.134	11.466	5.117	33.650	1.557
N	63	63	63	63	63	63

Note: Outcome variables are as indicated in shorthand on top of each column. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.12: Data Sources

Category	Variable	Source
1.Industry		
	Third Front Firms by industry	Directory of large and medium sized plants from 1985 manufacturing census
	County industrial output	1985 manufacturing census
	County industrial agricultural income	1982 population census
	County industrial output by ownership and industry	manufacturing part of 2004 economic census
	County employment by industry	1990, 2000, 2010 population censuses
	County average productivity for manufacturing firms	manufacturing part of 2004 economic census
	County average productivity for scale-and-up manufacturing firms	scale-and-up manufacturing firms survey
	County industrial employment and output	1936 manufacturing survey
2.Geography		
	County boundaries	population censuses
	County average alevation and ruggedness	China DEM file
	County area	GIS maps accompanying population censuses
3.Other social and economic characteristics		
	County population	population censuses
	Urban population	population censuses
	Education level of county residents	population censuses
	Per capita output	1982 population census

References

- Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490).
- Abadie, A. and Imbens, G. W. (2006). Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica*, 74(1):235–267.
- Abadie, A. and Imbens, G. W. (2008). On the Failure of the Bootstrap for Matching Estimators. *Econometrica*, 76(6):1537–1557.
- Alberto Abadie, A. D. and Hainmueller, J. (2009). Synthetic Control Methods for Comparative Case Study: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490):493–505.
- Alder, S., Shao, L., and Zilibotti, F. (2013). Economic Reforms and Industrial Policy in a Panel of Chinese Cities. *Available at SSRN 2369416*.
- Alesina, A. and Dollar, D. (2000). Who Gives Foreign Aid to Whom and Why? *Journal of Economic Growth*, 5:33–64.
- Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58(2):277–297.
- aus dem Moore, J. P. and Spitz-Oener, A. (2012). Bye Bye, G.I. - The Impact of the U.S. Military Drawdown on Local German Labor Markets.
- Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American*

- Economic Review*, 103(6):2121–2168.
- Bachman, D. (2013). China's Defence Industrial Base in 1985. *The China Quarterly*, 214:432–455.
- Banerjee, A., Duflo, E., and Qian, N. (2012). On the Road: Access to Transportation Infrastructure and Economic Growth in China. Technical report, National Bureau of Economic Research.
- Barro, R. J. and Redlick, C. J. (2011). Macroeconomic Effects from Government Purchases and Taxes. *Quarterly Journal of Economics*, 126:51–102.
- Bartik, T. J. (1991). Who Benefits from State and Local Economic Development Policies? *Upjohn Press*.
- Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2012). Roads, Railroads and Decentralization of Chinese Cities.
- Bazzi, S. and Clemens, M. (2013). Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth. *American Economic Journal: Macroeconomics*, 5(2):152–186.
- Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics*, 119:249–275.
- Black, D., McKinnish, T., and Sanders, S. (2005). The Economic Impact Of The Coal Boom And Bust. *The Economic Journal*, 115(503):449–476.
- Blanchard, O. J. and Katz, L. F. (1992). Regional Evolutions. *Brookings Papers on Economic Activity*, 1992(1):1–75.
- Bleakley, H. and Lin, J. (2012). Portage and Path Dependence. *The Quarterly*

- Journal of Economics*, 127(2):587–644.
- Boone, P. (1996). Politics and the Effectiveness of Foreign Aid. *European Economic Review*, 40(2):289–329.
- Bound, J. and Holzer, H. J. (2000). Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s. *Journal of Labor Economics*, 18(1):20–54.
- Bun, M. J. and Windmeijer, F. (2010). The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models. *Econometrics Journal*, 13(1):95–126.
- Burnside, C. and Dollar, D. (2000). Aid, Policies, and Growth. *American Economic Review*, 90(4):847–868.
- Busso, M., Gregory, J., and Kline, P. (2013). Assessing the Incidence and Efficiency of a Prominent Place Based Policy. *American Economic Review*, 103(2):897–947.
- Cameron, A. C., Gelbach, J., and Miller, D. (2008). Bootstrap-based Improvement for Inference with Clustered Errors. *Review of Economics and Statistics*, 90(3):414–427.
- Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust Inference with Multiway Clustering. *Journal of Business and Economic Statistics*, 29(2):238–249.
- Channing Arndt, S. J. and Tarp, F. (2010). Aid, Growth and Development: Have We Come Full Circle? *Journal of Globalization and Development*, 1(2):1948–1837.
- Chodorow-Reich, G., Feiveson, L., Liscow, Z., and Woolston, W. G. (2012). Does State Fiscal Relief during Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act. *American Economic Journal: Eco-*

- conomic Policy*, 4(3):118–145.
- Chong, A. and Gradstein, M. (2008). What Determines Foreign Aid: The Donors' Perspective. *Journal of Development Economics*, 87:1–13.
- Clemens, M., Steven Radelet, R. B., and Bazzi, S. (2012). Counting Chickens When They Hatch: The Short-Term Effect of Aid on Growth. *Economic Journal*, 122(561):590–617.
- Daron Acemoglu, S. J. and Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 91(5):1369–1401.
- Davis, D. R. and Weinstein, D. E. (2002). Bones, Bombs, and Break Points: The Geography of Economic Activity. *The American Economic Review*, 92(5):1269–1289.
- Davis, D. R. and Weinstein, D. E. (2008). A Search for Multiple Equilibria in Urban Industrial Structure. *Journal of Regional Science*, 48(1):29–65.
- Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica*, 78(6):1863–1903.
- Diamond, R. (2013). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000.
- Donaldson, D. (2012). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review*, Forthcoming.
- Eric Werker, F. Z. A. and Cohen, C. (2009). How is Foreign Aid Spent? Evidence from a Natural Experiment. *American Economic Journal: Microeconomics*, 1(2):225–244.

- Faber, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *Review of Economic Studies*, forthcoming:1–25.
- Feyrer, J., Sacerdote, B., and Stern, A. D. (2007). Did the Rust Belt Become Shiny? A Study of Cities and Countries that Lost Steel and Auto Jobs in the 1980s. *Brookings-Wharton Papers on Urban Affairs*, pages 41–102.
- Frot, E. and Santiso, J. (2011). Herding in Aid Allocation. *Kyklos*, 64(1):54–74.
- Glaeser, E. L. and Gyourko, J. (2005). Urban Decline and Durable Housing. *Journal of Political Economy*, 113(2):345–375.
- Gottlieb, J. D. and Glaeser, E. L. (2008). The Economics of Place-Making Policies. *Brookings Papers on Economic Activity*, 1:155–239.
- Greenstone, M. and Gallagher, J. (2008). Does Hazardous Waste Matter? Evidence from the Housing Market and The Superfund Program. *Quarterly Journal of Economics*, 123(3):951–1003.
- Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*, 118(3).
- Griliches, Z. and Hausman, J. (1986). Errors in Variables in Panel Data. *Journal of Econometrics*, 31(1):93–118.
- Guthrie, S. J. (1995). *Government Policy and Economic Behavior: A Case Study of National Defense*. PhD thesis, Harvard University.
- Hansen, H. and Tarp, F. (2001). Aid and Growth Regressions. *Journal of Develop-*

- ment Economics*, 64(2):547–570.
- Hausmann, R. and Rodrik, D. (2003). Economic Development as Self-discovery. *Journal of Development Economics*, 72(2):603–633.
- Headey, D. (2008). Geopolitics and the Effect of Foreign Aid on Economic Growth: 1970-2001. *Journal of International Development*, 20:161–180.
- Hirschman, A. O. (1958). *The Strategy of Economic Development*. Yale University Press.
- Hoeting, J. A., Madigan, D., and E., A. (1999). Bayesian Model Averaging: a Tutorial. *Statistical Science*, pages 382–401.
- Hooker, M. A. and Knetter, M. M. (2007). Measuring the Economic Effects of Military Base Closures. *Economic Inquiry*, 39(4):583–598.
- Hultquist, A. and Petras, T. L. (2012). An Examination of the Local Economic Impacts of Military Base Closures. *Economic Development Quarterly*, 26(2):151–161.
- Jerven, M. (2013). Comparability of GDP Estimates in Sub-Saharan Africa: The Effect of Revisions in Sources and Methods since Structural Adjustment. *Review of Income and Wealth*, 59:S16–S36.
- Kline, P. (2011). Oaxaca-Blinder as a Reweighting Estimator. *The American Economic Review Papers and Proceedings*, 101(3):532–537.
- Kline, P. and Moretti, E. (2013). Place based policies with unemployment. *American Economic Review, Papers and Proceedings*, 103(3):238–243.
- Kline, P. and Moretti, E. (2014a). Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley

- Authority. *Quarterly Journal of Economics*, 129(1):275–331.
- Kline, P. and Moretti, E. (2014b). People, Places and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs. *Annual Review of Economics*.
- Kovak, B. K. (2013). Regional Effects of Trade Reform: What is the Correct Measure of Liberalization? *American Economic Review*, 103(5):1960–1976.
- Krueger, A. B. (1999). Measuring Labor’s Share. *American Economic Review*, 89(2):45–51.
- Lewis, A. W. (1954). Economic Development with Unlimited Supplies of Labor. *The Manchester School*, 22(2):139–191.
- Marshall, A. (1890). *Principles of Economics*. London: Macmillan and Co., Ltd.
- McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics*, 142(2):698–714.
- Mekasha, T. J. and Tarp, F. (2013). Aid and Growth: What Meta-Analysis Reveals. *Journal of Development Studies*, 49(4):564–583.
- Miguel, E. and Roland, G. (2011). The Long-Run Impact of Bombing Vietnam. *Journal of Development Economics*, 96(1):1–15.
- Molloy, R., Smith, C. L., and Wozniak, A. (2011). Internal Migration in the United States. *Journal of Economic Perspectives*, 25(3):173–196.
- Moretti, E. (2010a). Local labor markets. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4B. North-holland.
- Moretti, E. (2010b). Local Multipliers. *American Economic Review, Papers and Proceedings*, 100(2):373–377.

- Moretti, E. (2013). Real Wage Inequality. *American Economic Journal: Applied Economics*, 5(1):65–103.
- Moss, T. and Majerowicz, S. (2012). No Longer Poor: Ghana’s New Income Status and Implications of Graduation from IDA.
- Murphy, K. M., Shleifer, A., and Vishny, R. W. (1989). Industrialization and the Big Push. *The Journal of Political Economy*, 97(5):1003–1026.
- Nakamura, E. and Steinsson, J. (2014). Fiscal Stimulus in a Monetary Union: Evidence from US Regions. *American Economic Review*, 104(3):753–792.
- Naughton, B. (1988). The Third Front: Defense Industrialization in the Chinese Interior. *The China Quarterly*, 115:351–86.
- Notowidigdo, M. J. (2013). *The Incidence of Local Labor Demand Shocks*. University of Chicago.
- Nunn, N. and Puga, D. (2012). Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics*, 94(1):20–36.
- Nunn, N. and Qian, N. (2014). U.S. Food Aid and Civil Conflict. *American Economic Review*, 104(6):1630–1666.
- Rajan, R. and Subramanian, A. (2008). Aid and Growth: What Does the Cross-Country Evidence Really Show? *Review of Economics and Statistics*, 90(4):643–665.
- Ramey, V. A. and Shapiro, M. D. (1998). Costly Capital Reallocation and the Effects of Government Spending. *Carnegie-Rochester Conference Series on Public Policy*, 48.
- Rauch, F. and Michaels, G. (2013). Resetting the Urban Network: 117-2012. Tech-

- nical report.
- Redding, S. J., Sturm, D. M., and Wolf, N. (2011). History and Industry Location: Evidence from German Airports. *Review of Economics and Statistics*, 93(3):814–831.
- Roback, J. (1982). Wages, Rents and the Quality of Life. *Journal of Political Economy*, 90(winter):1257–1278.
- Roback, J. (1988). Wages, Rents and Amenities: Differences among Workers and Regions. *Economic Inquiry*, 26(1):23–41.
- Rodrik, D. (1996). Understanding Economic Policy Reform. *Journal of Economic Literature*, 34:9–41.
- Roodman, D. (2007). The Anarchy of Numbers: Aid, Development, and Cross-Country Empirics. *World Bank Economic Review*, 21(2):255–277.
- Roodman, D. (2009a). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71(1):135–158.
- Roodman, D. (2009b). How to Do xtabond2: An Introduction to Difference and System GMM in Stata. *Stata Journal*, 9(1):86–136.
- Rosen, S. (1979). Wage-Based Indexes of Urban Quality of Life. In Miezowski, P. N. and Straszheim, M. R., editors, *Current Issues in Urban Economics*, pages 74–104. Johns Hopkins University Press.
- Rosenstein-Rodan, P. N. (1943). Problems of Industrialization of Eastern and South-eastern Europe. *The Economic Journal*, 53(210/211):202–211.
- Rostow, W. W. (1990). *The Stages of Economic Growth: A Non-Communist Manifesto*. Cambridge University Press.

- Samuelson, P. (1948). *Economics*. McGraw-Hill.
- Schultz, T. W. (1953). *The Economic Organization of Agriculture*. McGraw-Hill
New York.
- Serrato, J. C. S. and Wingender, P. (2014). Estimating Local Fiscal Multipliers.
- Severnini, E. R. (2012). The Power of Hydroelectric Dams: Agglomeration
Spillovers. *Working Paper*.
- Shoag, D. (2012). The Impact of Government Spending Shocks: Evidence on the
Multiplier from State Pension Plan Returns.
- Smith, A. (1776). *The Wealth of Nations*. London: Methuen and Co., Ltd.
- Story, L. (2012). As Companies Seek Tax Deals, Governments Pay High Price. *New
York Times*.
- Story, L., Fehr, T., and Watkins, D. (2012). United States of Subsidies
Searchable. [http://www.nytimes.com/interactive/2012/12/01/us/government-
incentives.html](http://www.nytimes.com/interactive/2012/12/01/us/government-incentives.html).
- Suarez-Serrato, J. C. and Zidar, O. (2014). Who Benefits from State Corporate Tax
Cuts? A local Labor Market Approach with Heterogeneous Firms.
- Topel, R. H. (1986). Local Labor Markets. *Journal of Political Economy*, 94(3):111–
143.
- Wang, J. (2012). The Economic Impact of Special Economic Zones: Evidence from
Chinese Municipalities. *Journal of Development Economics*, 101(2):133–147.
- William Easterly, R. L. and Roodman, D. (2004). Aid, Policies, and Growth: Com-
ment. *American Economic Review*, 94(3):774–780.
- Wilson, D. J. (2012). Fiscal Spending Jobs Multipliers: Evidence from the 2009

- American Recovery and Reinvestment Act. *American Economic Journal: Economic Policy*, 4(3):251–282.
- World Bank, . (2010). A Review of IDA’s Long Term Financial Capacity and Financial Instruments. Technical report, World Bank, Washington, DC.
- World Bank, . (2012). Review of IDA’s Graduation Policy” (IDA16 Mid-Term Review). Technical report, World Bank.
- Yao, Y. (2014). *The Chinese Growth Miracle*, volume 2, pages 943–1031. Elsevier.
- Yoon, C. (2014). Estimating a Dynamic Spatial Equilibrium Model to Evaluate the Welfare Implications of Regional Adjustment Processes: The Decline of the Rust Belt.
- Ziliak, J. P. (2012). The Appalachian Regional Development Act and Economic Change. *Appalachian Legacy: Economic Opportunity after the War on Poverty*.