ESSAYS ON LABOR AND PUBLIC ECONOMICS

Lichen Wang

A Dissertation submitted to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Economics

University of Virginia
August, 2023

Essays on Labor and Public Economics

Lichen Wang

ABSTRACT

Governments are increasingly responsible for resource allocation and policy implementation to improve the affordability of essential living expenses. However, the presence of market frictions, such as limited outside opportunities, imperfect mobility, and information asymmetry, can sometimes lead to unintended consequences of these policies. As a result, additional discretion may be needed to restore the desired outcomes.

The first chapter of the dissertation examines the impacts of government procurement drawdowns on the local labor markets. Specifically, it focuses on the period following the landmark Nuclear Treaty between the U.S. and the Soviet Union. Due to changing diplomatic circumstances, the U.S. Department of Defense (DoD) decreased its procurement budget by approximately 44% over a decade, from \$203 billion in 1987 to \$114 billion in 1997. The exposure to procurement drawdowns exhibits substantial variation across different locations and sectors within a given area. To analyze the local impacts of these drawdowns, I employ an instrumental variable (IV) approach to isolate the plausibly exogenous part of the procurement shocks across locations. The findings reveal that defense drawdowns significantly affected the flagship defense-related sectors. Evidence also suggests that the other manufacturing sectors were unable to entirely absorb the displaced workforce. Additionally, estimates show a net adverse spillover effect on the local non-tradable industry. However, results find

limited impacts on civilian wages and the number of establishments. Further explorations into labor supply responses reveal a swift reaction characterized by reduced in-migration, although with no evidence of increased out-migration. These findings lean towards advocating for targeted approaches, such as specialized job training, rather than pursuing a broad place-based policy.

The second chapter investigates how a national experiment aimed at alleviating individuals' financial burdens could end with unintended consequences. Many countries struggle with high medical expenses due to high prescription medicine prices. This chapter, in collaboration with Jiafeng Wu (University of Virginia), leverages a Chinese national experiment that implemented a zero-markup policy (ZMP) on medication sales in public hospitals to provide new evidence regarding its effectiveness in reducing patients' financial burdens. By employing a conditional difference-in-differences strategy, I examine average spending and service usage changes resulting from this plausibly exogenous government initiative, aiming to understand how physicians modify their behaviors in response. The findings reveal that the average patient's medication expenditure significantly decreased following the policy implementation. However, this reduction was accompanied by increased service spending, resulting in a marginal decline in the overall medical bill. Furthermore, hospitals may have employed various methods to compensate for the loss of drug revenue. Physicians may have leveraged their informational advantage to prescribe costlier treatment materials, which could have been obscured within the final medical bills. The results also indicate suggestive transfers of financial burdens between diagnosis groups.

The third chapter studies the local economic impacts of a positive demand shock from a particular immigrant group. Over the past two decades, the number of international students attending colleges and universities in the United States has nearly

doubled, increasing from 0.54 million in 2000 to a peak of 1.09 million in 2017. In this chapter, I examine the expenditure impacts resulting from the significant influx of international students pursuing higher education in the United States between 2000 and 2019. Focusing on college towns — the cities with a significant presence of college students — the analysis exploits the variation in international student enrollment. To estimate the causal effects on a selected non-traded industry, I employ an IV approach to account for potential confounding factors. The findings support and quantify the existence of a demand effect resulting from the presence of international students. For instance, an increase of 1,000 international college students in a city with a 100,000 baseline population brings about 9 - 13 more food and drinking places. The positive effects, however, are concentrated only in small-sized establishments, leading to relatively limited employment growth. This paper adds to the impact analysis of international students through the lens of local product markets.

Acknowledgments

I am immensely grateful to my advisors, Sarah Turner, Leora Friedberg, and Lee Lockwood, for their unwavering guidance and support throughout my Ph.D. journey. Their teaching and expertise ignited my passion for pursuing my field. During a challenging period of transitioning from a learner to a researcher, Sarah played a crucial role in helping me refine my research questions and expand the scope of my research. Leora was the first professor I interacted with at the outset of my research journey. Since then, she has been a constant source of support, providing countless help on my research and job-landing process. As my advisor and supervisor during my research assistantship, Lee has been a role model from whom I have greatly benefited, learning how to improve my research habits and foster both my academic and professional growth. In addition to offering insightful advice on my research, they have consistently demonstrated thoughtfulness and patience in helping me overcome challenges during the past years.

I want to extend my gratitude to William Shobe, John McLaren, Amalia Miller, Daniel Murphy, Simon Anderson, and all other faculty and staff at the Department of Economics who have shared helpful comments and insights. I am also fortunate to have the encouragement and support of John Rogers, Sam Ouliaris, Qiang Xu, and Behzad Diba, without whom I would have never begun my doctoral studies.

Special thanks to the colleagues and friends (The name list would be too long to be added here.) I met and discussed with in Charlottesville, D.C., Austin, Seattle, and virtually. They made the last six years a joyful and unforgettable memory.

I have been blessed to have the unconditional love and support of my parents. This dissertation would not be possible without the companionship of my husband, Xing Hong, who has always been my biggest cheerleader, offering a listening ear and words of wisdom. Their presence in my life has made all the difference. I am forever grateful for their love and belief in me.

Contents

Chapter 1

The Local Economic Effects of the Post-Cold War Procurement Drawdown

1.1 Introduction

The United States has witnessed a substantial decline in economic convergence across places in recent decades (Moretti, 2011; Ganong and Shoag, 2017). Some temporary economic fluctuations can lead to persistent impacts. While the severity of job losses often attracts attention in the media, whether and how a place returns to its previous trajectory extends beyond simple job numbers. Due to the diverse causes of economic downturns and the unique locations and populations affected, the existing literature offers mixed conclusions on how the shocks can shape the trajectories of economies. Consequently, it becomes essential to understand the nature of economic shocks and the adjustment process of the affected parties for effective policy implications in restoring economic activities.

In this paper, I investigate a downsizing government procurement, with a specific focus on defense cuts, and study their impacts on labor market disparities across geographical areas. Over the years, defense spending has played a crucial role in injecting substantial amounts of money into the economy, primarily through private business contracting, and has prompted scholars to explore its potential for job creation (e.g., Barro and Redlick, 2011; Nakamura and Steinsson, 2014; Auerbach, Gorodnichenko, and Murphy, 2019). However, the future of defense spending remains uncertain, as it is influenced by various factors, including changes in geopolitical circumstances, economic conditions, and the government's budgetary priorities. This uncertainty has led to a noticeable divide among the public regarding the direction of its future path. Yet, relatively little attention has been devoted to studying the impacts on local economies, given the potential for future defense drawdowns.

To address this gap, I exploit a unique historical period that followed the 1987 Nuclear Treaty between the U.S. and the Soviet Union. This landmark treaty marked the end of the Cold War arms race and resulted in substantial cuts to the U.S. Department of Defense (DoD) procurement budget. The aftermath of the treaty led to a remarkable 44% reduction in U.S. defense procurement. With the DoD's budget declining from $203 billion in 1987 to $114 billion in 1997, it represented the longest and most significant reduction in domestic defense procurement in recent history.[1] By studying this specific historical period, the research aims to gain insights into the challenges faced by local economies and the necessity for government engagement in local demand promotion.

[1] The DoD shrank its military spending recently in 2008, but neither the duration nor the drawdown's magnitude is comparable to the post-Cold War era. Domestic defense spending decreased by 25% from 2008 to 2015 and has been upward since then (Government Accountability Office, 2018).

Several aspects underline the unique features of the defense procurement drawdown and highlight the importance of studying the impacts empirically. It offers certain advantages in theoretically predicting the hiring decisions of the affected firms. Previous studies, such as Cohen, Coval, and Malloy (2011) and Hebous (2016), have discussed how government spending may crowd out labor supply in the private sector, with the extent depending on firms' ability to quickly adjust to the loss of government demand by finding other private sources of revenue. However, in the case of defense procurement drawdowns, the majority of the affected industries are involved in producing missiles, airplanes, and warships. These contractors typically have limited outside private alternatives, which hampers their capacity to compensate for the lost demand. Additionally, the downward wage rigidity (Dickens et al., 2007) suggests that firms may be more inclined to adjust their hiring decisions through extensive margin changes, such as reducing the number of employees, rather than adjusting wages. Consequently, the adverse shocks generated by defense procurement drawdowns allow for much sharper theoretical predictions of within-industry impacts, enabling a better evaluation of the quantitative estimates of the within-sector impacts.

On the other hand, the overall labor adjustments across industries remain unclear. The successful transition of workers from directly affected industries to other industries depends on the transferability of their skills and the demand for those skills in alternative fields. While we cannot directly observe the type of workers being affected, certain patterns indicate that the directly affected industries tend to rely heavily on high-skilled workers. It implies that workers facing job losses due to the drawdowns may encounter challenges in finding full absorption in other tradable industries.

When workers in these affected industries experience a decrease in income due to job losses, there may be limited impacts on demands for tradable products, but their diminished spending power can have ripple effects on other non-tradable sectors that depend heavily on local consumer activity. However, one may expect a dampened spillover resulting from adverse shocks. During a period of negative income shock, households are less likely to reduce their consumption to the same extent as they would increase it during favorable economic conditions.

In light of these implications, this paper provides quantitative estimates of the local labor market impacts of directly affected industries. Additionally, it investigates the extent of spillover effects on tradable and non-tradable sectors, respectively. Furthermore, it delves into the adjustment mechanisms through which the local workforce responds to these labor market shocks. Previous research suggests that the direct impacts on individual workers are minimal if they have the flexibility to move between employers or geographic locations. However, the collective response of the local population can significantly influence the geographical distribution of economic activity, potentially exacerbating economic disparities across locations. Therefore, understanding the mechanisms by which labor markets adapt to shocks is essential for developing effective policies to mitigate the adverse effects on affected communities.

To analyze the effects of procurement cutbacks on local labor markets, I exploit a granular dataset of procurement contracts signed by individual businesses with the DoD. This dataset covers nearly the entire population of firms engaged in contracts with the DoD, providing a comprehensive and representative sample. The detailed information of each contract includes specific procurement products, the county where each project is performed, and the exact timing of each project. Leveraging the comprehensive details is instrumental in identifying the exact industries involved and the precise magnitude of drawdowns at each county.

With this level of granularity, I assess the variations in demand shocks created by drawdowns on specific industries and locations. Given that the majority (more than 80%) of defense spending drawdowns occurred within the manufacturing sector, I specifically focus on the impacts of procurement shocks within key defense-related manufacturing industries. I combine the information from the procurement dataset with various Census datasets. This integration allows me to analyze the direct and spillover effects of defense drawdowns on other local sectors and study how the local labor supply adapts to the changes.

While federal decisions on the defense budget are influenced by diplomatic tensions, leading to plausibly exogenous national variation in defense spending over time, the local allocation of defense drawdown may be determined by transitory local economic shocks, introducing endogeneity issues that challenge the causal interpretation of its effects. To address these challenges and estimate the causal impacts of procurement drawdowns, I employ an instrumental variable (IV) approach. It utilizes the idea that national industry-specific procurement shocks propagate to local economies based on their historical procurement distribution, which is assumed to be uncorrelated with the local contemporaneous shocks with additional controls. I use two measures of instruments that leverage the geospatial and industrial variations across counties. Specifically, the first IV approach uses the past geographical distribution of national contracts to determine the exposure of industries across places. The second IV approach capitalizes on within-county variation in reliance on DoD procurement resulting from previous build-ups. For instance, areas that previously had a higher level of contracting with the DoD in industries such as missile or ship production - the ones with the most significant drawdowns - would be expected to have experienced relatively larger shocks compared to places with minimal reliance on procurement in such sectors. To enhance the causal interpretation of the strategy, I follow Goldsmith-Pinkham, Sorkin, and Swift (2020) and explicitly control for possible correlates through a series of specifications.

I begin by examining the effects of DoD procurement withdrawals on the local changes in total employment counts in the flagship industries that experienced the majority of drawdowns. The findings indicate that for every per-capita decrease of one thousand dollars in a county's procurement within the flagship defense sectors, there is a reduction of 0.3-0.4 percentage points in employment within the same sectors in the county. The national procurement drawdowns in these flagship sectors amounted to $73 billion, which translates to a national employment decline of 248 thousand between 1987 and 1997. It appears to be a small number of jobs lost in response to a substantial cutback in dollar terms, but consistent with the fact that these flagship defense sectors are highly capital-intensive, involving relatively fewer workers in producing expensive goods.[2]

However, the national average change per capita obscures the significant variation in employment impacts across localities due to varying levels of procurement exposure. The standard deviation of the two-year procurement change per capita in the sample, which was $205, indicates the considerable variation in defense procurement drawdowns experienced by different counties. With an average county population of 344.6 thousand, the counties in the 10th percentile alone experienced a decline in sector employment of 198 thousand. The sizable employment response in these most affected sectors confirms our hypothesis that these defense-contracting firms

[2] An alternative way to interpret the findings is through the cost per job estimation, which suggests that, on average, each job lost corresponds to approximately $294,000 ($73 billion/248 thousand workers) in procurement cutbacks.

were unable to fully replace their funding sources with private sources, highlighting their challenges in coping with the defense procurement drawdowns.

The empirical evidence further suggests the challenges workers face in the flagship defense sectors when attempting to transition smoothly to other manufacturing sectors. The estimates for other sectors within the local labor market indicate a net negative spillover effect, as evidenced by a 0.10 percentage-point decrease in the retail industry. Interestingly, despite the negative spillover effect on employment, there is no significant evidence of amenity loss, as the number of local establishments did not change correspondingly, and there is no strong evidence of a local wage effect.

Given the interconnected nature of economies, demand shocks would typically dissipate if the labor supply responded quickly. To explore how the local workforce adjusts to procurement shocks, I analyze the extent to which the local population migrates or exits the local labor force in response to the drawdowns. The estimates reveal a significant labor supply adjustment in response to the demand shocks, characterized primarily by a decline in inflows into the affected areas rather than individuals moving out of those areas. This pattern is consistent with the general conclusion of a reduced labor supply from Blanchard and Katz (1992) and aligns with findings from Dao, Furceri, and Loungani (2017), which demonstrated that inflows primarily drive the countercyclical population response to states with better economic performance during recessions. Furthermore, the larger impact of reduced in-migration suggests that affected areas become less attractive to labor market entrants.

The rest of the paper is organized as follows. Section 1.2 highlights the connection and contribution to literature. Section 1.3 introduces the historical background. Data sources and sample descriptions are in section 1.4. I then discuss the empirical approach in section 1.6. Section 1.7 presents the main results, and 1.8 provides further analysis. Section 1.9 concludes.

1.2 Literature Review

This paper first contributes to the literature that studies the impacts of local economic shocks. Previous studies, such as Feyrer, Mansur, and Sacerdote (2017), have investigated the transmission of income and employment effects of local shocks generated by the fracking boom. Autor, Dorn, and Hanson (2013) explores the consequences of China's import shocks, providing insights into the decline of manufacturing employment and the negative externalities on non-manufacturing jobs. Additionally, the direction and magnitude of spillovers across employers within the same industry have also been discussed extensively, particularly in the retail sector (e.g., Neumark, Zhang, and Ciccarella, 2008; Haltiwanger, Jarmin, and Krizan, 2010; and Shoag and Veuger, 2018). This paper contributes by examining local labor market shocks from a different perspective - a downsizing government procurement. It explores and provides new insights into the extent to which another type of economic shock can shape geographical disparities.

Another body of related literature studies the impacts of government stimulus plans on private-sector employment. This strand of literature has provided mixed evidence regarding the substitutability between public purchases and private activity, as well as the magnitude of the impact (e.g., Conley and Dupor (2013); Feyrer and Sacerdote (2011); Wilson (2012)). In addition to funding through state governments as an intermediate channel, another common role of stimulus plans is the direct provision of contract grants and loans to private employers. Defense procurement, which accounts for 70% of all federal procurement, is a suitable example within the context of this research. The existing literature has mainly focused on regional fiscal multipliers, with relatively little attention devoted to understanding the consequences of local workforce responses to such shocks. To this end, my analysis complements the literature by quantifying the impacts of an adverse shock created by government funding withdrawals. Moreover, I present evidence on the impacts from a micro-perspective, contributing to a more comprehensive understanding beyond the scope of a fiscal multiplier.

This study also contributes to the existing literature by offering new estimates of local labor force adjustments in response to adverse industrial shocks. Despite the extensive discussions on the effects of local labor markets, there is mixed evidence on individual adjustments to different industrial shocks. For example, Black, McKinnish, and Sanders (2005) find more migration responses in the coal bust period at the county level, although they measure the response time over a decade. Foote, Grosz, and Stevens (2019) focus on a massive layoff, discovering that moving out of the counties was the primary reaction of workers. In contrast, Autor, Dorn, and Hanson (2013) find weak labor mobility but an exit of the labor force as a response to import competitions. The differential workforce responses observed in these contexts may stem from variations in affected places and the demographic characteristics of the impacted population (Molloy, Smith, and Wozniak (2011). Contrary to prior research, my investigation of defense-related industries in the 1990s reveals that searching for jobs outside the local market was not the primary exit channel. The employment composition in these sectors differs from that in existing studies on regional shocks, underscoring the importance of empirically assessing the extent of local workforce adjustments and the means through which they occur. In comparison, the defense procurement shock was concentrated in heavy industries, which tend to be more capital-intensive and employ a higher proportion of high-skilled labor, as depicted in Figure 5. Conversely, light industries are more labor-intensive and have lower shares of college-educated workers.

Furthermore, the study highlights the importance of examining the impact of DoD procurement shocks on non-military counties, which have received limited attention in defense policy evaluations. Existing research primarily focused on Base Realignment and Closure (BRAC) and mainly discussed the impacts of military force reductions on the base counties (e.g., Poppert and Herzog Jr, 2003; Hultquist and Petras, 2012; Zou, 2018). However, these reductions are part of significant national defense policy change, and understanding the full impact of DoD spending shocks requires examining the effects beyond military counties. Indeed, approximately 80% of the counties that received DoD procurement did not have a military base. Therefore, this study provides novel insights into the impact of defense spending shocks on non-military counties, contributing to a better understanding of the broader economic consequences of a changing defense environment.

1.3 Background

The United States has a significantly higher defense spending than any other country, exceeding the combined defense spending of the following seven countries (Peter G. Peterson Foundation report, 2019.)[3] In addition to being the top spender in an international context, national defense spending is the second-largest source of government spending, second only to Social Security. In the Department of Defense (DoD) spending budget, procurement is a significant component of expenditure, ac-

[3] https://www.pgpf.org/chart-archive/0053_defense-comparison, accessed on March 1, 2020

counting for over 70% of all federal procurement, according to Federal Procurement Reports.[4] As a result, the DoD has long relied on private contractors to provide a wide range of products, from weapons and transportation to food and uniforms.

The recent history of a push for increased military spending in the United States began in 1997 and gained momentum during the Afghanistan and Iraq wars. However, the country did not always adopt an assertive defense doctrine, particularly when confronting a less formidable threat. With the escalating diplomatic tension between the United States and the Soviet Union in the early 80s, there was significant growth in military spending, particularly in weapons procurement. Nevertheless, the post-World War II Cold War era drew to a close during the second term of then-President Ronald Reagan when Mikhail Gorbachev assumed leadership of the Soviet Union in 1985. In December 1987, Reagan and Gorbachev officially signed the Intermediate-Range Nuclear Forces Treaty, eliminating all (short-range and middle-range) nuclear and conventional missiles. All the long-range missiles were abolished later in 1991. In reaction to the changing diplomatic regime, there has been a sharp fall in defense-related procurement since 1988. The end of the Cold War marked a significant decrease in the demand for missiles and other weapons, and the DoD shifted its focus to more efficiently and effectively supporting military forces.[5]

The sudden change of the national defense policy brought sizable uncertainties to the contractors and locations heavily dependent on DoD projects. The extent of their exposure to procurement drawdown varied depending on their industrial composition and historical relationship with the DoD. Section 1.5 provides a detailed analysis of contract allocation across industries and locations and the corresponding changes

[4] Specifically, the composition of defense spending is mainly categorized into the following sections: Operation and Maintenance (43%); Procurement (32%); Military Personnel (24%); and Military Construction and Family Housing (1%) (Office of Management and Budget, the Fiscal year 2020).

[5] see `https://www.acq.osd.mil/brac/`

over time. To summarize, California was the state that received the highest value of contracts in 1987, accounting for 18% of the national DoD procurement, while Wyoming received the least, with only 0.03%. The transportation-manufacturing industry experienced the most drastic drawdown, with California still ranking first in contracting activities. However, the variation in contracts within a particular sector did not solely reflect the overall changes in DoD's total procurement. Several geographic clusters of procurement in this industry were located in places where the total procurement amount was not among the highest, indicating a geospatial variation in the types of procurement.

1.4 Data

The DoD's contracts provide the basis for measuring county-level government procurement drawdowns. The prime contracts awarded by the DoD are collected from the Department of Defense Form 350 (DD350), part of the Federal Procurement Data System. The sample includes nearly all U.S. project-performing counties where firms have contracted with the DoD at least once. According to the United States General Accounting Office, the DD350 reports all contract actions greater than $25,000 since 1982.[6] I use the GDP price deflator from the Bureau of Economic Analysis with the base year 2000 to convert the contract amount over time.

The data used in this study has the advantage of providing detailed information on contractor identification, contract product, DoD funding obligations, project

[6] If a county does not report any values in any year, it either is because it does not satisfy the reporting threshold or is not assigned with any contracts. I included these observations (and treated them as 0 in my sample) to avoid sample selection bias. Although it could introduce measurement errors, the problem is not notable. Through aggregating contract obligations by the place of performance, I verified the total dollar amounts calculated from DD350 in 2000 with the Federal Procurement Report Year 2000, and the value matches well.

dates, and project work performance location at the county level. This level of granularity is instrumental in measuring economic shocks at the local level, which can sometimes be difficult to approximate in the literature. Commonly used indicators, such as local unemployment rates, may already reflect labor supply responses to demand fluctuations, making it challenging to accurately quantify the demand changes resulting from specific groups of industries.[7] In this study, the geographic distribution of different product types from DoD contracts allows for an identification of the origin of adverse labor demand shocks in defense industries and an accurate quantification of the extent of these shocks.

The procurement database organizes the products of contracts using the 1987 Standard Industrial Classification (SIC) system. For products not present in any other year in the sample, I use the Product Code Manual from Acquisition.gov and companies' SIC information from Mailinglist.com to fill in the missing information on the types and associated industries. I aggregate the contract-level information into the county-industry level and use a two-year average of contract obligations to smooth out any downward adjustment (de-obligations) at the county level.[8] The data pattern and research design in sections 1.5 and 1.6 illustrate the cross-sectional exposure variation to the nationwide procurement drawdown.

The primary measure of county employment is from the imputed County Business Patterns constructed by Eckert et al. (2020). It is based on the U.S. Census County Business Patterns (CBP) data but has two major advantages. Firstly, it imputes missing employment data for most county-industry cells that are suppressed in

[7] It may also be difficult to quantify the changes resulting from a particular group of industries in some contexts. To measure the differential exposures to the trade shock, Autor, Dorn, and Hanson (2013) allocated import values across locations based on indirect employment shares of industries.

[8] Small counties that sometimes reported zero employment in the MF sector are excluded, and the focus is on counties that have received at least one DoD contract.

the Census CBP.[9] This is done through a programming method that fills in missing data for county-industry cells where the number of employees working in an establishment is missing. Secondly, it resolves the inconsistencies of industry classifications over time. The rich information on county-industry employment is beneficial in the context of this paper, which examines the *local* labor market effects from procurement drawdown of the defense industries in the 1990s.[10] The wage and salary employment counts from the Bureau of Economic Analysis (BEA) Regional Economic Accounts provide additional information on county-level earning. These data are based on the U.S. Bureau of Labor Statistics and are adjusted in alignment with other BEA statistics.

To study the population adjustment and potential labor force exit channels, I investigate the data on the county-to-county migration flows and SSA filers. The former data is available from the Internal Revenue Service (IRS). The IRS calculates the summary of the population's inflows and outflows by tracking individual tax filers' address changes. It contains the number of returns filed — an approximation of households and the number of personal exemptions — and proxy measures of individuals moving in and out of a county. Using the aggregate number of individuals who changed county addresses from the last year, the county migration flows from 1980 to 1997 can be built. This information is useful in understanding how the population responds to economic shocks and whether labor force exit channels, such as migration, played a role in mitigating the negative effects of the procurement

[9]In the example world of only one big establishment in a county-industry cell, the number of employees working in this establishment is missing in the Census CBP.

[10]The imputed datasets increase the sample observations by 400 counties, from 1,600 to 2,000 counties that contracted with the DoD at least once before 1985. The correlation between the Census CBP and the imputed CBP is higher than 0.99 and 0.98 between the Census Quarterly Census Employment and Wages (QCEW).

drawdown.[11] The IRS-based migration flow data has a comparative advantage over other survey data in its broader coverage of finer spatial scales (counties) but is restrictive in representing the local population.[12] While I do not directly observe the labor force participation channels, I provide supporting evidence by measuring the number of people receiving disability or retirement benefits in each state using data from the OASDI Beneficiaries by State and County, U.S. Department of Health and Human Services. This data provides additional insights into how individuals respond to economic shocks and whether the drawdown resulted in changes in disability claims or retirement.[13]

I supplement these data sources with a vector of county demographics obtained from the National Vital Statistics System and IPUMS National Historical Geographic Information System (Manson et al., 2017). Specifically, the demographic controls include the county-level age, gender, racial, and educational composition. I use the Surveillance, Epidemiology, and End Results (SEER) from the National Cancer Institute to compile the annual counts of a county's population. To demonstrate that most contract-receiving counties are the ones without a military base, I identify the locations of military counties using the 1989 Base Structure Report, and conduct robustness check using only the non-base counties.

[11] The migration flows in IRS are tracked using the address changes of individual tax filers, so the migrants are those who filed different counties of residence in two consecutive years.

[12] A commonly-noted limitation is its under-representation of the people who do not file tax returns, usually those who are self-employed and have lower income, which could be less prevalent in defense-related industries. However, it is still true that such limitations may lead to an underestimation of the migration responses.

[13] All the archived OASDI files before 1999 were recorded in the annual statistical yearbooks. While optical character recognition (OCR) is usually used for digitizing scanned images, the image quality is too low and introduces measurement errors. I instead referred to the state-level statistics that were recognized and inputted in person.

1.5 Descriptive Analysis

Figure 1.1 plots the overview of the national DoD procurement since 1980, adjusted for inflation (2000 US dollars). After several years of build-up during the early 1980s, the total amount trended downward continuously through the late 1990s. There were substantial differences in the distribution of DoD contracts across sectors (Figure 1.2). For example, among all ten broad sectors, the manufacturing sector was where the majority of the obligations took place and was also the one with the largest drawdown during the sample period between 1987 and 1997. Figure 1.3 further breaks down the manufacturing sector by highlighting the top five sectors that received the most significant amount of DoD procurement. These sectors were responsible for manufacturing "equipment for transportation of passengers and cargo by land, air, and water" (SIC 37); "geophysical equipment; search, detection, navigation, and guidance systems" (SIC 38); "electronic and other electrical equipment and components, except computer equipment" (SIC 36); "fabricated metal products" (SIC 34); and "industrial and commercial machinery and computer equipment, e.g., engines" (SIC 35). Together, they (SIC 34 to 38) received more than 95% of all manufacturing procurement in 1987 and 90% of all manufacturing procurement in 1997. I refer to these five sectors as the five flagship manufacturing defense sectors, MF, in the remaining sections.

Out of the total $131 billion worth of contracts in the broad manufacturing sector in 1987, the transportation equipment industry (SIC 37) accounted for a substantial portion, specifically 61.1%, equivalent to $80 billion. However, over the following decade, there was a significant decline in total procurement in this industry. By 1997, the total contracts awarded dropped to $34 billion, even though the transportation equipment industry still represented over 57% of the total manufacturing contracts.

Within the transportation equipment industry (SIC 37), the Department of Defense (DoD) procurement was concentrated in a few specific industries. Figure 1.4 illustrates that the vast majority (more than 90%) of the contract value in SIC 37 was allocated to industries involved in building guided missiles and space vehicles (SIC 3761), military ships (SIC 3731), and aircraft (SIC 3721 and 3728). These industries experienced the most significant reduction in contract value during this period.

Regarding geographical distribution, the government procured missiles, ships, aircraft, and other transportation equipment primarily in certain locations. Figure 1.6 highlights counties in some states were among the primary beneficiaries of the government's procurement activities. In 1987, California received the largest share, accounting for 18% of the total contract amount. Following California, other significant recipients were Maryland, New York, Texas, and Virginia.

In Appendix, I conduct a case study to demonstrate the relationship between employment and procurement shocks in the manufacturing transportation equipment sector (SIC 37). The analysis focused on examining changes in the employment structure of counties that received higher shares of national procurement in SIC 37 since 1987, the period after the signing of the Nuclear Treaty. Counties with a procurement share of national spending in SIC 37 above 0.08% (representing the top 10% of the distribution of SIC 37 federal procurement share) in 1987 were identified as places that were intensive in the procurement of SIC 37. The event-study analysis, as shown in Appendix Figure 1.9, provides compelling evidence regarding the employment impact. Compared to the reference year of 1988, the employment shares of industry SIC 37 in counties with higher initial procurement shares were negatively affected by the DoD procurement shock during the post-Cold War era.[14]

[14] The impact estimated is larger when I set a higher threshold of the initial procurement share. It indicates a larger employment effect in counties initially receiving the most from DoD procurement

1.6 Empirical Strategy

This section outlines the identification approach used to evaluate the effects of procurement shocks on the local labor market, the county where a project is performed.[15] The primary objective of this analysis is to quantify the extent to which the drawdown in procurement led to a decline in local employment and to examine the responsiveness of the local workforce in mitigating these demand shocks.

I begin by presenting a baseline analysis of how DoD spending shocks affect the local labor market by changing total employment counts in the private sector. Given that the majority of drawdown occurred in the manufacturing sector, where more than 90% was concentrated in 5 broad groups (SIC 34 - 38), as shown in Figure 1.3, I focus on measuring the direct impacts of the procurement change in these industries, which I refer to as "MF." The vast majority of the drawdown took place in these MF sectors, making it less likely that the causal interpretation of the local labor market will be confounded by a general procurement reduction covering a broad range of product types.

Specifically, the baseline specification takes the following form:

$$\Delta Y_{k,c,t} = \beta_k \Delta Proc^{MF}_{c,t} + \Delta\tau_t X'_c + \gamma_t + \epsilon_{ct}. \tag{1.1}$$

The specification is in the stacked difference form, where each observation measures the periodic change. Here, c indicates county, and t indicates years. $Y_{k,c,t}$

in the transportation equipment industries.

[15]The analysis exploits variation at the county level because counties are the smallest geography units identified in a DoD contract. The coefficient estimates might show more conservative results than a broader unit such as MSA and CZ. On the other hand, it could also unmask changes that may not be salient in more expansive areas.

measures the outcome k, including county employment, wage, number of establishments, local population, and people's labor force adjustments. $\Delta Proc_{c,t}$ measures the value change of DoD procurement awarded to projects performed in county c, normalized by a county's population at the beginning of the period. Here, I use the periodic difference between the two-year averaged levels over pairs of endpoints.[16]The coefficient of interest, β_k, measures the impact of a two-year periodic change of DoD obligation on the outcome k during the same period. County fixed effects are excluded because the variables are expressed in the differenced form, and year fixed effects, τ_t, control for aggregate economic shocks common to all places.

Since larger places were more likely to experience more procurement drawdown and, at the same time, saw a more significant number of jobs added per year, failure to control for city size can result in biased estimates of the relationship. To mitigate scale effects, I follow Peri and Sparber (2011) and use the beginning-of-the-period population to normalize the level change of the total procurement and the total number of employment counts.

The OLS model faces a potential threat from the presence of confounding local factors that determine the extent of DoD procurement shocks and simultaneously influence local labor market conditions. The systematic and nationwide changes in DoD procurement, primarily driven by evolving diplomatic tensions during the analyzed period, offer a plausible foundation for the exogenous variation in national demand shifts. However, the politically motivated nature of these decisions may lead to an uneven distribution of drawdowns across different places, potentially correlating with local economic shocks (Demyanyk, Loutskina, and Murphy, 2019). For example, the

[16]For example, for the period endpoint 1997, I use a two-year moving average of a county's procurement in 1996 and 1997, $\overline{Proc_{c,1997}}$. Unless mentioned otherwise, $Proc_{c,t}$ represents $\overline{Proc_{c,t}}$. The periodic contract change between 1995 and 1997, $\Delta Proc_{c,1997}$ is then $\overline{Proc_{c,1997}}$ - $\overline{Proc_{c,1995}}$. Taking two-year average filters out the fluctuations while capturing the procurement trend.

DoD might be less inclined to withdraw funds from localities experiencing economic fluctuations, such as natural disasters or significant layoffs occurring simultaneously.

To address these sources of endogeneity, I apply an instrumental variable (IV) strategy that exploits plausibly exogenous variation in DoD procurement. Specifically, I use two types of shift-share instruments. The first is based on the initial geographical distribution of various industrial procurement across the country and a common national procurement change by product type. The first IV takes the form:

$$\Delta Proc_{ct}^{MF^{IV}} = \sum_{j \in MF} \frac{Proc_{c,1985}^{j}}{\sum_{c'} Proc_{c',1985}^{j}} \times \Delta Proc_{Natl,t}^{j}. \tag{1.2}$$

Here, the share is a location c's initial share of national procurement in the industry j (categorized by two-digit 1987 SIC level). The shift is driven by the federal procurement change in the same sector between years t-2 and t. To aggregate the predicted procurement in each industry j within the MF category, I assume that the later-year contracts will be allocated across counties in the same proportions as their initial distribution before the treaty was signed. Therefore, the *de facto* variation in the instrument across locations is determined by each place's historical share of the national procurement. The critical assumption underlying this instrumental variable is that the geographical distribution in the past is not correlated with later-period shocks affecting the changes in labor market outcomes once with the additional initial controls. The initial year for the instrument is set at 1985, which coincides with the peak of national defense procurement and prior to our sample that starts from 1987.

The second instrumental variable used in this analysis is a Bartik-type IV. It combines the local procurement shares with the national procurement growth rates across industries. To explain the concept, we can express a county's periodic change of procurement as:

$$\Delta Proc_{ct}^{MF} = Proc_{c,t-2}^{MF} \times g_{ct}^{MF}.$$

Multiplying the county's growth rate of procurement by its initial procurement amount yields a periodic change in procurement. Here, the first term $Proc_{c,t-2}^{MF}$ is the total procurement in industry j for location c. The second term g_{ct}^{MF} represents the procurement growth rate in industry MF for location c:

$$g_{ct}^{MF} = \sum_{j \in MF} \frac{Proc_{c,t-2}^{j}}{Proc_{c,t-2}^{MF}} \times \frac{\Delta Proc_{c,t}^{j}}{Proc_{c,t-2}^{j}}.$$

The decomposition of the periodic change $\Delta Proc_{ct}^{MF}$ reveals two potential sources of bias in our estimates. First, the beginning-of-the-period local procurement share, $\frac{Proc_{c,t-2}^{j}}{Proc_{c,t-2}^{MF}}$, may be correlated with contemporaneous local shocks that influence subsequent employment growth. Second, the growth rate of procurement across industries in a county c, $g_{c,j,t}$, may be influenced by contemporaneous county-specific shocks that affect employment outcomes. In fact, we can further decompose $g_{c,j,t}$ into a nationwide industrial trend $G_{j,t}$ and an idiosyncratic location-specific growth shock $g_{\tilde{c,j,t}}$, with the latter potentially correlated with unobserved determinants affecting changes in local labor market outcomes (Goldsmith-Pinkham, Sorkin, and Swift, 2020). By employing the Bartik-type IV, as shown in Equation 1.3, we aim to mitigate these potential sources of endogeneity.

$$\widehat{G_{ct}^{MF}} = \sum_{j \in MF} \frac{Proc_{c,1985}^{j}}{Proc_{c',1985}^{MF}} \times \frac{\Delta Proc_{Natl,t}^{j}}{Proc_{Natl,t-2}^{j}} \tag{1.3}$$

$$\Delta \widehat{Proc_{ct}^{MF}} = Proc_{c,t-2}^{MF} \times \widehat{G_{ct}^{MF}} \tag{1.4}$$

More specifically, the Bartik-type IV approach addresses endogeneity concerns by fixing the historical procurement share at an initial point in time, the year 1985, before the analysis sample begins. By using the growth rate of procurement across industries at the national level, denoted as $G_{j,t}$, the idiosyncratic shock of local procurement growth $g_{\tilde{c},j,t}$ is effectively filtered out.

Indeed, the Bartik-type IV approach used in this study shares a similar conceptual framework with the first shift-share IV, as both aim to capture a common procurement shock resulting from plausibly exogenous decisions made by the federal government. The key distinction lies in how the historical distribution of defense contracts is utilized to determine local exposure to this shock. The first IV relies on the geographic distribution of national procurement to estimate exposure. In contrast, instead of considering the size of a county's procurement relative to other counties, the Bartik-type IV focuses on the specific sector composition within a given county. By doing so, it helps mitigate the concern with the first IV, namely that areas with initially high levels of procurement in the base period may already be in the middle of a positive labor market shock unrelated to the procurement shock, even conditional on local characteristics (Smith, 2012). Further details regarding the condition for the identity of these two measures can be found in the Appendix.

The validity of the IV strategies relies on the exclusion assumption that differential reliance on industry-specific contracts at an initial point should not directly affect changes in labor market outcomes in future periods, despite potentially determining their future levels. However, the identification strategy is vulnerable to threats if the initial procurement structure is associated with other county-level characteristics that predict labor market changes through alternative pathways.

To address these concerns, I incorporate a set of county demographic characteristics X_c' that may be correlated with the historical procurement exposure and simultaneously shape the development trajectory of local amenities, thereby influencing changes in labor market outcomes.[17] The additional controls include birth rate, age composition, gender composition, racial composition, and educational attainment of the local population.[18] Adding time-varying demographic characteristics may lead to over-controlling as these outcomes could be the intermediate responses to the local procurement change and affect the labor market outcomes. Therefore, I follow Goldsmith-Pinkham, Sorkin, and Swift (2020) and control for the initial county-level demographic characteristics in 1985, which is also the time point used to compute the local exposure.

To account for the potential heterogeneity in county trajectories arising from differences in their industrial compositions, I introduce controls for the initial proportion of local employment in each of the nine broad industrial sectors.[19] The control variables are incorporated as interactions of the initial control with a time-fixed effect or as the initial level itself. In the latter scenario, I employ a parametric linear trend of the outcome variable, contingent upon local demographics and industrial compo-

[17] Instead of adding a fixed effect in the stacked difference form, I control for location-specific trend based on a vector of baseline county characteristics. It is because a shift-share instrument is used to isolate the exogenous variation of the local shocks. Thus, adding another county fixed effect would serve as a bad control as it absorbs the "share" variation of the instrument. See Burga and Turner (2023) for the discussion in a related context.

[18] The latest available data source of education composition before 1985 is the 1980 Census.

[19] These sectors, as defined by the Standard Industrial Classification system, include Agriculture, Forestry, and Fishing; Mining; Manufacturing; Construction; Transportation and Public Utilities; Wholesale Trade; Retail Trade; Finance, Insurance, and Real Estate; Services; and Public Administration. To calculate these sectors' initial employment shares, I divide the total employment in each sector by the county's private non-farm employment.

sition during the initial period. The underlying assumption is that counties would have followed the same trajectory path of the outcomes in the absence of procurement changes if all of these correlates were initially equal. Therefore, the coefficient on the local procurement change, in conjunction with a fitted county linear trend, identifies the effects via the deviations of outcome variables from the projected trend. Furthermore, in some specifications, I control for the initial local employment in the most impacted defense sectors, which is captured by the proportion of manufacturing employment in the MF sectors. In the subsequent main results, I cluster the standard errors at the county level.

1.7 Results

I begin the analysis by examining the impact of the procurement change in the defense-related manufacturing (MF) sector on local private-sector employment, directly and indirectly. An extension of the analysis follows this to consider changes in the number of establishments, wages, and local population adjustments in response to employment opportunity changes.

1.7.1 Defense-Related Heavy Sectors

Table 1.1 presents the estimation results for the baseline regression (Equation 1.1) on sector employment during the drawdown period from 1987 to 1997. The dependent variable is the biennial change of employment share in the flagship sectors of the county population in the sample period. Panel A reports the OLS estimates, and panels B and C report the IV estimations with the endogenous variable instrumented using Equations 1.2 and 1.4, respectively. The results show that the IV estimates in Panels B and C are larger in magnitude than the OLS estimates in Panel A. It supports our hypothesis that the political features of DoD contracts can lead to a downward bias in the OLS estimates.

The F-statistics, presented in the last row of panels B and C, indicate instrument strength for both IV estimations in the first stage. Both reveal a significant and positive relationship between the instrument and the endogenous variable.[20] Taken together, the first-stage strength of the two IVs provides consistent evidence for our hypothesis. Areas with higher proportions of national procurement or greater reliance on flagship defense sectors in the past were more exposed to the change in the national defense regime.

The empirical strategy in section 1.6 involves controlling for a vector of historical county demographics and industrial structures, which were fixed before the national defense drawdown. It is crucial because the predicted procurement changes should only capture the effects of the drawdown and not other correlates. The coefficient estimates are generally robust to the different forms of controls.

Column 2 of Table 1.1 includes controls for the vector of initial county demographics and their interactions with the time-fixed effects. The IV specifications in panels B and C show significant effects of the procurement drawdown on a county's employment in the directly affected sectors. The coefficient estimate of 0.315 in Table

[20] The rule-of-thumb critical value for weak instrument F-statistics is ten, based on the assumption of independent and identically distributed (iid) errors. In certain scenarios where the spatial correlation of errors affects the asymptotic distribution of the estimator, the conventional critical values for the weak instrument test change (Olea and Pflueger, 2013; Burga and Turner, 2023). While the results do not change much when I clustered standard errors at the state level to address the spatial correlation concerns, the statistical significance changes according to the new thresholds of critical values. The first-stage statistics is 28, higher than 23.11, the critical value at 10% significance level, but lower than 37.42, the critical value at 5% significance level.

1.1 (panel C) suggests that a one-thousand dollar per-capita exogenous decrease in a county's procurement in the five flagship defense sectors leads to a reduction in the county's employment in the same sectors directly by 0.315 percentage points. The estimated effects for the initial demographics are slightly smaller than those shown in Column 1, which only includes county- and year- fixed effects.

However, the pre-existing declining trend in manufacturing industries could be a concern. Numerous literature focuses on two primary explanatory factors: the concept of "technology-kills-jobs" (e.g., Brynjolfsson and McAfee, 2012; Acemoglu and Restrepo, 2020), and the notion of "trade-kills-jobs" (e.g., Autor, Dorn, and Hanson, 2013; Hakobyan and McLaren, 2016). Although the diminishing manufacturing employment can partly be attributed to heightened international trade-induced price competition, we are less concerned about the potential estimation biases within our context. Notably, the influence of trade was less pronounced before 2000. If anything, it may have exhibited a stronger correlation with regions reliant on labor-intensive industries, diverging from the conventional procurement-receiving counties.[21]

Differential technology adoption across regions or industries could lead to variations in pre-period trends, thereby influencing the estimates derived from instrumental variables (IV). For instance, places with greater dependence on flagship defense sectors before the nuclear treaty might have embraced automation to a larger extent. Consequently, the IV estimate could be prone to bias arising from factors unrelated to procurement changes, including spurious trends captured by employment fluctuations within these sectors. In response to this concern, I incorporate the initial proportion

[21] The impact of the China trade shock on various industries has been extensively studied, with certain sectors such as luggage, rubber, footwear, games and toys, die-cut paperboard, apparel, textile, and furniture being particularly affected (Autor, Dorn, and Hanson, 2013). In comparison, the defense procurement shock was concentrated in heavy industries, which tend to be more capital-intensive and employ a higher proportion of high-skilled labor. Conversely, light industries are more labor-intensive and have lower shares of college-educated workers.

of manufacturing employment within these sectors, or its interaction with time-fixed effects, as control variables. The coefficients on the initial employment share across these flagship sectors and their interactions with time-fixed effects consistently exhibit positive trends. The estimate suggests that areas with higher reliance on these sectors witnessed more pronounced positive growth in employment levels compared to a hypothetical scenario with lower reliance.[22] Moreover, areas featuring relatively elevated manufacturing employment in these sectors tend to experience greater procurement drawdown from the government. This circumstance could have potentially introduced a downward bias in the estimated effect if the initial industrial structure is not appropriately controlled for.

In Column 3 of Table 1.1, I control for the initial employment allocation across broad sectors in the county, which is particularly important when studying employment change in a more general sector. In Column 4, I instead control for the local reliance of manufacturing employment on the most affected sectors. The inclusion of these controls results in slightly larger estimated effects of procurement shocks, although the magnitudes remain largely comparable. The model in Column 5 of Table 1.1 incorporates both the within-manufacturing employment reliance and the across-sector employment distribution. Additionally, their respective interaction terms with the year-fixed effects are included. The results suggest slightly higher magnitudes of the estimated effect when these potential correlates are controlled for. Moving to Column 6, I introduce a linear county trend contingent on the initial county demographics and the local industrial composition. The estimates with these controls are of similar magnitudes to those in the previous specifications in Columns 4 and 5. This consistency in results across varying specifications underscores the robustness

[22]However, it is important to exercise caution in interpreting these coefficients as they show correlation rather than causality.

of our findings. Furthermore, the coefficient estimates derived from the instrumental variable analysis based on the geographical distribution broadly align with the confidence intervals of the estimates obtained from the Bartik IV method, albeit with larger magnitudes.

The average procurement decrease in the MF sectors per capita is \$20 per period and around \$100 over a ten-year span. Employing a more conservative instrumental variable (IV) estimate extracted from Column 5 in panel C of Table 1.1, we can translate this into a 0.007 percentage points decrease in the county employment share over two years and a reduction of 0.03 percentage points over a decade. Examining Figure 1.1, which illustrates the national procurement drawdown within the five flagship sectors, we observe an aggregate reduction of approximately \$73 billion. By applying the aforementioned conservative estimate from Column 5, we can infer a projected national decrease in employment by 248 thousand individuals within these sectors during the interval between 1987 and 1997.[23]

However, the aggregate change depicted by the national average per capita masks the considerable disparity in employment impacts stemming from varying levels of procurement exposure across different localities. Indeed, the standard deviation of the two-year per capita procurement change within the sample amounts to \$205, surpassing the average change by over tenfold. This statistical observation suggests that a one-standard-deviation reduction in procurement per capita brings about a 0.07 percentage points decrease in local sectoral employment over two years. Similarly, the standard deviation decline of \$672 resulting from a decade-long procurement change corresponds to a 0.23 (calculated as 0.672×0.34) percentage point reduction in local sector employment.

[23] $\$73$ billion $\times 0.0034$ per thousand $= \$248,200$

Looking at the counties positioned within the bottom tenth percentile of the procurement change distribution, which encountered the most pronounced per-capita drawdown, unveils an even more pronounced impact. The average drawdown per person across these 170 counties totaled \$912 between 1987 and 1997. When these figures are applied to the estimates provided in column 5 of Table 1.1, and considering the average county population of 344.6 thousand, it becomes evident that the counties within this 10th percentile cluster alone contributed to a decline in sector employment by 198 thousand individuals[24].

If a crowding-out effect of government spending were indeed in play, as indicated by studies like Cohen, Coval, and Malloy (2011) and Hebous (2016), one might anticipate observing a moderate employment impact. This scenario arises when there is a reduction in demand for output within specific industries, potentially prompting those who lose jobs to seek employment in related sectors. This phenomenon could occur if the labor supply is not perfectly inelastic and various industries have spare capacity to accommodate additional labor.[25] However, the significant employment response observed within the most affected manufacturing (MF) sectors does not suggest so.

1.7.2 Spillover Effects on Other Sectors

The previous subsection indicates that the procurement reduction in the five flagship industries leads to a notable contraction in labor demand, consequently resulting in a decline in employment. Yet, a question remains: do these shocks remain localized within the directly affected industries, or do they manifest broader consequences?

[24] $0.912 \times 0.34\% \times 344,555 \times 170 = 181,627$

[25] For example, Cohen, Coval, and Malloy (2011) find that an influx of government spending is a deterrent to firms' investment because it crowds out the labor supply available. Following this logic, the reverse hypothesis is that the drawdown of government spending relaxes the labor supply, and the rest of the industries in the same sector would absorb part of it.

I explore if the shocks cascade into the broader manufacturing category, encompassing industries with a 2-digit SIC code between 20 and 39. By assessing whether the overall employment change in the manufacturing category is smaller than that within the MF sectors, we can gauge the degree of skill transferability among displaced workers. If the change is smaller, it suggests that the skills of displaced workers are relatively adaptable, and the lost employment finds absorption within other local sectors. In Table 1.2, I present the coefficient estimates on the impacts of the procurement shocks in MF on the broad manufacturing sectors. The OLS estimates are presented in the columns of panel A, whereas the IV estimates are in the two panels below. Correspondingly, the table incorporates the same set of control variables as featured in Table 1.1.

Overall, the impacts on the entire manufacturing sector mirror those derived from the within-sector analysis presented in Table 1.1. Even using the more conservative estimates from the second IV, a one-standard-deviation drop or a drawdown of $672 per capita in the ten years decreases the total local manufacturing employment by 0.17 to 0.28 percentage points. The slightly larger magnitude in Columns 5-6 of Table 1.2 suggests that the other manufacturing sectors struggle to absorb the labor supply surplus stemming from the five heavily affected sectors, and there may be additional adverse spillover effects onto these sectors. One plausible scenario for such spillover effects is a reduction in demand for local suppliers that contribute to the production chain. Alternatively, the local population might adapt to the adverse shocks, leading to a diminished labor supply across various sectors due to households' collective decision-making. However, considering the limited spillover effects on other non-manufacturing sectors, as suggested by Table 1.10, the latter explanation seems less likely. However, the small differences in magnitudes between Tables 1.1 and 1.2

across specifications suggest limited net spillover effects within the manufacturing sector. This observation is consistent with the fact that the demand for manufacturing products is often not locally dependent and the potentially limited skill transferability of workers in the flagship sectors.

The lack of labor transition within the manufacturing sector is consistent with several potential explanations. First, workers may not confine their job search to local labor markets, opting for employment opportunities in other counties. However, as demonstrated in subsequent subsections, I find no evidence of strong population outflows as a reaction to the procurement shocks. A second plausible explanation centers on the possibility that labor markets within the broader manufacturing sectors are inherently limited and lack the capacity to accommodate additional labor. It is particularly likely in the late 1980s and 1990s, during which the landscape witnessed a surge in automation and offshoring and notably contributed to the elimination of routine manual jobs (Acemoglu and Autor, 2011).

To estimate the causal effects of procurement drawdown on the broader scope of manufacturing employment, the empirical approaches hinge on the assumption that manufacturing employment in the affected counties would have displayed comparable growth rates without procurement alterations. However, it is important to acknowledge the potential influence of confounding factors that may have contributed to the decrease in employment beyond the immediate impact of procurement shocks. One such factor could be the secular decline of routine jobs across the broader economy. Echoing concerns regarding the interpretation of direct impacts within the manufacturing sectors, one may argue that the decline in routine jobs influenced the overall decrease in employment.

While a significant structural shift in manufacturing employment has been evident since 1980, several factors mitigate the potential bias stemming from this shift. First, incorporating the initial industrial composition in the specifications helps account for the influence of pre-existing structural changes. Therefore, the coefficient on procurement change captures deviations in employment from the secular industrial trend. Second, high-skilled workers, particularly those with a college education, are generally less inclined to engage in routine tasks compared to their relatively low-skilled counterparts. While we cannot directly observe the workforce composition affected by the procurement shock, the descriptive analysis suggests that the most affected defense sectors typically exhibit a higher proportion of college-educated employees. Consequently, one would expect a more pronounced decline in other manufacturing sectors if structural shifts were the primary driver. The limited effects of procurement on these sectors mitigate such concerns. Another potential concern regarding causality identification arises from the patterns of globalization, wherein the loss of employment in manufacturing sectors could be attributed to increased competitiveness of imports, as indicated by Autor, Dorn, and Hanson (2013). However, there was a limited overlap between the industries affected by the defense procurement drawdown and those exposed to import competitiveness, and the coefficients are robust after we control for the local industrial composition.

Given the significant impact of the procurement drawdown on manufacturing employment, a pertinent question emerges: were non-manufacturing sectors equipped to absorb the resulting employment losses? To investigate this, I use the same specifications to assess whether there is a commensurate decrease in the county's total private-sector non-farm employment. The findings, as detailed in Table 1.10, indicate a slightly more pronounced decline in private-sector non-farm employment.

A further breakdown of these sectors shows significant impacts in one non-tradable sector, specifically the retail trade sector.[26] Detailed coefficient estimates within the retail sector are presented in Table 1.3. Columns 1 to 6 maintain the same vector of controls as Tables 1.1 and 1.2. The coefficients are consistent across columns, although they lose some precision when more controls are added. Broadly, these coefficients suggest that a $1,000 decrease in contract values per capita corresponds to a decline in retail-sector employment ranging from 0.09 to 0.13 percentage points. This magnitude approximates one-third of the within-sector effects. As the non-traded sector's labor demand is directly associated with the local product demand, I show below that the negative demand spillover is attributed to a diminished customer base, not a reduced individual spending power.

1.7.3 Establishment and Average Earning

To gain a comprehensive understanding, I extend the investigation to encompass the number of establishments and wage changes. The outcomes are detailed in Table 1.11 in the Appendix. The estimations delineate changes in the total establishment count and establishment numbers within the MF sector, manufacturing sector, and retail sector, respectively. The findings unveil no substantial evidence that the defense spending cuts significantly reduced manufacturing or other sector establishments. The estimates for the manufacturing sector are precisely zero, underscored by both minimal economic significance and standard errors. These coefficients suggest that firms within the MF sector tend to curtail their employment at the intensive margin, opting to reduce staff rather than closing operations at the extensive margin. Comparable conclusions extend to the retail sector, although one might anticipate

[26]The retail trade sector contains industries with SIC codes between 52 and 59. I checked using the sector employment directly from BEA's REA and found similar effects by aggregating the granular employment from the imputed CBP database.

more frequent turnovers due to the prevalence of smaller-scale establishments in this sector. In contrast to the pronounced responses observed in employment counts, no discernible change exists in the growth rates of a county's average personal income, as shown in Table 1.4.[27] I will discuss in the following subsection that besides a downward wage rigidity, a response on the supply side may drive the null effect on the average earning.

1.7.4 Population Relocation

Following an adverse labor demand shock, the equilibrium employment in the local labor market declined. To explore the responses in labor supply, I examine the effects on both out-migration and in-migration flows separately. Table 1.5 presents the coefficient estimates from equation 1.1. The outcome variable is the number change of individuals moving out of a county, normalized by the county's beginning-of-the-period population. Employing the Bartik-type IV method across various specifications in the columns, the results consistently indicate insignificance, suggesting a lack of discernible effects on out-migration behaviors. Similarly, when employing the first IV, which capitalizes on the initial geographical procurement variation, the results generally remain statistically insignificant, particularly when controlling for the initial local industrial structure.[28] Collectively, these estimates suggest limited evidence

[27]The average earning is computed by dividing the county's total personal income and salary employment from BEA. Same as the DoD contracts, the income was converted to the real value using the GDP price deflator with the base year 2000.

[28]The coefficients, if anything, reveal a positive relationship between procurement change and outflows, albeit largely statistically insignificant. A plausible interpretation could be that counties with a historically higher share of DoD contracts, such as those in California and New York, might have experienced relatively more substantial gross outflows due to weakened employment opportunities in the manufacturing sector. This perspective gains further support from the observation that the marginal significance disappears when incorporating a county-specific trend contingent on the industrial employment structure.

to support the notion that individuals directly affected by the labor demand shock sought alternative opportunities by relocating outside of the county.

In contrast to the relatively stagnant out-migration, there were positive impacts on the county-level in-migration rates. Table 1.6 summarizes the estimates of the effect of procurement drawdown on this aspect of labor supply pool transformation. The estimates illustrate that, in response to a $1,000 reduction in contract value per capita, the influx of individuals relocating to the county diminished by approximately 0.12 to 0.16 individuals per 100 residents. This number corresponds to approximately 2 to 3% of the sample's mean inflows-to-population ratio, which is 5 individuals per 100 residents. Similar to what we observed from the equilibrium employment, the two IV approaches address the downward bias in the OLS estimation. While the estimates exhibit relatively larger magnitudes in scenarios where no county characteristics are controlled, the coefficient estimates are generally robust to whether we assume a linear trend or a more flexible form of the time-varying impacts of these baseline characteristics.

The empirical evidence indicates a noteworthy labor supply adaptation in response to the labor demand shock. However, this adjustment does not appear to stem from an active pursuit of opportunities elsewhere by the affected individuals. Instead, the labor supply response manifests primarily through a substantial reduction in inward migration flows. Despite the prompt supply adjustment, the reduced population inflow could yield enduring consequences for these counties. Notably, a significant proportion of counties grappling with substantial per-capita drawdown were small to medium sized, with populations numbering fewer than 100,000 residents. While the contraction of manufacturing job opportunities may not be the sole contributing factor, the resultant decline in the attractiveness to individuals relocating to these places could pose challenges for the future trajectory of social and economic growth.

Geographical relocation, however, represents one among several conceivable strategies for local labor supply pool adjustments. An alternate approach to reducing the local labor supply involves the exit of individuals from the workforce. While the most straightforward way to investigate this hypothesis is to see how the county procurement drawdown affected the county labor force participation, the available dataset from the BLS does not provide the county-level participation rates for our study period. In light of this constraint, I adopt an alternative strategy by referring to the disability and retirement sampling data sourced from OASDI at the state level. The literature has discussed extensively how economic downturns originated from diverse market fluctuations can lead to specific forms of non-participation, such as retirement patterns or DI enrollment (Coile and Levine, 2011; Chan and Huff Stevens, 2001; Autor and Duggan, 2003). By examining changes in state procurement and their impacts on these two margins, we can extract insights into potential shifts in labor force participation. Tables 1.13 and 1.14 in the Appendix show the coefficients for this analysis. The specifications align with the same form as equation 1.1, with the distinction that variations have been aggregated at the state level.[29] The results reveal no significant evidence of opting for these forms of non-participation.

[29]The U.S. Department of Health and Human Services publishes the annual statistical book of Social Security Beneficiaries by State and County. While Optical Character Recognition (OCR) allows the extraction of county statistics, the relatively poor quality of earlier years creates measurement errors, especially in recognizing zeros and commas. In realization of the data accuracy issues, I instead extract the state-level person counts on those receiving social security benefits such as retirement and DI.

1.8 Robustness Checks

1.8.1 Falsification Test

To ensure that the employment changes observed after the Cold War were not simply a continuation of the pre-existing trends, I conducted a pre-period analysis by regressing the average biennial employment change from 1980 to 1985 against the future procurement drawdown commencing in 1987. The results of this analysis, as shown in Table 1.7, reveal a robust negative relationship between procurement cuts and employment growth within the flagship industries, as well as across the broader spectrum of manufacturing employment, during the pre-period. This finding indicates that the employment trajectories witnessed in the pre-period diverged from the subsequent employment shifts observed after the Cold War. The contrast in employment growth patterns between these segments helps rule out the possibility that the post-Cold War changes were merely a continuation of pre-existing trends in the local labor market. Moreover, this inverse impact on manufacturing employment aligns with the historical context that the Reagan administration's persistent escalation of defense spending led to higher employment growth, even during a period experiencing economic recession triggered by the energy crisis in the early 1980s.

1.8.2 Non-base Counties

Figures 1.6 and 1.7 illustrate that most counties that received DoD contracts did not have a military base. However, it is important to consider the possibility that the procurement impact may have been more severe and concentrated in counties with military bases, which could have absorbed the majority of the labor market consequences. To address this concern, I conduct a robustness check by focusing exclusively on counties without military bases. The Base Structure Report 1989 identifies military bases, and the sample is validated using military staff counts reported by the BEA's county-military industry employment data. Table 1.8 presents the results for estimated impacts on the primary outcomes, with the specifications using the IVs and including the county's initial correlates (as in Column 6 in the previous tables).

The coefficients in Table 1.8 remain significant, although the magnitudes are smaller than those estimated from the whole sample. The findings indicate a sizable drop (0.234 percentage points) in employment in the flagship defense industries. While no significant spillover effect is observed in other sectors, the analysis reveals that population inflows in these counties demonstrate a rapid adjustment. Despite experiencing a milder impact than their base-containing counterparts, the results still suggest that these places were less attractive to potential labor supply in the wake of a defense procurement shock.

1.8.3 Synthetic Controls Using Non-contracting Counties

Thus far, our sample has been confined to counties that procured a DoD contract in the flagship industries at some point within our sample period. This strategic approach encompasses around 1,700 counties and alleviates concerns of comparing counties with systematic differences associated with defense funding support. I extend our analysis further to bolster the robustness of our primary regression results. This expanded sample incorporates counties that remained unconnected by DoD contracts throughout the sample period. This augmented sample allows for the examination of outcomes in counties unaffected by the defense policy and also accounts for the historical trajectories of each county.

To address concerns surrounding comparability between counties that have procured DoD contracts (treated) and those that have not (control), a matching approach is employed. These control counties are matched to contracting (treated) counties based on their previous outcomes and local characteristics, following the approach used by Abadie, Diamond, and Hainmueller (2010), Deryugina, MacKay, and Reif (2020) and Zou (2018). This matching process establishes comparisons while accounting for variations in local characteristics and historical trends. Practically, I utilize a set of county correlates that might engender differences between the two categories of counties. These characteristics include variables such as the share of the population employed in the flagship defense industries in 1980 - 1985 and the county-level demographic and socioeconomic controls. After incorporating these county-level characteristics, I reweigh each control county unit to construct a synthetic control for each treated county. By aligning the post-drawdown trends with their pre-drawdown counterparts, as indicated by these selected controls, the synthetic control provides a counterfactual scenario that the local labor market outcomes would have prevailed without the drawdown. The underlying assumption of such approach, nevertheless, relies on the extent to which these selected controls help predict the changes in the outcome paths.

Specifically, for a DoD-contracting county $i \in$ 1, 2, ... I and non-contracting county $j \in$ 1, 2, ... J in year t, I first estimate the following specification:

$$\Delta Y_{it} = \alpha + \beta \Delta Proc_{it} + \theta_{it} + \Delta\epsilon_{it} \tag{1.5}$$

$$
\begin{aligned}
w_{i1}\Delta Y_{1t} &= w_{i1}\alpha + \beta w_{i1}\Delta Proc_{1t} + w_{i1}\theta_{1t} + \Delta w_{i1}\epsilon_{1t} \\
w_{i2}\Delta Y_{2t} &= w_{i2}\alpha + \beta w_{i2}\Delta Proc_{2t} + w_{i2}\theta_{2t} + \Delta w_{i2}\epsilon_{2t} \\
&... \\
w_{iJ}\Delta Y_{Jt} &= w_{iJ}\alpha + \beta w_{iJ}\Delta Proc_{Jt} + w_{iJ}\theta_{Jt} + \Delta w_{iJ}\epsilon_{Jt},
\end{aligned}
\tag{1.6}
$$

where Y_{it} is a set of outcome variables of a county i. $\Delta Proc_{jt}$ is zero because the non-contracting counties did not receive any contracts. Both are normalized by the county's population. For each contracting county i, I set the conditional change of its outcome over time to be θ_{it}, which varies both at the county and the time levels and is assumed to be determined by a vector of county's initial characteristics.[30] For each given contracting county i, I construct a series of weights $\{w_{i1}, w_{i2}, ...w_{iJ}\}$ assigned to each non-contracting county $j \in J$ so that the outcome trajectories conditional on the procurement amount are similar across the counties in the pre-drawdown period. Mathematically, the weights are calculated so that:

$$
W_{iJ} = \underset{w_{ij} \in W_{iJ}}{\operatorname{argmin}} \left\| \theta_{it} - \sum_j w_{ij}\theta_{jt} \right\| \tag{1.7}
$$

These weights are determined by minimizing the distance (sum of least squares) between the outcomes of these counties and the contracting county before the drawdown occurs. The obtained weights are then used to create a synthetic control group for the contracting counties. Practically, the objective function is to minimize the distance in a vector of outcomes and the initial county characteristics between a treated

[30] The selected controls follow the ones in the main specification, and include: the industrial, gender, age, race, education composition of the local population, and the share of the population in the flagship manufacturing industries in the pre-drawdown period.

county and the control county units in the pre-drawdown period. The outcomes of the non-contracting counties are aggregated based on their weights to form a synthetic control group that exhibits a similar pre-drawdown trend in outcomes to the contracting county.[31]

I then estimate the impacts of the procurement changes conditional on these county-specific trends. Specifically, I employ the TSLS approach with the following specification:

$$
\begin{aligned}
(\Delta Y_{it} - \sum_{j \in J} \hat{w_{ij}} \Delta Y_{jt}) &= \widetilde{\alpha} + \beta(\Delta Proc_{it} - \sum_{j \in J} \hat{w_{ij}} \Delta Proc_{jt}) + (\theta_{it} - \sum_{j} \hat{w_{ij}} \theta_{jt}) + \Delta \widetilde{\epsilon_{it}} \\
\widetilde{\Delta Y_{it}} &= \widetilde{\alpha} + \beta \Delta Proc_{it} + \Delta \widetilde{\epsilon_{it}}
\end{aligned}
\tag{1.8}
$$

The specification incorporates the equations 1.5 - 1.7. Using the weights calculated from equation 1.7 and the observed outcomes of the non-contracting counties, I first estimate the counterfactual outcome for county i, denoted as $\Delta \hat{Y_{it}}(0) = \sum_{j \in J} \hat{w_{ij}} \Delta Y_{jt}$. It would represent the estimated change in outcome for county i if the defense policy did not undergo any changes. In this estimation, $\widetilde{\alpha}$ and $\Delta \widetilde{\epsilon_{it}}$ represent

[31]However, it is important to acknowledge certain limitations of this approach. While it introduces flexibility in capturing the effects of local characteristics on outcome paths, it also relies on additional assumptions. For instance, it assumes that the comparability between the contracting and non-contracting counties remains consistent during the later period. Therefore, the synthetic control strategy measures the impacts of procurement in the flagship industries to the extent that the contracting counties would have followed a similar path as the non-contracting counties without any policy change. One potential issue with this methodology is the existence of unobserved factors that may have caused the contracting counties to experience slower growth in the early 1980s. It could occur in our context if the employment growth rates of some contract-receiving counties were more influenced by macroeconomic conditions during the recession in the early 1980s. In this case, tracking the outcome paths for these counties in the earlier period might have inadvertently over-captured the downturn that was not directly related to downsizing procurement, leading to an underestimation of the labor market impacts resulting from the procurement.

the differenced intercept and residual terms, respectively. As a result, the dependent variable becomes the difference between the observed outcome and the imputed outcome for a contracting county i, and the key explanatory variable remains the same as in equation 1.1.

To address the endogeneity issue of the local procurement change, I apply the instrumental variable (IV) approach similar to our main analysis. However, one concern with this two-step approach is that the usual county-level clustered standard errors in the second stage do not fully account for the uncertainties associated with the weights estimated by the synthetic control approach. To mitigate this concern, I employ bootstrapped standard errors obtained from the empirical distributions of the estimated coefficients.[32]

Table 1.9 presents the coefficient estimates of the primary outcomes examined in this study. The results indicate a significant adverse effect on local employment due to the reduction in procurement contracts in the flagship industries. Specifically, Column 1 shows that a $1,000 contract cut per person decreases local employment in the flagship industries by approximately 0.44 percentage points, slightly larger than the estimate of 0.35 obtained from equation 1.1. Interestingly, with this specification, despite a larger impact on the directly affected industries, there appears to be an insignificant change in other sectors. Moreover, consistent with earlier findings on labor supply adjustments, this non-parametric method also unveils a noteworthy reduction

[32]The bootstrapped standard errors are empirically obtained from 100 iterations of the pseudo samples. Specifically, I perform 100 iterations of bootstrap samples. In each iteration, I independently resample the contracting (treated) and non-contracting (control) county groups, taking a random subsample of 50% from each group with replacement. These randomly selected non-contracting counties form the synthetic control for each contracting county in the pseudo sample. Using the pseudo samples, I estimate equations 1.7 and then 1.8 using the IV approach in equation 1.4. The bootstrapped coefficients reveal the standard deviation of the estimated coefficients from the pseudo-samples. This procedure empirically accounts for the uncertainties in the weights and provides more accurate standard errors, enhancing the reliability of the estimation results.

in population inflows as a reaction to the local procurement shock. In summary, although the estimates have wider confidence intervals, they lead to a similar conclusion as the previous findings: The reduction in employment in the flagship industries prompts an immediate labor supply response but is steered mainly by a decline in labor inflows.

1.9 Conclusion

In the last twenty years, the overall defense expenditure by the U.S. Department of Defense has undergone a noteworthy expansion, more than doubling its previous levels. This upward trajectory in defense spending is commonly attributed to assessments of potential threats to U.S. national security. However, with substantial defense spending and mounting pressure to reduce federal budget deficits, the future of U.S. defense spending has become progressively uncertain.

Despite the importance of defense spending, the literature has under-explored how a changing defense policy might affect local job markets. This paper aims to fill this gap by examining a unique historical period characterized by a significant and prolonged defense drawdown. To address potential endogeneity in contract distribution across locations, I employ an instrumental variable (IV) approach to isolate the plausibly exogenous portion of the procurement shocks. This approach is based on the concept that national industry-specific procurement shocks propagate to local economies based on their initial procurement composition. The empirical strategy uses various techniques to consider potential county-specific trends, ensuring that the changes in employment after the Cold War do not merely reflect a continuation of existing trends.

The findings from this study suggest that shifts in defense policy could result in lasting demand disruptions for industries previously reliant on defense procurement. There is also suggestive evidence that other manufacturing sectors could not absorb the displaced employment in those directly affected sectors. Furthermore, the results indicate a potential negative influence on the local non-tradable. Nonetheless, the employment impacts appear to be limited. Moreover, the findings reveal a relatively prompt labor supply response. However, when exploring how the local population responded to the reduced labor demand, the results indicate a reduction in inward migration, while there is no noticeable increase in outward migration. These findings lean towards the usage of targeted approaches, such as specialized job training or re-employment initiatives, as opposed to a broad place-based policy.

1.10 Tables

Table 1.1: Estimated Effects of Procurement Drawdown on Local Employment in the Flagship Manufacturing Sectors, MF

Dependent variable: change in employment (MF sector) to population (per 100 people)						
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
ΔProcurement in MF sector,	0.256***	0.225***	0.232***	0.234***	0.240***	0.237***
$1,000 per person	(0.043)	(0.041)	(0.042)	(0.040)	(0.042)	(0.041)
Panel B: IV (Geographical distribution)						
ΔProcurement in MF sector,	0.513***	0.446***	0.459***	0.450***	0.461***	0.475***
$1,000 per person	(0.119)	(0.121)	(0.120)	(0.120)	(0.119)	(0.116)
First stage F statistic	191.895	192.211	195.199	193.654	198.304	194.067
Panel C: IV (Bartik)						
ΔProcurement in MF sector,	0.402***	0.315***	0.325***	0.328***	0.337***	0.348***
$1,000 per person	(0.113)	(0.087)	(0.092)	(0.089)	(0.095)	(0.098)
First stage F statistic	28.026	28.754	28.604	29.010	28.906	28.859
Demographic controls ×year FE		Y	Y	Y	Y	
Share of private non-farm employment in the broad sectors ×year FE			Y		Y	
Share of manufacturing employment in the most affected sectors ×year FE				Y	Y	
Demographic controls (trend)						Y
Share of private non-farm employment in the broad sectors (trend)						Y
Share of manufacturing employment in the most affected sectors (trend)						Y
Year FE	Y					Y
No. obs.	8510	8510	8510	8510	8510	8510

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the county level in parenthesis.

Table 1.2: Estimated Effects of Procurement Drawdown on Local Employment in Manufacturing Sector

Dependent variable: change in employment (manufacturing sector) to population (per 100 people)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
ΔProcurement in MF sector,	0.261***	0.195***	0.210***	0.208***	0.222***	0.230***
$1,000 per person	(0.054)	(0.052)	(0.051)	(0.050)	(0.050)	(0.051)
Panel B: IV (Geographical distribution)						
ΔProcurement in MF sector,	0.691***	0.543***	0.590***	0.553***	0.598***	0.642***
$1,000 per person	(0.141)	(0.117)	(0.120)	(0.119)	(0.123)	(0.131)
First stage F statistic	191.895	192.211	195.199	193.654	197.586	194.067
Panel C: IV (Bartik)						
ΔProcurement in MF sector,	0.495***	0.302***	0.335***	0.325***	0.356***	0.406***
$1,000 per person	(0.161)	(0.098)	(0.110)	(0.105)	(0.116)	(0.134)
First stage F statistic	28.026	28.754	28.604	29.010	28.906	28.859
Demographic controls ×year FE		Y	Y	Y	Y	
Share of private non-farm employment in the broad sectors ×year FE			Y		Y	
Share of manufacturing employment in the most affected sectors ×year FE				Y	Y	
Demographic controls (trend)						Y
Share of private non-farm employment in the broad sectors (trend)						Y
Share of manufacturing employment in the most affected sectors (trend)						Y
Year FE	Y					Y
No. obs.	8510	8510	8510	8510	8510	8510

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the county level in parenthesis.

Table 1.3: Estimated Effects of Procurement Drawdown on Local Employment in Retail Sector

Dependent variable: change in employment (retail sector) to population (per 100 people)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
ΔProcurement in MF sector,	0.040	0.042	0.038	0.025	0.024	0.026
\$1,000 per person	(0.029)	(0.026)	(0.026)	(0.028)	(0.028)	(0.029)
Panel B: IV (Geographical distribution)						
ΔProcurement in MF sector,	0.133***	0.145**	0.133**	0.119**	0.118**	0.118**
\$1,000 per person	(0.047)	(0.059)	(0.055)	(0.052)	(0.052)	(0.051)
First stage F statistic	191.895	192.211	195.717	193.654	197.586	194.067
Panel C: IV (Bartik)						
ΔProcurement in MF sector,	0.100**	0.129**	0.118**	0.103**	0.099**	0.098**
\$1,000 per person	(0.042)	(0.052)	(0.049)	(0.046)	(0.045)	(0.044)
First stage F statistic	28.026	28.754	28.604	29.010	28.906	28.859
Demographic controls ×year FE		Y	Y	Y	Y	
Share of private non-farm employment in the broad sectors ×year FE			Y		Y	
Share of manufacturing employment in the most affected sectors ×year FE				Y	Y	
Demographic controls (trend)						Y
Share of private non-farm employment in the broad sectors (trend)						Y
Share of manufacturing employment in the most affected sectors (trend)						Y
Year FE	Y					Y
No. obs.	8510	8510	8510	8510	8510	8510

*** p<0.01, ** p<0.05, * p<0.1 County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the county level in parenthesis.

Table 1.4: Estimated Effects of Procurement Drawdown on County's Average Earning

Dependent variable: percentage change in county's average wage

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
ΔProcurement in MF sector,	-0.001	0.001	0.001	-0.000	-0.000	0.000
$1,000 per person	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Panel B: IV (Geographical distribution)						
ΔProcurement in MF sector,	-0.001	0.002	0.000	0.000	0.001	0.001
$1,000 per person	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
First stage F statistic	191.895	192.211	199.097	193.654	197.586	194.067
Panel C: IV (Bartik)						
ΔProcurement in MF sector,	-0.005	-0.002	-0.003	-0.002	-0.002	-0.002
$1,000 per person	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
First stage F statistic	28.026	28.754	28.604	29.010	28.906	28.859
Demographic controls ×year FE		Y	Y	Y	Y	
Share of private non-farm employment in the broad sectors ×year FE			Y		Y	
Share of manufacturing employment in the most affected sectors ×year FE				Y	Y	
Demographic controls (trend)						Y
Share of private non-farm employment in the broad sectors (trend)						Y
Share of manufacturing employment in the most affected sectors (trend)						Y
Year FE	Y					Y
No. obs.	8510	8510	8510	8510	8510	8510

*** p<0.01, ** p<0.05, * p<0.1 The dependent variable is the periodic change in $ln(wage_t)$. County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the county level in parenthesis.

Table 1.5: Estimated Effects of Procurement Drawdown on Migration Outflows

Dependent variable: change in migration outflow to population (per 100 people)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
ΔProcurement in MF sector,	0.014	0.010	0.004	-0.012	-0.013	-0.009
$1,000 per person	(0.028)	(0.032)	(0.031)	(0.029)	(0.029)	(0.028)
Panel B: IV (Geographical distribution)						
ΔProcurement in MF sector,	0.061*	0.071*	0.058	0.037	0.038	0.026
$1,000 per person	(0.034)	(0.039)	(0.036)	(0.035)	(0.035)	(0.032)
First stage F statistic	191.895	192.211	195.199	193.654	197.586	194.067
Panel C: IV (Bartik)						
ΔProcurement in MF sector,	0.021	0.016	0.003	-0.008	-0.009	-0.014
$1,000 per person	(0.042)	(0.047)	(0.044)	(0.044)	(0.044)	(0.039)
First stage F statistic	28.026	28.754	28.604	29.010	28.906	28.859
Demographic controls ×year FE		Y	Y	Y	Y	
Share of private non-farm employment in the broad sectors ×year FE			Y		Y	
Share of manufacturing employment in the most affected sectors ×year FE				Y	Y	
Demographic controls (trend)						Y
Share of private non-farm employment in the broad sectors (trend)						Y
Share of manufacturing employment in the most affected sectors (trend)						Y
Year FE	Y					Y
No. obs.	8510	8510	8510	8510	8510	8510

*** p<0.01, ** p<0.05, * p<0.1 County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the county level in parenthesis.

Table 1.6: Estimated Effects of Procurement Drawdown on Migration Inflows

Dependent variable: change in migration inflow to population (per 100 people)						
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
ΔProcurement in MF sector,	0.138***	0.107***	0.098***	0.098***	0.091***	0.108***
$1,000 per person	(0.029)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Panel B: IV (Geographical distribution)						
ΔProcurement in MF sector,	0.182**	0.129*	0.116*	0.103	0.095	0.114*
$1,000 per person	(0.085)	(0.071)	(0.070)	(0.064)	(0.063)	(0.066)
First stage F statistic	191.895	192.211	195.199	193.654	197.586	194.067
Panel C: IV (Bartik)						
ΔProcurement in MF sector,	0.223***	0.144***	0.126**	0.132***	0.119**	0.158***
$1,000 per person	(0.074)	(0.053)	(0.051)	(0.051)	(0.050)	(0.054)
First stage F statistic	28.026	28.754	28.604	29.010	28.906	28.859
Demographic controls ×year FE		Y	Y	Y	Y	
Share of private non-farm employment in the broad sectors ×year FE			Y		Y	
Share of manufacturing employment in the most affected sectors ×year FE				Y	Y	
Demographic controls (trend)						Y
Share of private non-farm employment in the broad sectors (trend)						Y
Share of manufacturing employment in the most affected sectors (trend)						Y
Year FE	Y					Y
No. obs.	8510	8510	8510	8510	8510	8510

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the county level in parenthesis.

Table 1.7: Falsification Tests: Pre-period Employment Change on Procurement Changes

	M5		M		R	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV (Geographical distribution)						
ΔProcurement in MF sector,	-0.592**	-0.587**	-0.546***	-0.542***	-0.210**	-0.208**
$1,000 per person	(0.238)	(0.237)	(0.173)	(0.172)	(0.095)	(0.094)
First stage: KP Wald F statistic	198.326	194.092	198.326	194.092	198.326	194.092
Panel C: IV (Bartik)						
ΔProcurement in MF sector,	-0.357***	-0.352***	-0.352***	-0.348***	-0.170**	-0.168**
$1,000 per person	(0.110)	(0.109)	(0.099)	(0.098)	(0.076)	(0.075)
First stage: KP Wald F statistic	28.914	28.863	28.914	28.863	28.914	28.863
Demographic controls ×year FE	Y		Y		Y	
Share of private non-farm employment in the broad sectors ×year FE	Y		Y		Y	
Share of manufacturing employment in the most affected sectors ×year FE	Y		Y		Y	
Demographic controls (trend)		Y		Y		Y
Share of private non-farm employment in the broad sectors (trend)		Y		Y		Y
Share of manufacturing employment in the most affected sectors (trend)		Y		Y		Y
Year FE		Y		Y		Y
No. obs.	8495	8495	8495	8495	8495	8495

*** p<0.01, ** p<0.05, * p<0.1 County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the county level in parenthesis.

Table 1.8: Estimated Effects of Procurement Drawdown, Non-base Counties

	(1) Flagship Industries	(2) Manufacturing	(3) Retail	(4) Out-migration	(5) In-migration
Panel A: OLS					
ΔProcurement in MF sector,	0.206***	0.187***	0.014	-0.020	0.098***
$1,000 per person	(0.042)	(0.056)	(0.036)	(0.026)	(0.026)
Panel B: IV (Geographical distribution)					
ΔProcurement in MF sector,	0.386***	0.527***	0.073***	0.008	0.067
$1,000 per person	(0.145)	(0.121)	(0.028)	(0.030)	(0.065)
First stage: KP Wald F statistic	269.584	269.584	269.584	269.584	269.584
Panel C: IV (Bartik)					
ΔProcurement in MF sector,	0.234***	0.253***	0.033	-0.044	0.103**
$1,000 per person	(0.071)	(0.093)	(0.033)	(0.032)	(0.043)
First stage: KP Wald F statistic	37.415	37.415	37.415	37.415	37.415
Demographic controls (trend)	Y	Y	Y	Y	Y
Share of private non-farm employment in the broad sectors (trend)	Y	Y	Y	Y	Y
Share of manufacturing employment in the most affected sectors (trend)	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
No. obs.	7391	7391	7391	7391	7391

*** p<0.01, ** p<0.05, * p<0.1 County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the county level in parenthesis.

Table 1.9: Estimated Effects of Procurement Drawdown, Non-parametric Trend

	Flagship Industries	Manufacturing	Retail	Out-migration	In-migration
IV (Geographical distribution):					
Delta Procurement in MF sector, $1,000 per person (mean)	0.513	0.011	0.002	0.013	0.196
90% Confidence Interval	[-0.080, 1.062]	[-0.002, 0.026]	[0.000, 0.004]	[-0.223, 0.308]	[-0.193, 0.621]
IV (Bartik):					
Delta Procurement in MF sector, $1,000 per person (mean)	0.439	0.008	0.001	0.054	0.220
90% Confidence Interval	[0.167, 0.851]	[-0.001, 0.018]	[0.000,0.003]	[-0.248, 0.216]	[-0.042, 0.553]
Secular trend using synthetic control	Y	Y	Y	Y	Y
Estimation method	TSLS	TSLS	TSLS	TSLS	TSLS

County-level controls used for weight calculation in the control group include the share of population employed in the flagship defense industries (MF) in 1980 - 1985, age, gender, race and education composition of the population, and the industrial composition in the year of 1985. Coefficients in the brackets are extracted from the 100 bootstrapped samples that are randomly drawn from 50% of the contracting and non-contracting counties with replacement.

1.11 Figures

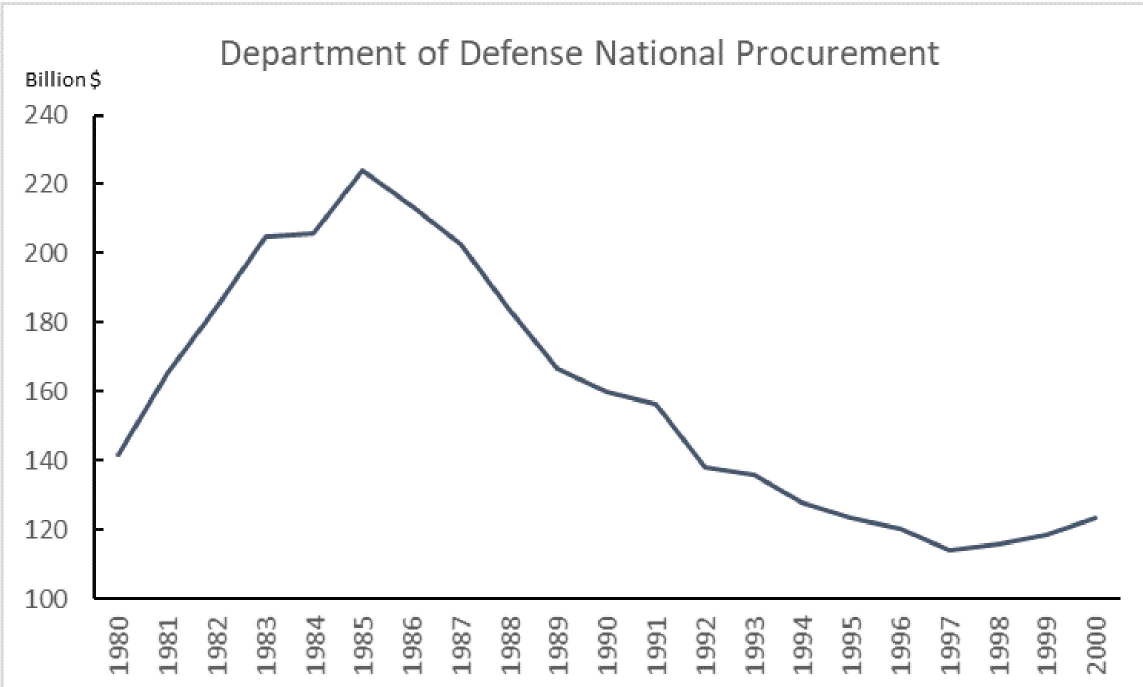

Figure 1.1: The Department of Defense National Procurement, billion $, 1980 - 2000

Note: The Department of Defense national procurement 1980 - 2000, billion $, adjusted for inflation using the base year 2000. *Data source: Department of Defense 350 forms, and author's tabulation*

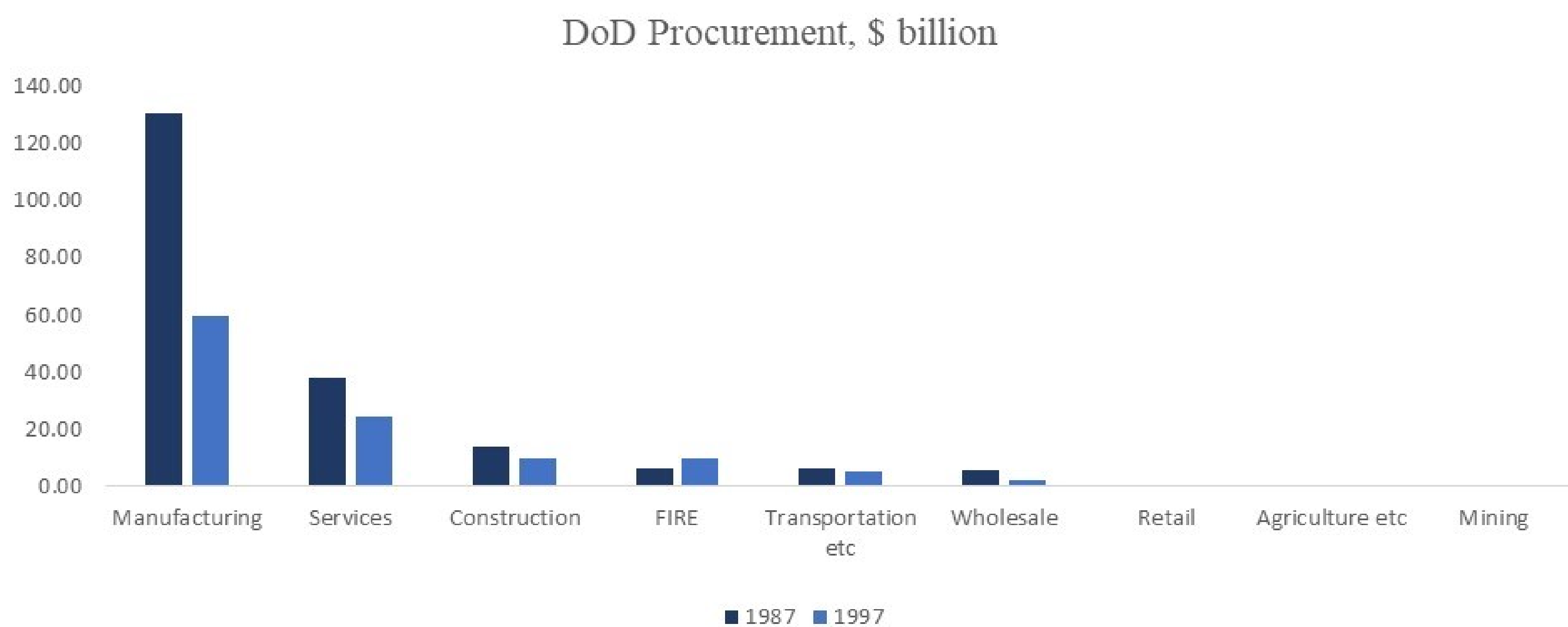

Figure 1.2: National Amount of DoD Procurement by Sector, billion $, 1987 and 1997

Note: The Department of Defense national procurement 1980 - 2000, billion $, adjusted for inflation using the base year 2000. *Data source: Department of Defense 350 forms, and author's tabulation*

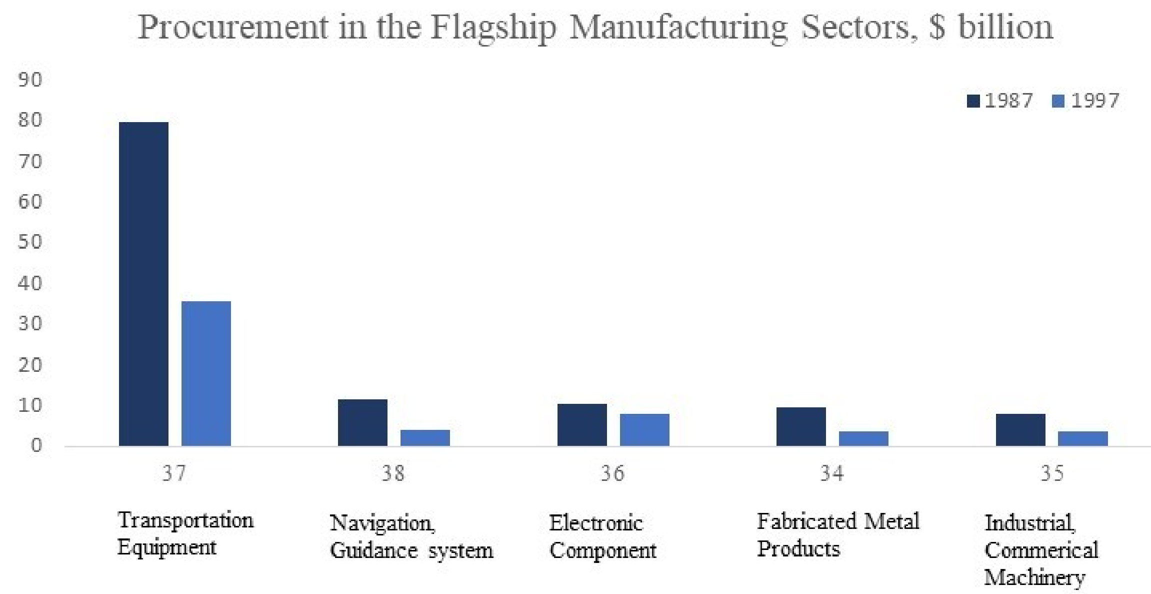

Figure 1.3: National Amount of Procurement Across Five Manufacturing Industries Contracting the Most with DoD, billion $, 1987 and 1997

Note: The top five manufacturing industries that received the most from the DoD procurement, 1987 and 1997, billion $, adjusted for inflation using the base year 2000. The horizontal axis indicates the 1987 4-digit SIC. Industries from left to right cover establishments in: "manufacturing equipment for transportation of passengers and cargo by land, air, and water" (SIC 37); "manufacturing geophysical equipment; search, detection, navigation, and guidance systems" (SIC 38); "manufacturing electronic and other electrical equipment and components, except computer equipment" (SIC 36); "manufacturing fabricated metal products" (SIC 34); and "manufacturing industrial and commercial machinery and computer equipment, e.g. engines" (SIC 35). *Data source: Department of Defense 350 forms, and author's tabulation*

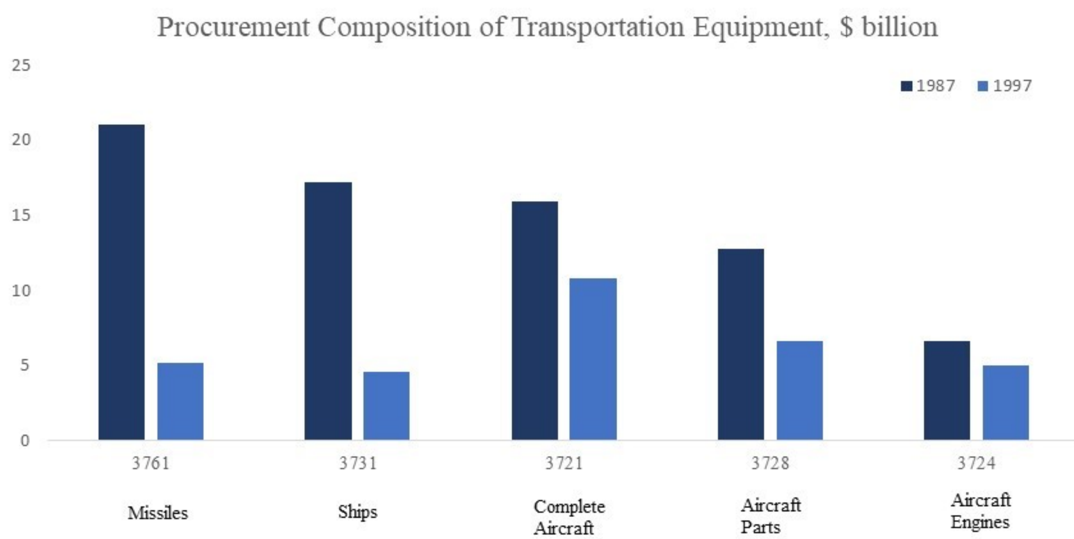

Figure 1.4: National Amount of Procurement Across Five Top Industries in the Transportation Equipment Industry (SIC 37), billion $, 1987 and 1997

Note: The graph presents the five top industries that belong to the industry group (SIC 37) and received the most from the DoD procurement, 1987 and 1997, billion $, adjusted for inflation using the base year 2000. The horizontal axis indicates the 1987 4-digit SIC. Industries from left to right cover establishments in: “building Guided missiles and space vehicles” (SIC 3761); “building and repairing ships” (SIC 3731); “manufacturing complete aircraft” (SIC 3721); “manufacturing aircraft parts and auxiliary equipment” (SIC 3728); and “manufacturing aircraft engines and engine parts” (SIC 3724). *Data source: Department of Defense 350 forms, and author’s tabulation*

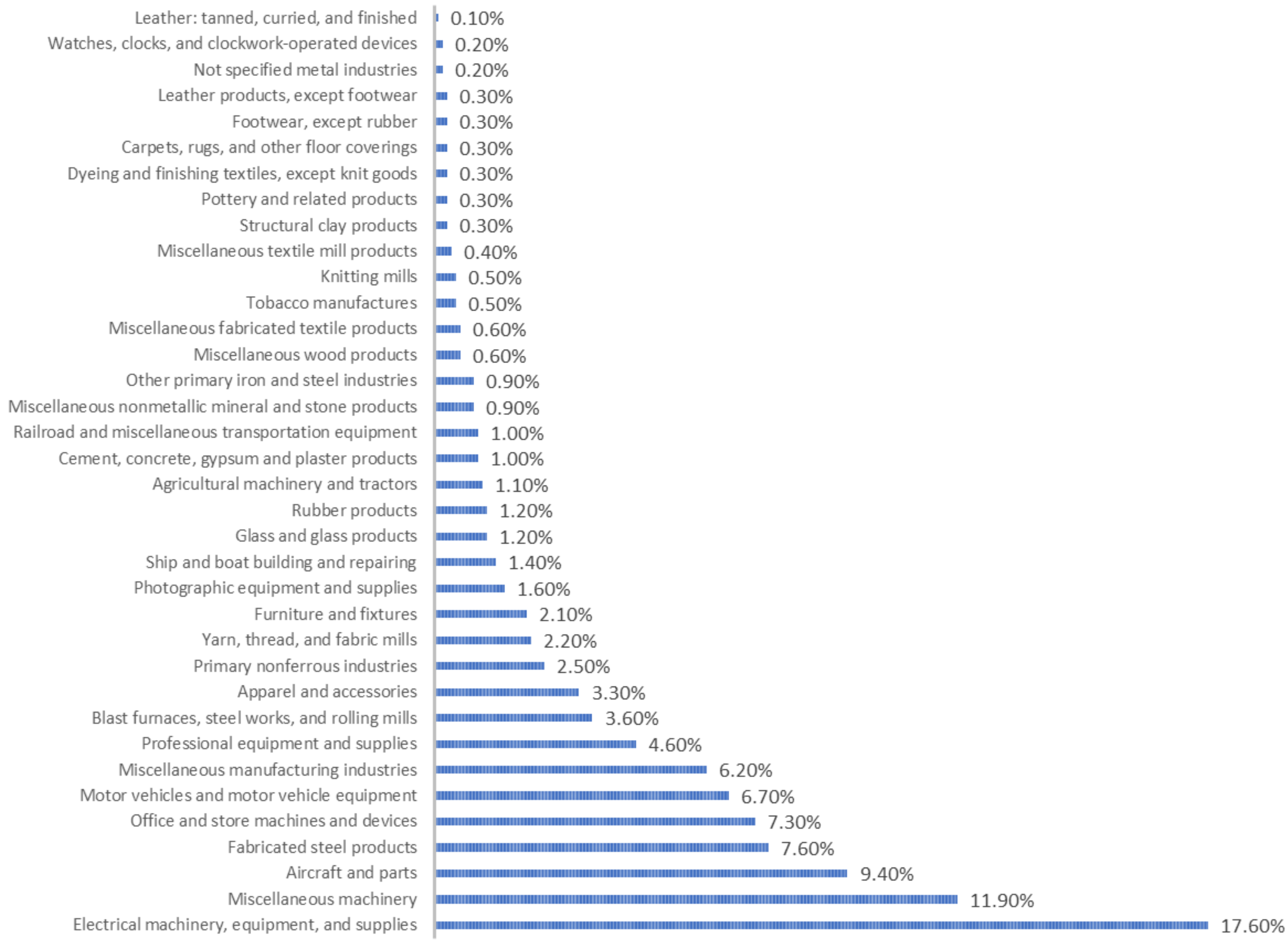

Figure 1.5: Share of Workers in Manufacturing Sector with Educational Attainment 4 Years of College, 1980

Note: The horizontal axis shows the percentage of workers with educational attainment of 4 years in 1980. The vertical axis shows the 36 manufacturing industries reported from 1950 Census Industry Code. *Data source: the 1980 5% Census obtained via IPUMS (Ruggles et al. 2019)*

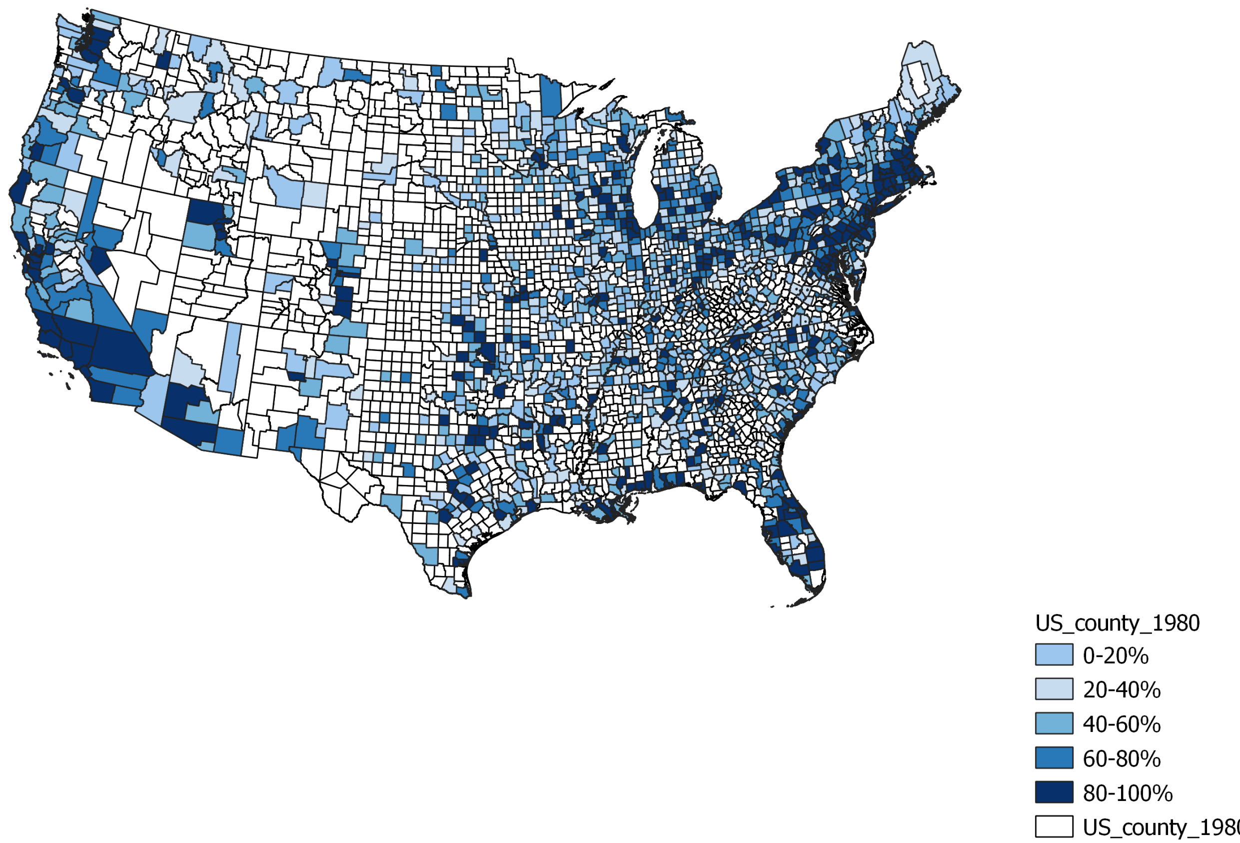

Figure 1.6: Geographical Distributions of Counties receiving DoD Procurement in Industries SIC 34-38, 1987

Note: The Department of Defense procurement in 1987, billion \$, adjusted for inflation using the base year 2000. Counties marked with color received procurement in industries SIC 34- 38. Color intensity measures the relative amount of contract obligations in counties. The five color categories indicate the quintile groups of county-level procurement amount in SIC 34-38.*Data source: Department of Defense 350 forms, and author's tabulation*

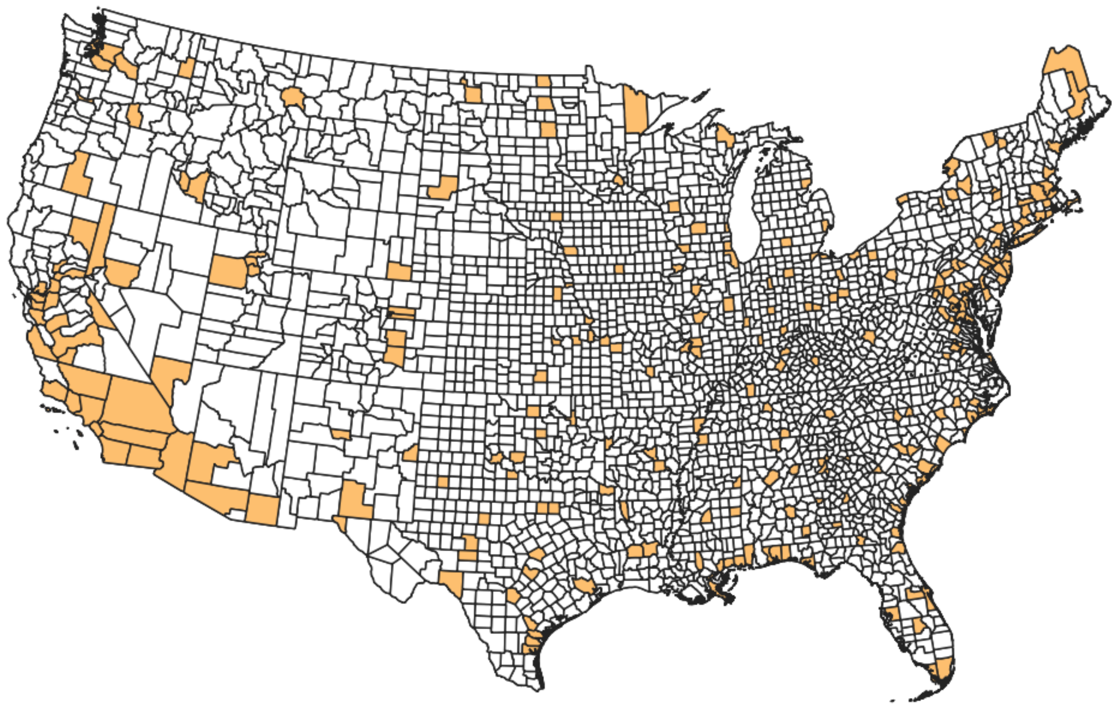

Figure 1.7: Geographical Distribution of Military Counties, 1988

Note: The graph shows the locations of counties (marked with color) with military installations in 1988. *Data source: Military Base Report Fiscal Year 1989*

Appendix

Additional Tables

Table 1.10: Estimated Effects of Procurement Drawdown on County's Non-farm Employment

Dependent variable: change in employment (total non-farm) to population (per 100 people)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
ΔProcurement in M5 sector,	0.209*	0.243**	0.229**	0.216**	0.212**	0.222**
$1,000 per person	(0.110)	(0.102)	(0.102)	(0.091)	(0.091)	(0.099)
Panel B: IV (Geographical distribution)						
ΔProcurement in M5 sector,	0.497	0.679**	0.676**	0.589**	0.627**	0.639**
$1,000 per person	(0.324)	(0.278)	(0.277)	(0.269)	(0.277)	(0.276)
First stage F statistics	191.895	192.211	195.717	193.654	197.586	194.067
Panel C: IV (Bartik)						
ΔProcurement in M5 sector,	0.155	0.459*	0.432*	0.411*	0.419*	0.438*
$1,000 per person	(0.290)	(0.247)	(0.243)	(0.240)	(0.240)	(0.241)
First stage F statistics	28.026	28.754	28.600	29.010	28.908	28.859
Demographic controls ×year FE		Y	Y	Y	Y	
Share of private non-farm employment in the broad sectors ×year FE			Y		Y	
Share of manufacturing employment in the most affected sectors ×year FE				Y	Y	
Demographic controls (trend)						Y
Share of private non-farm employment in the broad sectors (trend)						Y
Share of manufacturing employment in the most affected sectors (trend)						Y
Year FE	Y					Y
No. obs.	8510	8510	8510	8510	8510	8510

*** p<0.01, ** p<0.05, * p<0.1 County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the county level in parenthesis.

Table 1.11: Estimated Effects of Procurement Drawdown on Establishment Numbers

Dependent variable: change in establishments to population (per 100 people)

	(1) Total	(2) Flagship	(3) Manufacturing	(4) Retail
Panel A: OLS				
Procurement in MF sector,	0.000	0.000	0.000	0.000
$1,000 per person	(0.000)	(0.000)	(0.000)	(0.000)
Panel B: IV (Geographical distribution)				
Procurement in MF sector,	0.000	0.000	-0.000	0.000
$1,000 per person	(0.000)	(0.000)	(0.000)	(0.000)
First stage F statistics	197.586	197.586	197.586	197.586
Panel C: IV (Bartik)				
Procurement in MF sector,	0.000	0.000*	0.000	0.000
$1,000 per person	(0.000)	(0.000)	(0.000)	(0.000)
First stage F statistics	28.906	28.906	28.906	28.906
Demographic controls ×year FE	Y	Y	Y	Y
Share of private non-farm employment in the broad sectors ×year FE	Y	Y	Y	Y
Share of manufacturing employment in the most affected sectors ×year FE	Y	Y	Y	Y
No. obs.	8510	8510	8510	8510

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the county level in parenthesis.

Table 1.12: Main Outcome Estimates with Standard Errors Clustered at the State Level

	(1) Flagship Industries	(2) Manufacturing	(3) Retail	(4) Out-migration	(5) In-migration
Panel A: OLS					
ΔProcurement in M5 sector,	0.240***	0.222***	0.024	-0.013	0.091***
$1,000 per person	(0.046)	(0.047)	(0.025)	(0.029)	(0.024)
Panel C: IV (Geographical distribution)					
ΔProcurement in M5 sector,	0.461***	0.598***	0.118**	0.038	0.095
$1,000 per person	(0.109)	(0.114)	(0.056)	(0.038)	(0.060)
First stage: KP Wald F statistic	185.082	185.082	185.082	185.082	185.082
Panel C: IV (Bartik)					
\Delta Procurement in MF sector,	0.336***	0.356***	0.099**	-0.009	0.119***
$1,000 per person	(0.103)	(0.123)	(0.043)	(0.044)	(0.041)
First stage: KP Wald F statistic	28.384	28.384	28.384	28.384	28.384
Demographic controls (trend)	Y	Y	Y	Y	Y
Share of private non-farm employment in the broad sectors (trend)	Y	Y	Y	Y	Y
Share of manufacturing employment in the most affected sectors (trend)	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y

*** p<0.01, ** p<0.05, * p<0.1 County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the state level in parenthesis.

Table 1.13: Estimated Effects of State Procurement Drawdown on the State Counts of Disability Insurance Enrollees

Dependent variable: change in state DI enrollees to state population (per 100 people)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
ΔProcurement in MF sector,	-0.012	-0.016	-0.018	0.013	0.013	-0.013
$1,000 per person	(0.017)	(0.023)	(0.025)	(0.019)	(0.020)	(0.021)
Panel B: IV (Geographical distribution)						
ΔProcurement in MF sector,	0.006	-0.004	-0.009	0.018	0.003	0.027
$1,000 per person	(0.017)	(0.017)	(0.020)	(0.026)	(0.020)	(0.040)
First stage F statistics	59.497	35.297	26.183	30.462	21.950	48.994
Panel C: IV (Bartik)						
ΔProcurement in MF sector,	0.013	-0.006	-0.012	0.016	-0.004	0.049
$1,000 per person	(0.020)	(0.017)	(0.025)	(0.022)	(0.022)	(0.053)
First stage F statistics	35.014	23.489	18.889	25.577	18.215	25.682
Demographic controls ×year FE		Y	Y	Y	Y	
Share of private non-farm employment in the broad sectors ×year FE			Y		Y	
Share of manufacturing employment in the most affected sectors ×year FE				Y	Y	
Demographic controls (trend)						Y
Share of private non-farm employment in the broad sectors (trend)						Y
Share of manufacturing employment in the most affected sectors (trend)						Y
Year FE	Y					Y
No. obs.	250	250	250	250	250	250

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the state level in parenthesis.

Table 1.14: Estimates Effects of State Procurement Drawdown on the State Counts of Retirees

Dependent variable: change in retired people to state population (per 100 people)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
ΔProcurement in MF sector,	-0.020	-0.032	-0.023	-0.026	-0.026	-0.024
\$1,000 per person	(0.019)	(0.044)	(0.049)	(0.066)	(0.068)	(0.024)
Panel B: IV (Geographical distribution)						
ΔProcurement in MF sector,	0.034*	0.031	0.026	0.061	0.021	0.064
\$1,000 per person	(0.020)	(0.026)	(0.026)	(0.037)	(0.042)	(0.044)
First stage F statistics	59.497	35.297	26.183	30.462	21.950	48.994
Panel C: IV (Bartik)						
ΔProcurement in MF sector,	0.009	-0.017	-0.022	0.012	0.002	0.005
\$1,000 per person	(0.006)	(0.030)	(0.042)	(0.030)	(0.043)	(0.010)
First stage F statistics	35.014	23.489	18.889	25.577	18.215	25.682
Demographic controls ×year FE		Y	Y	Y	Y	
Share of private non-farm employment in the broad sectors ×year FE			Y		Y	
Share of manufacturing employment in the most affected sectors ×year FE				Y	Y	
Demographic controls (trend)						Y
Share of private non-farm employment in the broad sectors (trend)						Y
Share of manufacturing employment in the most affected sectors (trend)						Y
Year FE	Y					Y
No. obs.	250	250	250	250	250	250

*** p<0.01, ** p<0.05, * p<0.1 County-level demographic controls include birth rates, the percentage of population as male, the percentage of population as black, Asian American, the percentage of population having high school, some college, and college and above degree. County-level controls of industrial composition include the proportion of employment in the service sector, the FIRE sector, the retail trade sector, the wholesale trade sector, the manufacturing sector, the construction sector, the transportation sector and the agricultural sector. Standard errors clustered at the state level in parenthesis.

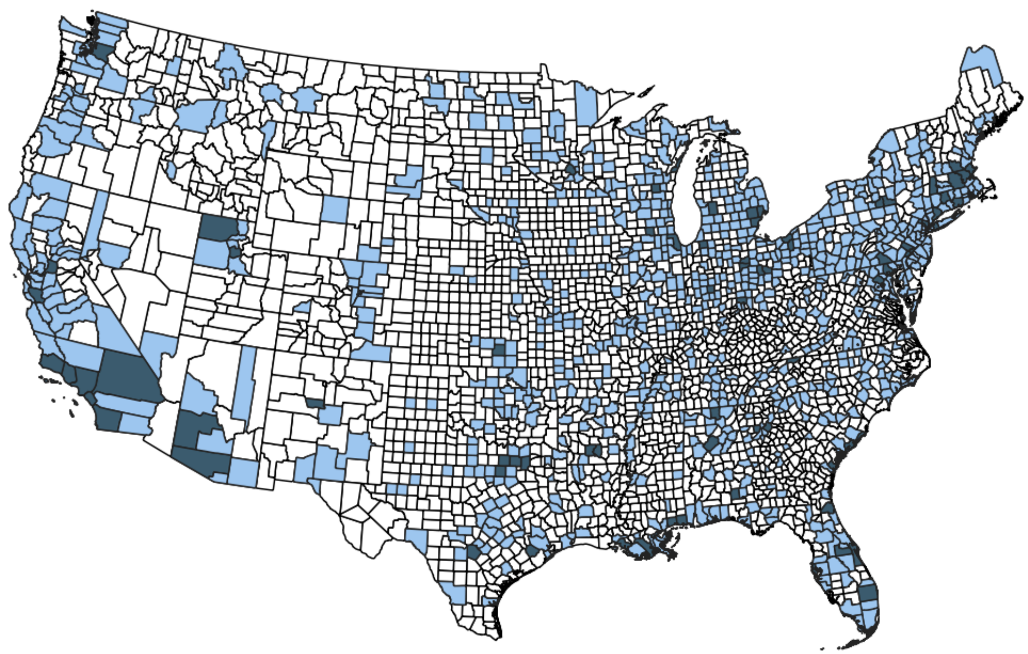

Figure 1.8: Geographical Distribution of Counties Receiving DoD Procurement in the Industry Transportation Equipment (SIC 37)

Note: The graph shows the locations of counties (marked with color) that received DoD procurement in industry SIC 37. Places with darker color are the counties with national shares of procurement in SIC 37 above 0.08% (top 10% of the distribution of SIC 37 procurement share in 1987). Places with light color are the remaining counties receiving SIC 37 procurement.*Data source: Department of Defense 350 forms, and author's tabulation*

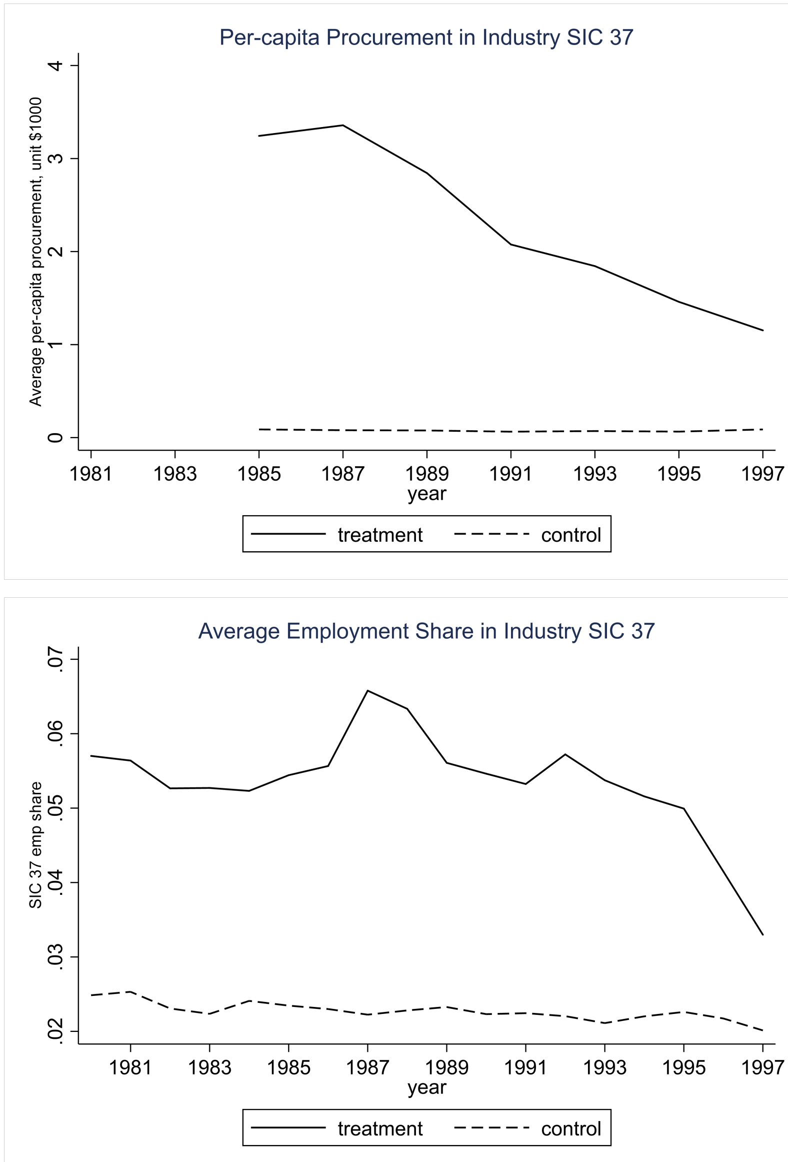

Figure 1.9: Average DoD Procurement Amount and Average Employment Share in Industry SIC 37

Note: The upper figure shows the average procurement in the industry Transportation Equipment (SIC 37). The unit is thousand $, adjusted for inflation using the base year 2000, normalized by the county population in 1980 Census. The lower figure plots the average employment share in the the industry Transportation Equipment (SIC 37). The "treatment" group includes counties with national shares of procurement in SIC 37 being above 0.08% (top 10% of the distribution of SIC 37 procurement share in 1987). The "control" group includes the remaining counties receiving SIC 37 procurement. *Data source: Department of Defense 350 forms, the QCEW, and author's tabulation.*

Illustration of the two Instrumental Variables

Without the loss of generality, I demonstrate the decomposition of the two instruments using a simplified example. Suppose we have two counties, {A, B} in two periods of time $\{t_1, t_2\}$. Each county has three industries $i \in \{1, 2, 3\}$, and I use $\{a_1, a_2, a_3\} \in A$; $\{b_1, b_2, b_3\} \in B$ to demonstrate the amount of contract in each place.

One can compute the predicted amount of contract change in county A in the spirits of two IVs. Our first IV predicts the local amount of change by allocating the same national change following the historical and geographical distribution. Specifically, for place A, the change of its total contract is:

$$
\begin{aligned}
\Delta A = & \frac{a_1^{t_1}}{a_1^{t_1} + b_1^{t_1}} \times ((a_1^{t_2} + b_1^{t_2}) - (a_1^{t_1} + b_1^{t_1})) + \\
& \frac{a_2^{t_1}}{a_2^{t_1} + b_2^{t_1}} \times ((a_2^{t_2} + b_2^{t_2}) - (a_2^{t_1} + b_2^{t_1})) + \\
& \frac{a_3^{t_1}}{a_3^{t_1} + b_3^{t_1}} \times ((a_3^{t_2} + b_3^{t_2}) - (a_3^{t_1} + b_3^{t_1})),
\end{aligned}
$$

where ΔA is the predicted total contract change of all its three industries $\{a_1, a_2, a_3\}$ in county A.

The second IV exploits the within-county industrial composition of the contract and predicts the total change based on each industry's national growth rate of the contract amount. For example, the national contract growth rate of each industry i between the period is:

$$
g\hat{r}(i) = \frac{(a_i^{t_2} + b_i^{t_2}) - (a_i^{t_1} + b_i^{t_1})}{a_i^{t_1} + b_i^{t_1}},
$$

where $i \in \{1,2,3\}$. For county A specifically, the predicted local growth rate is:

$$\%\Delta A = \frac{a_1^{t_1}}{a_1^{t_1}+a_2^{t_1}+a_3^{t_1}} \times g\hat{r(1)} + \frac{a_2^{t_1}}{a_1^{t_1}+a_2^{t_1}+a_3^{t_1}} \times g\hat{r(2)} + \frac{a_3^{t_1}}{a_1^{t_1}+a_2^{t_1}+a_3^{t_1}} \times g\hat{r(3)}.$$

With the rearrangement of the terms, the total dollar amount contract change in county A is:

$$\begin{aligned}
\Delta A =& (a_1^{t1}+a_2^{t1}+a_3^{t1}) \times \%\Delta A \\
=& \frac{a_1^{t_1}}{a_1^{t_1}+b_1^{t_1}} \times ((a_1^{t_2}+b_1^{t_2}) - (a_1^{t_1}+b_1^{t_1})) + \\
& \frac{a_2^{t_1}}{a_2^{t_1}+b_2^{t_1}} \times ((a_2^{t_2}+b_2^{t_2}) - (a_2^{t_1}+b_2^{t_1})) + \\
& \frac{a_3^{t_1}}{a_3^{t_1}+b_3^{t_1}} \times ((a_3^{t_2}+b_3^{t_2}) - (a_3^{t_1}+b_3^{t_1})),
\end{aligned}$$

In this simplified scenario with only two periods, the two IVs produce the same predictions because the initial share at the historical period t_1 is the same as the beginning-of-the-period amount to compute the national growth rate. However, the two IVs will be different when we fix the "shart" part of the IV to an earlier period, t_0.

Chapter 2

Lower Price but Higher Bill? Evidence from the Zero-Markup Policy in China

, with Jiafeng Wu (UVA)

2.1 Introduction

The financial burdens of high medical spending are a challenge many countries face. Among all factors driving high health care costs, increasing expenditure on prescription drugs plays a key role. In the United States, for instance, the health care system spending on prescription drugs amounted to 603 billion dollars in 2021, accounting for approximately 2.5% of the country's GDP.[1] The situation could be even more daunt-

[1] https://aspe.hhs.gov/sites/default/files/documents/88c547c976e915fc31fe2c6903ac0bc9/sdp-trends-prescription-drug-spending.pdf

ing for some other countries. In China, for example, total pharmaceutical expenditure reached 3.6% of the GDP in 2015.[2]

China's high medication costs have drawn significant attention from the media, researchers, and policymakers. Multiple studies (e.g., Li et al., 2012; Currie, Lin, and Meng, 2014; Chen et al., 2014) have found the operational practices of hospitals in China contributed to high retail prices. In China, public hospitals provide health care services and have a prominent role in dispensing medications. They purchased medicines at a fixed price set at the provincial level and were permitted to sell these prescriptions to patients at a higher price, with a maximum markup cap of 15%. This profit-generating tool has been believed to create strong financial incentives for hospitals to prioritize the sale of drugs that yield higher profits. Physicians, as direct employees of hospitals, play an essential role in this process since part of their total income could be tied to bonuses, which might depend on the revenue that they brought in, including prescriptions (Yip and Hsiao 2008; Yip, Hsiao, Meng, et al. 2010). Furthermore, despite the presence of local pharmacies, public hospitals remained significant providers of medications to the general public. In fact, public hospitals in China relied heavily on drug sales. Prior to the health care reform that we study, the sale of medications accounted for over 40% of a hospital's outpatient and inpatient revenue on average (Fu, Li, and Yip, 2018).

In realization of the concerning incentives for physicians, the Chinese government launched a national reform, commonly known as the "Zero-Markup Policy" (ZMP), in 2009. The primary objective was to alleviate patients' financial burdens due to high retail prices and potential over-prescriptions. The policy targeted nearly all medicines sold at public hospitals and mandated that they be sold at their procurement costs.[3]

[2] `https://www.healthaffairs.org/doi/10.1377/hlthaff.2018.05324`

[3] This meant that hospitals could no longer add a markup to the medications they dispensed to

The implementation of the policy began with trial programs at primary clinics in local townships and gradually expanded to include all city hospitals. By 2017, the ZMP had been fully implemented in all public hospitals across China.

While it may seem appealing to assume that removing medication markups would alleviate patients' financial burdens, a comprehensive evaluation of the overall effectiveness of such reforms requires careful empirical analysis. It is because physicians' behaviors could be affected by the two opposing forces brought by a lower medication price.[4] On one hand, the elimination of profits from medication sales may reduce the incentives for over-prescribing. Physicians may opt for a cheaper alternative, a lower dosage, or a combination of both, resulting in reduced costs. On the other hand, the loss of profits due to price controls may incentivize physicians to intensify non-medication treatments, particularly if they have a greater say in the diagnosis process. These incentives may be further amplified when other services have higher profit margins. Previous studies have documented the financial incentives and behavior of health care providers in various contexts (e.g., Yip, 1998; Gruber, Kim, and Mayzlin, 1999; Dafny, 2005; Ho and Pakes, 2014; Fang and Gong, 2017; Alexander, 2020). These studies have shown that changes in reimbursement rates, whether higher or lower, can lead physicians to prescribe more services to compensate for revenue loss or extract more profits.

In this paper, we exploit a unique individual-level administrative database from

patients, with the exception of a specific category of liquid Chinese herbal medicine.

[4] Double marginalization could be another concern in the long run. For example, suppose a model environment with only one monopoly upstream and one monopoly downstream firm. If the downstream firm suddenly cannot charge any markup and becomes a pure distributor, then the upstream monopoly firm would charge a lower price to maximize its profit. However, it is less of a concern for our context. The main reason is that the procurement cycle overlaps with our sample period, in which the medicine procurement cost was already fixed and would not change in the post-ZMP in our sample period. In addition, we do not observe a visually noticeable change in the average medicine expenditure from our data from Oct 2016 to the official ZMP implementation date.

the Health Care Security Administration (HSA) of a representative city in China. The city has a population of around 1 million urban residents and a per-capita income level slightly higher than the average prefecture city in China, making it a suitable representative sample of a Chinese prefecture city.[5] We obtained the full records of inpatient visits with urban employer-sponsored insurance.[6] It allows us to examine the specifics of a person's hospital visits, including the reasons for visits and the associated costs for a detailed set of diagnoses and treatments received. By leveraging this rich dataset and the ZMP implementation in 2017 in this location, we provide insights into the specific patterns of health care utilization and costs among the urban population in this representative city.

We employ a conditional Difference-in-Differences (DiD) approach to investigate how physicians responded to the removal of medication markups under the ZMP. Specifically, we focus on estimating the impacts of the ZMP by comparing changes in medicine expenditure, medical service expenditure, and service utilization rates between the post-ZMP calendar months in 2017 and the same calendar months in the previous year, 2016, relative to the differences in the earlier months. Our findings indicate a significant decrease of 18% to 20% in the average spending on medication in the two general-purpose hospitals following the implementation of the ZMP. However, this decrease was completely offset by a corresponding increase in non-drug services, resulting in no change in the average patient's medical bill. This suggests that physicians compensated for the lost revenue from medication sales by increasing the provision of non-drug medical services. Furthermore, we examine the changes in the utilization of specific medical services, such as exams, surgeries, or the dura-

[5] An administrative division in China is ranked as Province - Prefecture City - County.

[6] This group includes all urban residents who are employed. We include more details in section 2.2.

tion of inpatient care. We find no strong evidence that physicians directed patients towards increased service utilization. However, we do observe that physicians prescribed more expensive supplementary materials or personal care items, leading to a significant increase in average patient spending on consumables in their treatment process or personal care. Importantly, as the official government documents indicate, this increase in consumable expenditure is unrelated to any procurement price changes in the comprehensive hospitals. Moreover, the expense increase for specific disease groups was substantial enough to offset the benefits of reduced medication prices.

We also document a heterogeneous response across hospital types. In particular, the integrated hospital, the only hospital in the city that offers traditional Chinese treatment, followed a different strategy to compensate for the lost revenue from medication sales. While we also observed a significant decline in medicine revenue of 32%, there were no substantial rises in consumable expenses. Instead, the physicians in this hospital attempted to prescribe more supplementary herbal therapies or traditional physical treatments.

The high health care expenditures have been extensively studied in the literature. Demographic changes, such as population aging and the prevalence of chronic diseases, along with income growth, contribute to the rising demand for health care services (e.g., Newhouse, Group, and Staff, 1993; Hall and Jones, 2007). While this increased demand drives some health care costs, institutional factors within the health care system also play a significant role. The principal-agent relationship between physicians and patients, which involves information asymmetry and uncertainty, explains why physicians may provide excessive treatments, leading to higher costs (Arrow, 1965). Empirical evidence has demonstrated physicians' responses to financial incentives through various research designs (Currie, Lin, and Zhang, 2011; Currie, Lin, and Meng, 2014; Lu, 2014; Alexander, 2020) or through quasi-experiment settings (Gruber and Owings, 1994; Yip, 1998; Dafny, 2005; Fang and Gong, 2017). Recent discussions have also focused on the effects of broader access to health care services on increased overall utilization and higher prices (e.g., Finkelstein, Taubman, Wright, et al., 2012; Taubman et al., 2014; Finkelstein, Taubman, Allen, et al., 2016). These studies collectively provide valuable insights into the drivers of health care expenditures and the role of financial incentives in shaping physicians' behavior.

Our research makes several contributions to the literature on the response of medical providers to financial incentives. By exploring a unique policy experiment and utilizing a novel and comprehensive individual-level insurance dataset, we provide empirical evidence on how physicians adjust their behaviors when the government eliminates profit margins on drug prescriptions. Our findings support the notion of a principal-agent relationship between physicians and patients, where financial incentives can lead to excessive treatment. Moreover, our results uncover a relatively implicit channel through which physicians compensate for their profit loss. We also find suggestive evidence of a shift in financial burdens across different diagnosis groups, with patients suffering from specific diseases bearing higher costs. It has implications for understanding the welfare consequences for different patient groups.

Furthermore, our paper contributes to the strand of literature studying health care reforms in China and other similar contexts with a top-down policy design. Over the past few decades, China has implemented a series of health care reforms to reduce medical burdens for individuals. Among these reforms, particular attention has been given to the drug sector due to its significant contribution to total health care expenditure.[7] However, formal evaluations of these policies are limited, probably because of the data challenges. The few existing studies (e.g., Xiang, 2021; Fang, Lei, et al., 2021) on the impacts of the Zero-Markup Policy (ZMP) have focused on township or county hospitals, which may have limited capacity for more complex treatments such as surgeries or specific diseases. In contrast, our analysis utilizes data from comprehensive city hospitals, which cover a wide range of medical conditions and allow us to examine physicians' responses to different types of medical services.

The rest of the paper is organized as follows. Section 2.2 introduces the institutional background of China's health care system and the medical reform. Section 2.3 discusses data and summarizes our sample. Section 2.4 outlines the framework of our research design. The presentation and interpretation of results are in Section 2.5. Section 2.6 concludes.

2.2 Institutional Background

2.2.1 Public Hospitals in China

Public hospitals in China are the primary providers of healthcare services, delivering more than 90% of the country's inpatient and outpatient services (Yip, Hsiao, Chen, et al., 2012). In addition to providing diagnosis and treatment, they have played a critical role in distributing medicines, representing an average market share of 80% of all retail drug sales (National Medical Products Administration, 2014).

The dominant market coverage of public hospitals in medicine sales is related to

[7]For example, to ensure the need for certain essential medicines are satisfied due to disease prevalence, China has been updating its reimbursement drug list and the national essential drug list. The ZMP targets reducing drug prices and hospitals' reliance on drug-selling revenue.

how the healthcare insurance system operates in China. The Chinese central government launched a universal health insurance program in 2007 to provide comprehensive coverage for all residents. By 2011, the public health insurance system had covered more than 95% of the population.[8] Each local government (usually the provincial administration) designs and provides guidance, including details on individual copayment and reimbursement rates for each type of healthcare service. Participants who locate in the same city and are registered under the same insurance group share the same insurance plan and thus have the same policy reimbursement rates.

Like many other countries, an individual must show an official prescription issued at hospitals to purchase the corresponding prescription medicines. However, to receive insurance reimbursement for medications, individuals need to buy from places authorized in the government's health insurance network, which in most cases are public hospitals. Retail pharmacies, another option for customers to obtain medications, usually are not qualified for reimbursement and are preferred chiefly for convenience reasons.

2.2.2 Hospitals' and Physicians' Incentives

Although the name suggests government ownership, the government does not fully fund public hospitals in China. In fact, on average, government funding only accounted for less than 10% of a hospital's total revenue (China Health Statistic Yearbook), rendering medicine sales and other services their primary income sources. Like health insurance plans, the government actively regulates the prices of medication and

[8]The plan designated the residents into three main classifications: Urban Employer Sponsored, Urban Non-Employer Sponsored, and Rural Group. The two urban groups cover only city residents, with the first group consisting of people whose employers sponsor the insurance through the Social Security Administration. The HSA data covers all hospital visits of people in this group.

services. Specifically, public hospitals were allowed to charge up to 15% markup over the drug procurement price when selling drugs to patients. Besides, the government sets a price ceiling for each type of healthcare service. It is commonly known that labor-related services have been heavily under-priced, which creates strong incentives for hospitals to rely on its medicine sales.[9] In 2011, for example, the national pharmaceutical revenue, on average, accounted for about 40% of total health expenditure. The average ratio of medicine to the total medical spending of the three comprehensive hospitals in this study ranged from 27% to 35% before the medical reform. As employees of hospitals, physicians' incentives were also aligned with the hospital's in prescribing more profitable medicines since they could be rewarded with bonus payments and promotions based on the revenue they generated (Yip, Hsiao, Meng, et al. 2010).

2.2.3 Healthcare System Reform

To cut the linkages between physicians and medicines, the central government launched a nationwide healthcare system reform, commonly known as the Zero Markup Policy (ZMP), in 2009. The reform prohibited any markup profits made by public hospitals for dispensing medications. As a result, the retail prices of all medicines sold at public hospitals must be set just their procurement cost, a cost fixed at the province level through a centralized procurement process. Nevertheless, the policy allowed public hospitals to adjust the medical service prices subject to a regulated price cap for each category.

[9] According to a 2009 statement from the National Health Commission of China, one of the healthcare system issues is that prices of some medical services had been lower than their costs for a long time. For example, in the public hospitals of our city, the consultation fee could be as low as 2 - 3 Yuan ($0.3 - $0.46) each time.

In the same spirit as the Chinese regional experiment regime, the policy was initially piloted at a smaller scale and later rolled out sequentially to cover a broader population base. The first phase (between 2009 and 2012) launched a pilot program targeting primary healthcare institutions (e.g., township clinics). In the second phase (2012 to 2015), all county-level hospitals were required to implement the policy. The program was then rolled out to different cities across the county so that by the end of September 2017, all city-level hospitals had removed their markup from medicine sales.

Complying with this planned national schedule, the provincial governments then determined the exact timing of policy implementation for each city under its administration. For instance, the ZMP implementation date in our sample city was July 31, 2017. Following the reform, all public hospitals in the city must sell medications at the procurement price determined in the provincial centralized drug procurement system. [10] Apart from the medicine price change under the ZMP, the city government issued price adjustment guidance for all medical services. The advice updated a list of cap prices that each public hospital can charge for each type of service, including inpatient and post-surgery care, examination, surgical treatment, and all other supplementary treatment. Hospitals are allowed to set prices of medical services as long as they do not exceed the cap prices set by the government. We refer to the official document published by the city government for the ZMP in 2017 when aggregating the granular service items. Generally, the guidance allowed an upward price adjust-

[10]To lower prices, each provincial government established a central bidding platform to which pharmaceutical companies bid for the wholesale price for each designated drug category. Once the winning bids are finalized, the local (provincial) centralized procurement requires all public hospitals to purchase medicines from the winning producers at the bid price. The centralized procurement process was launched nationally in 2009 and renewed every two years. The most recent medicine procurement in our sample province happened in December 2015, and the new wholesale price was issued in Oct 2016. We do not observe a visually noticeable change in the average medicine expenditure from our data from Oct 2016 to the official ZMP implementation date.

ment for the diagnosis, inpatient care services, and surgery treatment and required a downward price adjustment for exams.[11] Unlike these service groups, there were no price changes in the medical consumables during our sample period.[12]

2.3 Data and Sample Description

Our primary data is from the Healthcare Security Administration (HSA) of a prefecture city in a central Chinese province. The city is generally representative of a median Chinese prefecture city – it has a population of around 1 million urban residents and an annual GDP per capita of roughly USD 8,000, both of which are slightly higher than a median and average prefecture city in China, according to the 2010 Chinese Population Census.

We use the daily-individual-level patient healthcare claims data from HSA, which covers the city's urban residents' healthcare expenditure. Patients visiting public hospitals are identified using their government-sponsored medical insurance IDs and are digitally recorded in the healthcare system. The broad coverage of such government insurance means that almost all urban residents are included in the design, minimizing a sample selection risk commonly seen in survey data.[13] We conduct the analysis

[11] Each broader category of medical services in the document contains tens or hundreds of service items, and each may vary by unit. For example, during the entire hospitalization experience, there could be multiple times of nursing services provided, whereas there is generally at most one surgery conducted. The document also exhibits a variety of price cap changes across service types. The adjustments range from -6% to -15% for medical examination services and are from 20% to 30% for surgical treatments.

[12] Medical consumables may include general medical equipment such as syringes, needles, tubing, sealants for wounding, etc., or high-value medical supplies such as vascular catheters or artificial joints typically used in surgeries. The medical consumables were also procured through a centralized procurement system.

[13] China's government medical insurance covered more than 95% of the population as of 2013. https://web.archive.org/web/20150328095843/http://www.mckinsey.com/insights/health_systems_and_services/health_care_in_china_entering_uncharted_waters

using the city's three public comprehensive hospitals. These hospitals provide both inpatient and outpatient medical services and have licensed health professionals to offer consultative, diagnostic, and therapeutic services to almost all types of disease categories.[14] Specifically, three city-level comprehensive hospitals form the basis of our sample. Among them are the general types of comprehensive hospitals: The City No.1 People's Hospital and The City No.2 People's Hospital. We refer to them as Hospitals A and B in the following sections. There is also an integrated hospital that provides traditional Chinese physical and herbal treatment as an additional option, with the official name being the City's Hospital of Traditional Chinese Medicine. We denote it as Hospital C in the following sections.

The HSA data provides rich expenditure information on each patient's visits to these comprehensive hospitals. For each patient visit, we observe the visit date, the person's characteristics such as age, gender, if the visit is for an inpatient service, and the duration of stay in the hospital. We focus on inpatient records for the completeness of diagnosis information.[15] Pairing with each individual-daily level visit record, we observe a rich vector of prescription spending and service expenditures covering from consultation to therapies. Together, we can examine the behavioral responses of physicians in prescribing non-medicine services when there is a drop in the profit margin of medical prescriptions.

[14]The comprehensive hospitals are different from specialty hospitals in that the latter only offer medical services to a particular disease group. For example, the city has nine public hospitals, out of which three are city-level comprehensive hospitals, four are specialty hospitals (A Stomatology, a Maternity, a Dermatology, and a Rehabilitation hospital), and two district hospitals for clinics.

[15]Since inpatient services are eligible for insurance reimbursement, the disease name is well-recorded by physicians. In contrast, many outpatient services may not qualify for insurance coverage and are only vaguely reported as "general" in the claim item "reason of visit." Therefore, we observe massive missing values of the disease types for outpatient observations.

2.3.1 Disease Classification

Before moving to investigate the summary statistics of our sample, we first aggregate the granular disease names up to a consistent broader category based on the tenth revision of the International Classification of Diseases (ICD-10) from the National Clinical Disease Classification Code.[16] Grouping granular diseases into broader categories allows us to control for the common factors of diseases that lead to systematically higher (or lower) expenditures. Without grouping, the system contains thousands of unique disease names, some of which could have been treated as different diseases simply because one includes a few more words of description. By recognizing that certain groups of diseases are fundamentally related in an independent classification system, we can use fixed effects to control for all disease-related time-invariant factors affecting the outcome levels. We describe in Appendix the algorithm we used to match the diseases recorded in the local health administrative system to those from the official ICD-10.

2.3.2 Summary Statistics

Patients in our city sample have three choices when deciding among the comprehensive facility list. While all three hospitals are general-purpose and are classified by the provincial government as the top tier group[17], each hospital differs in the number and the types of patients. For example, a hospital with an integrated traditional Chinese

[16]China has adopted the ICD-10 standard since 2003. We obtained the ICD-10 mapping file from the National Clinical Disease Classification Code 2.0 published by the National Health Commission of the People's Republic of China in 2019

[17]Each public hospital is evaluated by the government with a grade based on their service and facility quality. The grading scale includes three groups: primary (bottom-tier), secondary (middle-tier), and tertiary (top-tier). Within each group, there are letter grades from A to C, with A being the best subdivision. The three hospitals in our sample are classified as a tertiary A group, indicating that they can serve more than 500 beds and are considered the best city-level general hospitals.

treatment (Hospital C) is believed to attract patients who at least value parts of the conventional treatment. Probably due to a longer establishment history, Hospital A received more inpatient visits than the other two. The HSA data allows us to see all inpatient visits to these three hospitals on a daily-person level from 2016-01-01 to 2017-12-31. We focus on relatively common disease types that receive at least 300 visits annually to avoid measurement biases related to a small set of uncommon diseases.

Tables 2.1, 2.2, and 2.3 show the annual distribution of diseases diagnosed by each hospital. Among all the inpatient services provided by the three hospitals during the two sample years, the most common disease groups are “Circulatory system” or “Neoplasms” -related. We see from Table 2.1 that Hospital A received more than double the number of patients than Hospital B, as Hospital A is the largest hospital in the city. Regarding disease distribution, more than a quarter of the inpatient services were diagnosed as related to “Circulatory system” for both Hospital A and B. While “Neoplasm” diseases were also a common visit reason in Hospital A, we can see that the proportion of Tumor-related diseases is much higher in Hospital B than Hospital A.

Table 2.4 provides further details of patients’ demographics in each hospital. We observe a relatively stable distribution of disease groups across hospitals. We divide our sample into four age groups. The sample similarity in terms of the diagnosis types and demographic distribution indicates no evidence of a changing patient composition across the years within the same hospital.

Figures 2.1, 2.3, and 2.5 demonstrate the basic patterns of the average medicine expenditure per patient in the two years. The average medicine expenditure was comparable across these two years from January through July. However, a sharp contrasting pattern has followed since August, when the ZMP took effect. After August 2017, the average medicine expenditure per patient experienced a noticeable visual decline relative to the 2016 levels in all three hospitals. The trends of a patient's total medical spending are shown in Figures 2.2, 2.4, and 2.6. In contrast to a declining trend of medicine expenditure, there is not much disparity in the total medical bill an average person received between the two years. The decreasing medicine price and a similar medical bill lead us to formally examine the channels through which the expenditures are altered.

2.4 Empirical Strategy

To study the impacts of the ZMP on an individual's medical spending, we leverage a conditional Difference-in-Differences (DID) strategy that compares a patient's bill between the post- and pre-ZMP period in the same hospital. Our empirical strategy is partially illustrated in Figures 2.1 - 2.6. Since all city hospitals implemented the ZMP simultaneously, we cannot rely on a conventional approach in which we compare the outcomes of the treated group to those of the control group unaffected by the hospital.[18]

In light of these empirical complications, we compute the counterfactual out-

[18]One tempting option is to use the neighboring county hospitals as a control group. However, the complications introduced in the identification strategy made us decide not to pursue such a strategy. Hospitals in the nearby counties adopted the reform earlier than the city hospitals, consistent with a typical reform in China that was widely implemented only after seeing successful examples from smaller-scale experiments. The experimental reforms took place in areas that are perceived to be less risky (often places with smaller populations) – which, in the context of ZMP, are the county hospitals. Therefore, the control group candidates are the hospitals that have already implemented the ZMP. More importantly, each county has only one comprehensive hospital, and the serving size is much smaller than a city hospital. For example, we find that few treatments involve an inpatient service. Moreover, patients who expected more complicated medical treatment for Tumor- and Heart-related diseases are more likely to visit the city's tertiary hospitals, further limiting the sample.

comes following Miller, Segal, and Spencer (2022). We use the most comparable outcomes on a daily basis from the year just before the ZMP to construct the outcomes that would have been observed for each hospital had the policy not occurred.[19] Specifically, our general DID strategy takes the following form:

$$Exp_{iwt} = \beta_1 ZMP_{iwt} + m_t + dow_t + wave_t + \delta X_{iwt} + \epsilon_{iwt}, \tag{2.1}$$

where Exp_{iwt} is an individual patient's expense on a particular medical service or medication, measured at the individual-wave-day level. The two waves are the years 2016 and 2017, respectively. The primary explanatory variable is ZMP_{iwt} for which an individual observed is assigned with value 0 if the patient is admitted before 2017-07-31 and is of 1 afterward. $wave_t$ represents the wave fixed effect and control for systematic time-invariant factors at the year level. Differences in expenditures could also occur in the timing of hospital visits. For example, people are expected to be less likely to visit hospitals during the lunar new year holidays unless necessary, so the average medical during the new year is typically different from the rest of the year. The fluctuations in hospital visits could also be observed for days within a week. To account for seasonal and weekday variation, we control for month-fixed effects m_t and day-of-week fixed effects, dow_t. Moreover, we also include a vector of individuals' observed characteristics such as gender, age, and a categorical variable covering the diagnosis groups. We estimate Eqn. 2.1 and separately for each comprehensive hospital.

In addition to expense changes, we also examine the effects of ZMP on medical

[19]While we could observe partial visit records before 2016, we chose the sample period from the first month of 2016. The city hospitals did not adopt complete digitization until late 2015, and there exists a strong selection bias of the observations in earlier periods. Moreover, the procurement costs of medical consumables were fixed during the sample period from 2016 Jan to 2017 Dec.

service utilization. As described in Section 2, a ZMP policy is linked to a corresponding price adjustment in other service categories.[20] To exclude the possibility that the total expenditure change of service categories merely reflects price change, we estimate a second model that studies the effects on service utilization rate. The specification is:

$$\mathbb{1}(\text{Service Utilization})_{iwt} = \beta_1 ZMP_{iwt} + m_t + dow_t + wave_t + \delta X_{iwt} + \epsilon_{iwt}. \quad (2.2)$$

The second model is shown in Equation 2.2. Here, the dependent variable is a binary indicator that takes a value of 1 if an individual was observed taking an exam or surgery during the diagnosis process. We impute the service utilization through a positive service charge for any exam or surgery-related categories. Therefore, a positive value of β_1 means that an individual is more likely to be charged with these services between the months after ZMP and December compared to a similar individual who visited the hospital between these months when ZMP was in effect.

The identification of the causal impacts of ZMP hinges on several assumptions. First, we assume that the ZMP was exogenously determined so that no changes within the hospital were associated with the reform's timing and medical service provisions. The central government initiated the ZMP reform, and the provincial government then planned and carried out county and city rollouts. Therefore, local bureaucrats (the city government in this paper) were not allowed discretion. Second, we also assume that the counterfactual average medical service utilization would have been the same as in the past year without any reform. Since all patients in the city were affected by the policy at a clear-cut time, we rely on the prior year to serve as a control

[20]For example, there is a systematic price decrease of 15% in MRI and CT exam fees, a 6% reduction of blood test fees, and a 20 or 30 % increase of surgery fees.

wave, and we assume that the calendar days of admission before (after) Jul-31 are the typical "pre" ("post") for each wave. Third, we assume that the broad diagnosis groups do not change in response to the treatment.

While one cannot verify such assumptions, we formally compare the differences in the outcomes of interests during the periods before the ZMP in the two years through a visual examination and an event-study framework. The event-study analysis is of the following form:

$$Y_{iwt} = \alpha + \sum_{\tau \neq 7} \beta_\tau \times \mathbb{1}(ZMP_i) \times \mathbb{1}(t = \tau) + m_t + dow_t + wave_t + \delta X_{iwt} + \epsilon_{iwt}, \quad (2.3)$$

where $1(t = \tau)$ is a group of indicator variables for each calendar month from January to December, except for July. It is the month since the ZMP started to take effect and is the baseline group we compare our coefficients. The framework incorporates a set of monthly indicators interacting with the ZMP year. Thus, these coefficients β_1 to β_6 capture the differences in outcome variables between the year 2016 and 2017 for months ((January - June) when ZMP was not in effect, and the coefficients β_8 to β_{12} measure the outcome differences relative to July for the months since ZMP ((August - December). Our outcome variables include expenses, utilization of service categories, and the length of stay at the hospital. The rest of the variables are the same as those in Equations 2.1 and 2.2. We show in section 2.5 that the control (wave 2016) and treatment (2017) groups exhibit parallel pre-trends on the outcome variables.

2.5 Results

2.5.1 Expenditure

Tables 2.7 and 2.8 present the main findings on aggregate medical spending for two comprehensive hospitals, Hospital A and Hospital B. The outcome variables are presented as the average per-person medical bills in local currency (Chinese Yuan, CNY), and the coefficients in the table indicate the relative change in spending for medicine, non-medicine, and total expenditure. Column 1 shows that following the ZMP enforcement, there were noticeable decreases in medicine expenditure, with an average reduction ranging from ¥599 (17.8%) to ¥669 (19.9%) for the two general-purpose hospitals, respectively. Although we cannot observe the price and quantity changes separately, the significant decline in medicine prescriptions suggests that the price reduction percentage outweighed demand responses, leading to an overall decrease in drug expense. Furthermore, the large average percentage drop indicates that over-prescription for drugs could be a problem in these city hospitals before the ZMP. However, there was a large and significant increase in the sum of non-medicine categories, as shown in Column 2. While the absolute amounts of these increases were not as large as the medicine expense drop, the changes were still statistically significant. As a result, we only observed a slight, statistically insignificant decrease in the final bill amount (Column 3). The ZMP changed the relative reliance of hospitals' revenue on drug sales.

2.5.2 Service Utilization

Since our results show a large adjustment of non-medicine expenditure in both hospitals, it is worth examining whether the increased bill was due to more physician-induced service provision. We categorize the granular diagnosis and treatment services into the following main categories: (1) inpatient and post-surgery care, (2) Exam, (3) surgical treatment, and (4) medical consumables. For example, consultation services included in the first category are mandatory for every patient. Additionally, the usages of exams or medical consumables, although not necessarily 100%, are almost seen for all patients in this inpatient sample. Therefore, the utilization was already extensive for these groups, and any changes could only occur at the intensive margins. In contrast, we can exploit the differences in service utilization in the other categories to examine the effects of the ZMP on physician-induced demand.

We present the expenditure changes in Tables 2.10 and 2.11 for Hospitals A and B. There were significant changes in the average expenditure for all categories. Take Hospital A as an example, relative to the pre-ZMP months, there was an average increase of ¥350 in the inpatient and post-surgery care expenditure (Column 1), an average drop of ¥310 in exam expenditure (Column 2), and a small increase in surgery expenditure of ¥95 (Column 3), respectively. Moreover, medical consumables increased by ¥332, or more than 30% of the pre-ZMP level.

While there was a rise in expenses in specific categories, one cannot conclude if this was due to a price adjustment or a change in service utilization. Recall from section 2 that hospitals were *required* to reduce exam fees (if the former ceiling was binding) whereas were *allowed* to increase prices charged for medical services subject to a cap. Does the increase in bills associated with inpatient and post-surgery care mean that physicians in our sample hospitals keep patients longer at the hospital, or was it purely a reflection of price changes? To answer this question, we need to look at the utilization rates of the service, as mentioned earlier.

Tables 2.13 and 2.14 present the coefficient estimates for Equation 2.2 for inpatient and post-surgery care and surgical treatment for the two general-purpose hospitals, A and B, respectively. Regarding the inpatient and post-surgery care services, we use the admission and discharge dates to calculate the length of hospital stay and refer to it as the duration of inpatient care. None of these services experienced a significant increase in utilization rate after the ZMP. If anything, we observed a slight decrease of 0.5 days in Hospital A's average length of stay post the ZMP. Our estimates differ from the few studies that examined clinics at a more primary level. For example, Fang, Lei, et al. (2021) found that physicians in township health centers increased their charges on bed and exam fees.[21]

In Column 1 of Table 2.13, we find an increase in the average surgery rates for Hospital A. We show in the later subsection that the surges in surgery rates in Hospital A were primarily driven by a specific group of diseases in "Musculoskeletal system and connective tissue" and "Endocrine and metabolic diseases" (E). While the evidence may be suggestive given the relatively small sample size, the result is consistent with Xiang (2021) in which the patients with Spondylosis were seen with a higher surgery rate.

Our findings from the two hospitals suggest that physicians in these two city hospitals tried to maximize surplus to compensate for their earnings losses from lower

[21] However, we do not know if there was a price change in the township clinics for the bed or exam fees. If those clinics were allowed to raise fees, the conclusion of a shift towards more nights in the hospital or more exams should have been reflected in a change in the usage rates rather than the fees.

medicine profits. However, they did not necessarily have the power to induce more service utilization, such as extended hospital stays, exams, or surgeries. Surgical treatment often requires discretion from both sides, and higher prices may lead to reduced demand. Moreover, since the exam rates for inpatient services in our sample were already sufficiently high (as shown in Table 2.6), the even lower prices for the already under-priced exam fees leave no incentives for physicians to direct toward them. On the other hand, the physicians have much more discretion in charging medical consumables, which were procured through a centralized system at a fixed cost during our sample period. Specifically, the estimates in Columns 4 of Tables 2.10 and 2.11 show that the average spending increase of medical consumables was large enough to significantly weakened the markup reduction of the retail drugs. Consequently, an average patient's medical bill was barely lowered. Unlike exams or surgeries, since there was no cost adjustment in the medical consumables, all expenditure changes were caused by an increased prescription of the corresponding items or a substitution towards more expensive materials.

2.5.3 Further Evidence on Supplementary Treatment

Our study of two general-purpose comprehensive hospitals has revealed that physicians in the sample city tend to compensate for revenue loss by charging more for medical consumables, the item over which they have complete discretion. To gain further insights into how physicians could leverage the additional treatment options, we analyze a third comprehensive hospital (Hospital C) – the only hospital in the city that provides traditional Chinese treatment options. Our analysis concludes a second finding, presented in Tables 2.12 and 2.15, that depending on their hospital type, physicians may adopt different approaches to increase their charges on supplementary treatment fees. For example, the coefficient estimates in Table 2.12 show that patients at Hospital C spent significantly more on the unique service option (e.g., the traditional Chinese herbal treatment) provided by the integrated hospital rather than on other categories.

After ZMP, the average spending on medicines declined by 33% in Hospital C. Different from the two previous hospitals, Column 2 of Table 2.9 does not show a sizable increase in spending in a few medical service categories, which resulted in only a partial offset to decreases in medicine expenditure. A further examination of service utilization also did not suggest strong evidence of a systematic increase in service utilization, except for one category. Similar to Hospital B, the coefficient estimates on the utilization rates in Table 2.15 illustrate that there were no significant changes in the probability of a patient getting surgery. We also observe that the average length of stay decreased by 1.3 days, possibly due to higher prices for personal care during hospitalization. Interestingly, this hospital does not see a systematic expense increase in medical consumables. However, we find consistent evidence that physicians in Hospital C attempted to increase surplus from other services to break even the lost revenue. Specifically, the utilization rate of traditional Chinese physical treatment significantly increased for Hospital C after the ZMP. The coefficient in Column 3 of Table 2.15 shows an increase of 8.5 percentage points relative to an average utilization rate before the ZMP. This finding is also consistent with our previous interpretation that physicians are likely to induce patients to spend more on items over which they have more discretion in the treatment procedure.

2.5.4 Variation by Diagnosis

The findings indicate that physicians may pass their financial pressure onto patients by overcharging for other services, and the significant increase in additional charges following the ZMP almost offsets the spending decreases from medicine prescriptions. A related question is whether all patients are affected similarly or whether there is a shift in burdens between diagnosis groups with higher medicine expenditure and those who spent less before the ZMP.

To answer this question, we provide suggestive evidence on how the average expenditure of individuals diagnosed with different groups varies after the implementation of ZMP. Table 2.16 presents the ZMP impacts on medical expenses for Hospital A, and we observe that the policy impacts varied across diagnosis groups. Specifically, the estimates reveal that most diagnosis groups experienced significant decreases in medicine spending, although the estimates for some groups were less precisely estimated.[22] Similarly, we find significant decreases in an average patient's medicine spending across diagnoses at Hospital B, as shown in Table 2.17. Among these groups, patients diagnosed with digestive system diseases (K) had the most considerable dollar (percentage) decrease in spending compared to the pre-ZMP period mean spending.

We then report separate estimates for the ZMP impacts on other service expenditures for each diagnosis group. The tables in 2.8 show the corresponding estimates for changes for Hospital A. To summarize, relative to the pre-ZMP period, the policy led to a consistent increase (decrease) in inpatient care and diagnosis (exam) expenditure. In contrast, the impact on surgery expenditure was insignificant for most

[22]It is possible that the average prices of the drugs prescribed for specific diseases were already at the price ceilings set by the government, leaving little room for additional markup. Therefore, the removal of the 15% markup was not binding.

diagnosis groups.[23]

It is interesting to see variations in the ZMP impacts on different diagnosis groups' expenditures. Increases in other services' expenditures might have offset the decrease in drug expenses for some diagnosis groups. For example, there is an increase in the average surgery expenditure for three diagnosis groups, "Neoplasm (CD)", "Endocrine and metabolic diseases (E)", and "Genitourinary System (N)" in Hospital A. For the medical consumables, we find that six diagnosis groups in Hospital A experienced significant increases in their expenditure. It's also noteworthy that some diagnosis groups, such as those with "Musculoskeletal system and connective tissue (M)" diseases, ended up paying more for medical consumables, which could have imposed additional financial burdens on them.

To examine whether the spending increases are due to higher service utilization rates, we analyze the average utilization rates after the ZMP. The results in Table 2.13, Columns 4 and 7 show that there was an increase of surgical rates by 10% and 8%, respectively, for diagnosis groups "Musculoskeletal system and connective tissue (M)" and "Endocrine and metabolic diseases (E)" in Hospital A.[24] Additionally, we do not find significant changes in the length of stay at Hospital A across diagnosis groups. If anything, the average hospital stay decreased by 1.2 days for "Genitourinary System (N)", possibly due to a higher price and the resulting lower demand.

Similarly, while the changes in service utilization are not substantial for Hospital B, there are significant increases in the average medical consumables expenditure across almost all diagnosis groups. Moreover, for the diagnosis group "Injury, poison-

[23] As the service prices were not recorded in the HSA data, we could not directly observe if the surgical prices were adjusted.

[24] While we attempted to mitigate the effects of small-sample bias by targeting relatively common diagnosis groups, the estimates presented here may still be vulnerable to such bias, and should therefore be interpreted with care.

ing and certain other consequences of external causes (S)", the spending increase on this category is so high that it results in an even greater overall cost. Many patients in this group were diagnosed with "Bone Fracture", which was commonly reported for charging high prices for fixation plates.

Given that the compensation mechanism for Hospital C was different from the previous two, we would expect a different expense distribution. Other than the "Digestive System (K)" group, there was no significant change in the expenses on medical consumables. A breakdown of the diagnosis groups based on the traditional Chinese treatment category reveals a rise in the average spending for many diagnosis groups. Furthermore, at least part of the higher average expenditure was due to the increased usage rate of this supplementary treatment. Columns 2 and 3 of Table 2.15 show that after the ZMP, the average utilization rate increased by 22 and 12.5 percentage points for the diagnosis groups "Circulatory system (I)" and "Digestive System (K)," respectively. These are substantial usage rate increases compared to the pre-ZMP sample mean of 75 and 65 percentage points, resulting in almost a universal treatment plan for nearly all patients.

Taken together with the estimates for all hospitals, our results suggest that the reduced medicine expenditure may benefit certain patients more than others. However, the increases in spending on other services partly or entirely offset the potential benefits of the ZMP, resulting in disproportionate impacts across different diagnosis groups.

2.5.5 Additional Analyses

Our empirical strategy assumes that the timing of local ZMP implementation is plausibly uncorrelated to the potential medical expenditure trajectories. The provincial government determined the rollout timeline so that there were no simultaneous unobserved changes in the local hospitals that affected both physicians' behaviors and the policy implementation date. A primary threat to the identification in this context is that physicians have anticipated the ZMP timing and adjusted their prescription behaviors before the ZMP. For example, realizing the stop of future profits from medications upon ZMP, they may prescribe more profitable medicines in advance.

We use an event-study strategy to see if any empirical pattern supports this hypothesis, as stated in equation 2.3. The event-study design identifies estimates in a dynamic framework based on within-hospital changes in outcome variables in the calendar before and after a policy change. If physicians attempted to increase patients' spending by prescribing more profitable drugs, we should have seen a spike, at least in the months near the ZMP dates. However, the insignificant coefficients β_1 to β_7 preclude such a possibility, as we do not see any differences in the outcomes in the pre-ZMP months.

We present the coefficient estimates from the event study analysis for each hospital and each outcome in Figures 2.7 to 2.23. We normalize the values of the expenses using the mean and the standard deviation of the corresponding outcome costs in 2016. The coefficient estimates, therefore, measure the number of standard deviations that the outcomes deviate from the mean of the previous year. While there is a sharp decline in the average medicine expense, as indicated by the negative and significant coefficients for the post-ZMP months from August to December, the coefficient estimates of all pre-ZMP months are insignificant. The results show that physicians, although they may have anticipated the ZMP, did not react accordingly by inducing patients to purchase more medications. In addition, we also detect supportive evidence for parallel pre-trend in the primary outcomes for each hospital in the months leading up to the ZMP implementation date.

Another potential threat to identification is that the ZMP implementation timing happened at the same time as other changes in the patient pool. For instance, the hospitals may not adjust the service prices upward completely to the price cap allowed by the local government. Since all hospitals in our sample are comprehensive, the patients who expected a more complicated treatment may have switched to the one with a relatively lower treatment price, all else equal. In this case, the estimates based on ZMP are confounded by the changes in the patients' characteristics. While we cannot observe the price tags for each service in each hospital, we adopt an implicit strategy to address this concern. We separately examine the patients' demographic composition by disease for each hospital. The results are presented in Table 2.24. In summary, for each hospital, we do not detect a significant change in the age and gender composition of patients by diagnosis groups.

2.6 Conclusion

The high cost of pharmaceuticals has raised concerns in China, with many pointing to public hospitals and the financial incentives tied to physicians' prescribing behavior as crucial contributing factors. To address this issue and alleviate the financial burden on patients, the Chinese government implemented a nationwide program to reduce over-prescription and control medication costs. As part of this program, the previously permitted 15% profit margin on drug sales at public hospitals was eliminated. By the end of 2017, all public hospitals across the country had adopted this policy change, signaling a significant shift in the approach to pharmaceutical pricing and prescribing practices.

This paper investigates the effectiveness of government-led health care reform in reducing health care expenditure for individuals. Using a unique dataset of admission-level health care claims from a representative city, we examine the effectiveness of the reform in light of the plausibly exogenous change in retail medicine prices resulting from the ZMP. Our estimates suggest that while the policy led to a decrease in medicine expenses, it did not result in a reduction in the final medical bill for an average individual.

The evidence presented in this study leads to several main conclusions. First, we find that hospitals may have employed different strategies to compensate for their drug revenue loss resulting from the ZMP, including physicians prescribing more expensive materials for treatment, which may be buried in the final medical bills. For example, the patients at the two general-purpose hospitals in our sample paid more for medical consumables. In contrast, the patients in the integrated hospital were prescribed more traditional supplementary treatment. It highlights the need for transparency and oversight in medical billing practices to ensure patients are not burdened with unnecessary costs. Second, while there was a slight increase in surgery rates in one hospital for specific diseases, overall, we did not observe a significant change in surgery rates across hospitals. It suggests that patients and physicians may exercise caution when considering surgical interventions, and decisions regarding surgeries are influenced by multiple factors beyond financial incentives. Last, our study highlights the need to monitor unintended outcomes, such as medical resource utilization across diagnosis groups, in addition to targeted medicine expenditure. Depending on the severity of the illness, some patients may be more susceptible to unnecessary medical resource usage than others, leading to an even higher financial burden than before ZMP.

From a policy perspective, the findings of this study suggest that simply implementing price controls for drugs and services may not be sufficient to address the underlying incentives and behaviors of physicians within the health care system. As long as physicians' remuneration remains dependent on hospital revenue, the distorted principal-agent incentives in the Chinese health care system cannot be eliminated. One potential avenue for further research and policy consideration is the implementation of lump sum compensation or alternative compensation models that detach physicians' income from the revenue generated by the services they provide.

2.7 Tables

Table 2.1: Diagnosis Distributions in Hospital A

No. Obs	Hospital A	
Diagnosis Groups	2016	2017
I (Circulatory system)	3,323 (29%)	3,161 (27%)
CD (Neoplasms)	1,532 (13%)	1,712 (15%)
K (Digestive system)	1,679 (15%)	1,556 (13%)
J (Respiratory system)	1,015 (8.8%)	1,111 (9.5%)
N (Genitourinary System)	943 (8.2%)	1,071 (9.1%)
R (Abdominal Pain)	953 (8.3%)	1,048 (8.9%)
E (Endocrine and metabolic diseases)	693 (6.0%)	775 (6.6%)
H (Ear and Mastoid Process)	762 (6.6%)	633 (5.4%)
M (Musculoskeletal system and connective tissue)	649 (5.6%)	647 (5.5%)
Total	11549	11714

Note: Daily admissions of the inpatient visits are from 2016-01-01 to 2017-12-31. Column 1 breaks down the diseases into nine categories based on the ICD-10 Classification system. Columns 2 and 3 summarizes the total number of patient visits to Hospital A (The City No.1 People's Hospital) in years 2016 and 2017, respectively. The numbers in parentheses shows the proportions of diseases in each category group.

Table 2.2: Diagnosis Distributions in Hospital B

No. Obs	Hospital B	
Diagnosis Groups	2016	2017
CD (Neoplasms)	1,457 (31%)	1,863 (32%)
I (Circulatory system)	1,336 (28%)	1,797 (31%)
K (Digestive System)	862 (18%)	1,007 (18%)
N (Genitourinary System)	541 (11%)	544 (9.5%)
S (Injury, poisoning and more)	524 (11%)	543 (9.4%)
Total	4720	5754

Note: Daily admissions of the inpatient visits are from 2016-01-01 to 2017-12-31. Column 1 breaks down the diseases into nine categories based on the ICD-10 Classification system. Columns 2 and 3 summarizes the total number of patient visits to Hospital B (The City No.2 People's Hospital) in years 2016 and 2017, respectively. The numbers in parentheses shows the proportions of diseases in each category group.

Table 2.3: Diagnosis Distributions in Hospital C

No. Obs	Hospital C	
Diagnosis Groups	2016	2017
M (Musculoskeletal system and connective tissue)	873 (34%)	1,198 (39%)
I (Circulatory system)	748 (29%)	830 (27%)
K (Digestive system)	600 (23%)	680 (22%)
CD (Neoplasms)	352 (14%)	356 (12%)
Total	2573	3064

Note: Daily admissions of the inpatient visits are from 2016-01-01 to 2017-12-31. Column 1 breaks down the diseases into nine categories based on the ICD-10 Classification system. Columns 2 and 3 summarizes the total number of patient visits to Hospital C (City's Hospital of Traditional Chinese Medicine) in years 2016 and 2017, respectively. The numbers in parentheses shows the proportions of diseases in each category group.

Table 2.4: Patient Demographics

	Hospital A			Hospital B			Hospital C		
	All	Before	After	All	Before	After	All	Before	After
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Age (%)									
Below 46	0.25	0.26	0.24	0.28	0.29	0.25	0.31	0.32	0.28
Age 46 - 54	0.28	0.29	0.28	0.25	0.25	0.25	0.32	0.32	0.33
Age 54 -65	0.25	0.24	0.26	0.24	0.23	0.24	0.21	0.21	0.23
Above 65	0.22	0.22	0.23	0.23	0.22	0.26	0.15	0.15	0.16
Gender (%)									
Male	0.55	0.56	0.55	0.58	0.58	0.6	0.53	0.52	0.55
Obs	23263	18155	5108	10474	8001	2473	5637	4320	1317

Note: Daily admissions of the inpatient visits are from 2016-01-01 to 2017-12-31. We define the "Before" period to be the calendar days from 2016-01-01 to 2016-07-31; and from 2017-01-01 to 2017-07-31. The "After" period is the calendar days from 08-01 to 12-31 in both two years. Columns 1, 4, 7 summarize the patients' demographic characteristics based on age and gender of all inpatient visits. The rest columns break down the inpatient sample into the before- and after ZMP calendar days and summarize the sub-sample patient characteristics.

Table 2.5: Expenditure By Category

	All		Before		After	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Hospital A						
Total Exp (CNY)	9088	9930	9035	9233	9278	9541
Medicine Exp (CNY)	3230	4417	3348	4576	2809	3772
General Treatments (CNY)	1367	2392	1269	2117	1713	3161
Surgery Exp (CNY)	675	1911	627	1925	843	1852
Examination Exp (CNY)	2940	1887	2965	1908	2848	1805
Medical Consumables Exp (CNY)	878	2752	825	2622	1065	3163
Hospital B						
Total Exp (CNY)	8574	7669	8598	7273	8493	8829
Medicine Exp (CNY)	3156	3792	3361	3914	2494	3282
General Treatments (CNY)	1420	2513	1317	1884	1753	3889
Surgery Exp (CNY)	538	1238	532	1231	560	1259
Examination Exp (CNY)	2724	1605	2738	1642	2680	1475
Medical Consumables Exp (CNY)	735	2168	651	1859	1006	2939
Hospital C						
Total Exp (CNY)	7373	5663	7442	5979	7148	4450
Medicine Exp (CNY)	2227	3336	2388	3600	1698	2179
Traditional Chinese Treatments (CNY)	850	1065	803	980	1007	1292
General Treatments (CNY)	1272	1267	1265	1266	1296	1272
Surgery Exp (CNY)	412	922	370	897	551	988
Examination Exp (CNY)	2177	1324	2179	1332	2171	1298
Medical Consumables Exp (CNY)	434	1459	437	1610	426	782

Note: Daily admissions of the inpatient visits are from 2016-01-01 to 2017-12-31. The three panels show the mean and standard deviation of the sub-category expenditures for Hospitals A (The City No.1 People's Hospital), B(The City No.2 People's Hospital), C(City's Hospital of Traditional Chinese Medicine), respectively. The expenses are denominated in the local currency (Chinese Yuan, CNY).

Table 2.6: Service Utilization By Category

Variables	All Mean (1)	Before Mean (2)	After Mean (3)
Hospital A			
Exam Rates	0.99	0.99	1
Surgery Rates	0.39	0.39	0.42
Medical Consumable Rates	1	1	1
Duration (days)	12.3	12.4	11.7
Hospital B			
Exam Rates	1	1	1
Surgery Rates	0.37	0.38	0.34
Medical Consumable Rates	1	1	1
Duration (days)	12.6	12.7	12.1
Hospital C			
Exam Rates	0.99	0.99	0.99
Traditional Chinese Treat Rates	0.84	0.82	0.92
Surgery Rates	0.41	0.41	0.43
Medical Consumable Rates	0.99	1	0.99
Duration (days)	16.5	16.7	15.9

Note: Daily admissions of the inpatient visits are from 2016-01-01 to 2017-12-31. The three panels show the mean and standard deviation of the sub-category service utilization rates for Hospitals A (The City No.1 People's Hospital), B(The City No.2 People's Hospital), C(City's Hospital of Traditional Chinese Medicine), respectively. The utilization rate represents the proportion of inpatient visits charged with the corresponding services. Duration measures the days of stay at the hospital. The expenses are denominated in the local currency (Chinese Yuan, CNY).

Table 2.7: ZMP Impacts on Medicine and Total Expenditure in Hospital A

Hospital A	(1) Medicine Expenditure	(2) Non-Medicine Expenditure	(3) Total Expenditure
ZMP	-598.988***	446.899***	-152.089
	(114.080)	(155.441)	(243.146)
month FE	Y	Y	Y
day-of-week FE	Y	Y	Y
wave FE	Y	Y	Y
Demographics	Y	Y	Y
Mean	3348	5687	9035
N	23263	23263	23263

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average medicine, non-medicine, and total expenditure charged to patients in Hospital A (The City No.1 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average patient's expense. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.8: ZMP Impacts on Medicine and Total Expenditure in Hospital B

Hospital B	(1) Medicine Expenditure	(2) Non-Medicine Expenditure	(3) Total Expenditure
ZMP	-668.920***	518.583***	-150.337
	(146.636)	(189.364)	(300.551)
month FE	Y	Y	Y
day-of-week FE	Y	Y	Y
wave FE	Y	Y	Y
Demographics	Y	Y	Y
Mean	3361	5237	8598
N	10474	10474	10474

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average medicine, non-medicine, and total expenditure charged to patients in Hospital B (The City No.2 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average patient's expense. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.9: ZMP Impacts on Medicine and Total Expenditure in Hospital C

Hospital C	(1) Medicine Expenditure	(2) Non-Medicine Expenditure	(3) Total Expenditure
ZMP	-788.487***	135.141	-653.346**
	(168.159)	(176.075)	(301.417)
month FE	Y	Y	Y
day-of-week FE	Y	Y	Y
wave FE	Y	Y	Y
Demographics	Y	Y	Y
Mean	2388	5054	7442
N	5637	5637	5637

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average medicine, non-medicine, and total expenditure charged to patients in Hospital C (City's Hospital of Traditional Chinese Medicine). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average patient's expense. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.10: ZMP Impacts on Service Expenditure in Hospital A

Hospital A	(1) Inpatient and Post-Surgery Care	(2) Examination	(3) Surgical Treatment	(4) Medical Consumable
ZMP	350.368***	-310.832***	95.353*	312.010***
	(62.244)	(48.676)	(49.778)	(72.352)
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y
Mean	1269	2965	627	825
N	23263	23263	23263	23263

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the expenditure of the sub-service groups charged to patients in Hospital A (The City No.1 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average patient's expense. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.11: ZMP Impacts on Service Expenditure in Hospital B

Hospital B	(1) Inpatient and Post-Surgery Care	(2) Examination	(3) Surgical Treatment	(4) Medical Consumable
ZMP	385.188***	-289.170***	2.348	420.216***
	(98.288)	(59.916)	(46.079)	(85.455)
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y
Mean	1317	2738	532	651
N	10474	10474	10474	10474

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the expenditure of the sub-service groups charged to patients in Hospital B (The City No.2 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average patient's expense. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.12: ZMP Impacts on Service Expenditure in Hospital C

Hospital C	(1) Inpatient and Post-Surgery Care	(2) Examination	(3) Surgical Treatment	(4) Medical Consumable	(5) Traditional Chinese Physical Therapy
ZMP	106.009	-307.069***	105.920**	51.562	178.719***
	(66.620)	(68.957)	(47.907)	(77.722)	(49.008)
month FE	Y	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y	Y
wave FE	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y
Mean	1265	2179	370	437	803
N	5637	5637	5637	5637	5637

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the expenditure of the sub-service groups charged to patients in Hospital C (City's Hospital of Traditional Chinese Medicine). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average patient's expense. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.13: ZMP Impacts on Service Utilization in Hospital A

Hospital A	(1) Duration	(2) Surgery rates
ZMP	-0.537**	0.030**
	(0.25)	(0.01)
Mean	11.4	0.39
month FE	Y	Y
day-of-week FE	Y	Y
wave FE	Y	Y
Demographics	Y	Y
N	23263	23263

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the utilization of inpatient services used by patients in Hospital A (The City No.1 People's Hospital). The variable ZMP measures the impacts of the zero markup policy on the average service utilization. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.14: ZMP Impacts on Service Utilization in Hospital B

Hospital B	(1) Duration	(2) Surgery rates
ZMP	-0.242	-0.024
	(0.34)	(0.02)
Mean	11.7	0.38
month FE	Y	Y
day-of-week FE	Y	Y
wave FE	Y	Y
Demographics	Y	Y
N	10474	10474

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the utilization of inpatient services used by patients in Hospital B (The City No.2 People's Hospital). The variable ZMP measures the impacts of the zero markup policy on the average service utilization. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.15: ZMP Impacts on Service Utilization in Hospital C

Hospital C	(1) Surgery Rate	(2) Duration	(3) Chinese Treatment rate
ZMP	-0.016	-1.267**	0.085***
	(0.025)	(0.589)	(0.018)
Mean	0.41	15.7	0.81
month FE	Y	Y	Y
day-of-week FE	Y	Y	Y
wave FE	Y	Y	Y
Demographics	Y	Y	Y
N	5637	5637	5637

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. The dependent variables are the utilization of inpatient services used by patients in Hospital C (City's Hospital of Traditional Chinese Medicine). The variable ZMP measures the impacts of the zero markup policy on the average service utilization. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.16: ZMP Impacts on Medicine Expense by Diagnosis Groups, Hospital A

Hospital A	(1) CD	(2) I	(3) K	(4) M	(5) R
ZMP	-1287.086***	-466.845*	-384.097	-586.847**	-671.583*
	(422.984)	(247.262)	(287.332)	(262.125)	(347.819)
N	3244	6484	3235	1296	2001
Mean	4711.05	3249.8	3653.8	2175.7	2983.9
	(6) J	(7) E	(8) H	(9) N	
ZMP	-741.465**	-361.797	-281.279	-501.170*	
	(363.121)	(223.071)	(213.375)	(297.772)	
N	2126	1468	1395	2014	
Mean	3865.2	2418.22	1878.216	3261.9	
month FE	Y	Y	Y	Y	
day-of-week FE	Y	Y	Y	Y	
wave FE	Y	Y	Y	Y	
Demographics	Y	Y	Y	Y	

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average medicine expenditure charged to patients in Hospital A (No.1 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average medicine expenditure by diagnosis groups. The diagnosis groups are categorized based on ICD-10 classification systems. Refer to Table 2.1 for details. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.17: ZMP Impacts on Medicine Expense by Diagnosis Groups, Hospital B

Hospital B	(1) CD	(2) I	(3) K	(4) N	(5) S
ZMP	-658.595*	-646.827***	-859.104***	-513.803**	-547.846**
	(338.818)	(249.951)	(323.998)	(260.149)	(254.501)
N	3320	3133	1869	1085	1067
Mean	3948.7	3124.1	3561.4	2557.3	2702.3
month FE	Y	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y	Y
wave FE	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average medicine expenditure charged to patients in Hospital A (No.1 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average medicine expenditure by diagnosis groups. The diagnosis groups are categorized based on ICD-10 classification systems. Refer to Table 2.2 for details. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.18: ZMP Impacts on Medicine Expense by Diagnosis Groups, Hospital C

Hospital C	(1) CD	(2) I	(3) K	(4) M
ZMP	-1846.938	-776.191***	-677.499***	-548.155***
	(1135.050)	(238.991)	(249.177)	(103.895)
N	708	1578	1280	2071
Mean	4811.1	2611.9	2592	1208.5
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average medicine expenditure charged to patients in Hospital C (City's Hospital of Traditional Chinese Medicine). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average medicine expenditure by diagnosis groups. The diagnosis groups are categorized based on ICD-10 classification systems. Refer to Table 2.3 for details. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.19: ZMP Impacts on Consumable Expense by Diagnosis Groups, Hospital A

Hospital A	(1) CD	(2) I	(3) K	(4) M	(5) R
ZMP	255.342*	391.436*	401.261***	973.985***	244.735
	(138.567)	(226.457)	(104.285)	(232.709)	(152.948)
N	3244	6484	3235	1296	2001
Mean	631.1	1276.97	690.6	601.7	561.9
	(6) J	(7) E	(8) H	(9) N	
ZMP	-35.211	392.196***	-68.085	311.048**	
	(113.403)	(117.197)	(101.441)	(134.115)	
N	2126	1468	1395	2014	
Mean	622.9	672.1	538.2	807.2	
month FE	Y	Y	Y	Y	
day-of-week FE	Y	Y	Y	Y	
wave FE	Y	Y	Y	Y	
Demographics	Y	Y	Y	Y	

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average consumable expenditure charged to patients in Hospital A (City No.1 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average consumable expenditure by diagnosis groups. The diagnosis groups are categorized based on ICD-10 classification systems. Refer to Table 2.1 for details. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.20: ZMP Impacts on Consumable Expense by Diagnosis Groups, Hospital B

Hospital B	(1) CD	(2) I	(3) K	(4) N	(5) S
ZMP	373.242***	209.584	298.090**	277.452**	1769.625***
	(118.607)	(206.764)	(116.129)	(131.498)	(378.976)
N	3320	3133	1869	1085	1067
Mean	498.7	706.1	675.5	720	847.8
month FE	Y	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y	Y
wave FE	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average consumable expenditure charged to patients in Hospital B (City No.2 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average consumable expenditure by diagnosis groups. The diagnosis groups are categorized based on ICD-10 classification systems. Refer to Table 2.2 for details. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.21: ZMP Impacts on Consumable Expense by Diagnosis Groups, Hospital C

Hospital C	(1) CD	(2) I	(3) K	(4) M
ZMP	-65.702	-60.553	242.776*	11.451
	(137.420)	(241.622)	(138.835)	(34.459)
N	708	1578	1280	2071
Mean	476.4	613.5	613.8	183
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average consumable expenditure charged to patients in Hospital C (City's Hospital of Traditional Chinese Medicine). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average consumable expenditure by diagnosis groups. The diagnosis groups are categorized based on ICD-10 classification systems. Refer to Table 2.3 for details. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.22: ZMP Impacts on Chinese Treatment Expense by Diagnosis Groups, Hospital C

Hospital C	(1) CD	(2) I	(3) K	(4) M
ZMP	239.052**	180.735*	122.657**	170.414*
	(114.337)	(103.896)	(53.873)	(95.505)
N	708	1578	1280	2071
Mean	308.6	581.3	239.6	1490.2
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average Chinese treatment expense charged to patients in Hospital C (City's Hospital of Traditional Chinese Medicine). The expenses are denominated in the local currency (Chinese Yuan, CNY). The variable ZMP measures the impacts of the zero markup policy on the average Chinese treatment expense by diagnosis groups. The diagnosis groups are categorized based on ICD-10 classification systems. Refer to Table 2.3 for details. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.23: ZMP Impacts on Chinese Treatment Rates by Diagnosis Groups, Hospital C

Hospital C	(1) CD	(2) I	(3) K	(4) M
ZMP	0.003	0.220***	0.125**	-0.002
	(0.062)	(0.040)	(0.052)	(0.009)
N	708	1578	1280	2071
Mean	0.74	0.75	0..65	0.99
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the average utilization rate of Chinese treatment of patients in Hospital C (City's Hospital of Traditional Chinese Medicine). The variable ZMP measures the impacts of the zero markup policy on the average Chinese treatment utilization rate by diagnosis groups. The diagnosis groups are categorized based on ICD-10 classification systems. Refer to Table 2.3 for details. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

Table 2.24: ZMP Impacts on Age and Gender Composition by Hospitals

	Hospital A		Hospital B		Hospital C	
	(1) Age	(2) Female proportion	(3) Age	(4) Female proportion	(5) Age	(6) Female proportion
ZMP	0.012	0.004	0.340	-0.007	-0.232	0.028
	(0.345)	(0.013)	(0.536)	(0.019)	(0.619)	(0.026)
month FE	Y	Y	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y	Y	Y
wave FE	Y	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y	Y
N	23263	23263	10474	10474	5637	5637

*** p<0.01, ** p<0.05, * p<0.1. The dependent variables are the age and female proportion of patients in Hospital A (City No.1 People's Hospital), B (City No.2 People's Hospital), and C (City's Hospital of Traditional Chinese Medicine), respectively. The variable ZMP measures the impacts of the zero markup policy on the patient composition across hospitals. The analyses include calendar year, month, day-of-week fixed effects, and individual's demographic controls such as diagnosis groups, gender and age.

2.8 Figures

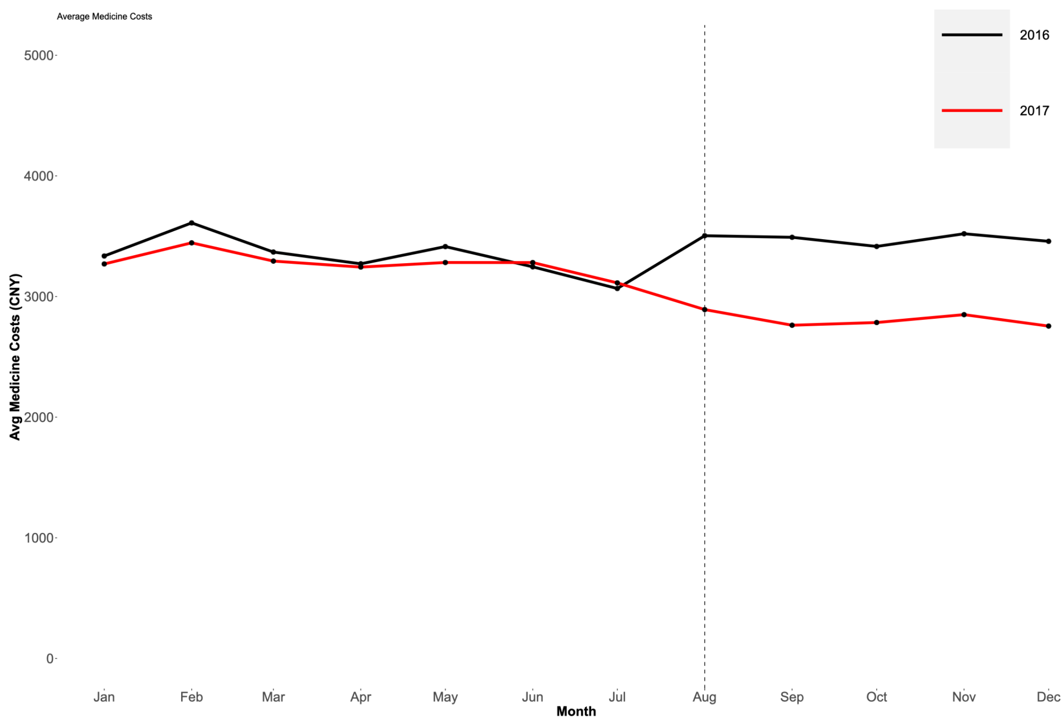

Figure 2.1: Average Medicine Expenditure in Local Currency, Hospital A

Note: This figure depicts trends of the average patient's spending on medicine prescriptions in Hospital A (City No.1 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). Daily trends were calculated as the monthly average of daily records, aggregated across patients in Hospital A. The dashed vertical line on August 1st indicates the date after the official implementation of ZMP. The black (red) line indicates the trend of patient records in the calendar year 2016 (2017).

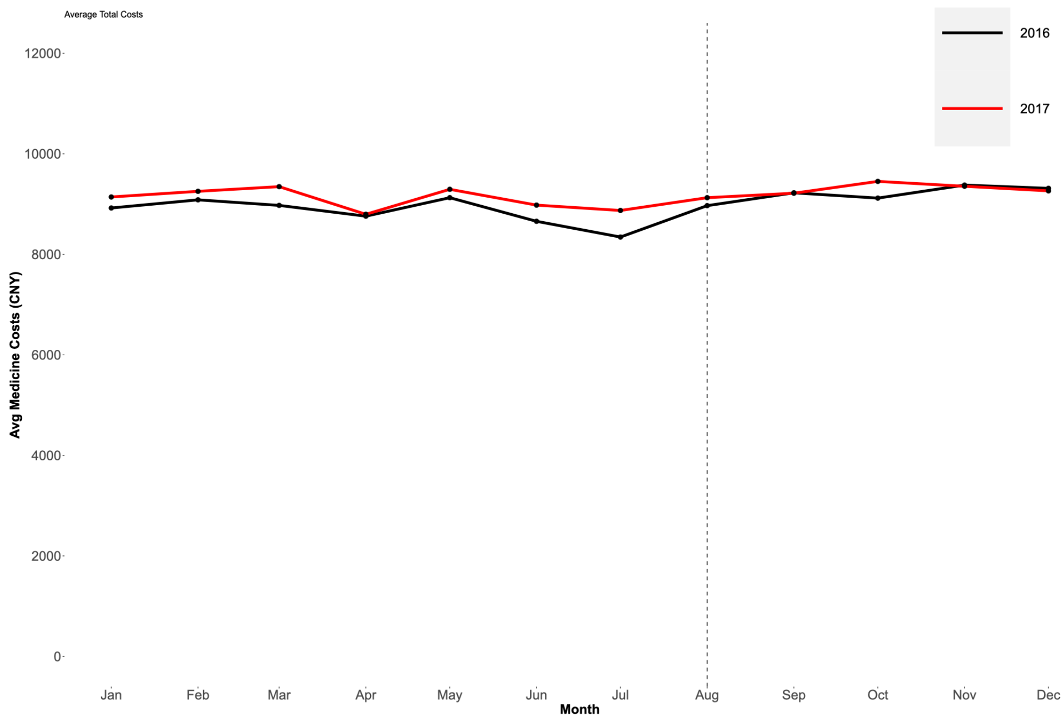

Figure 2.2: Total Medical Expenditure in Local Currency, Hospital A

This figure depicts trends of average patient's total spending in Hospital A (City No.1 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). Daily trends were calculated as the monthly average of daily records, aggregated across patients in Hospital A. The dashed vertical line on August 1st indicates the date after the official implementation of ZMP. The black (red) line indicates the trend of patient records in the calendar year 2016 (2017).

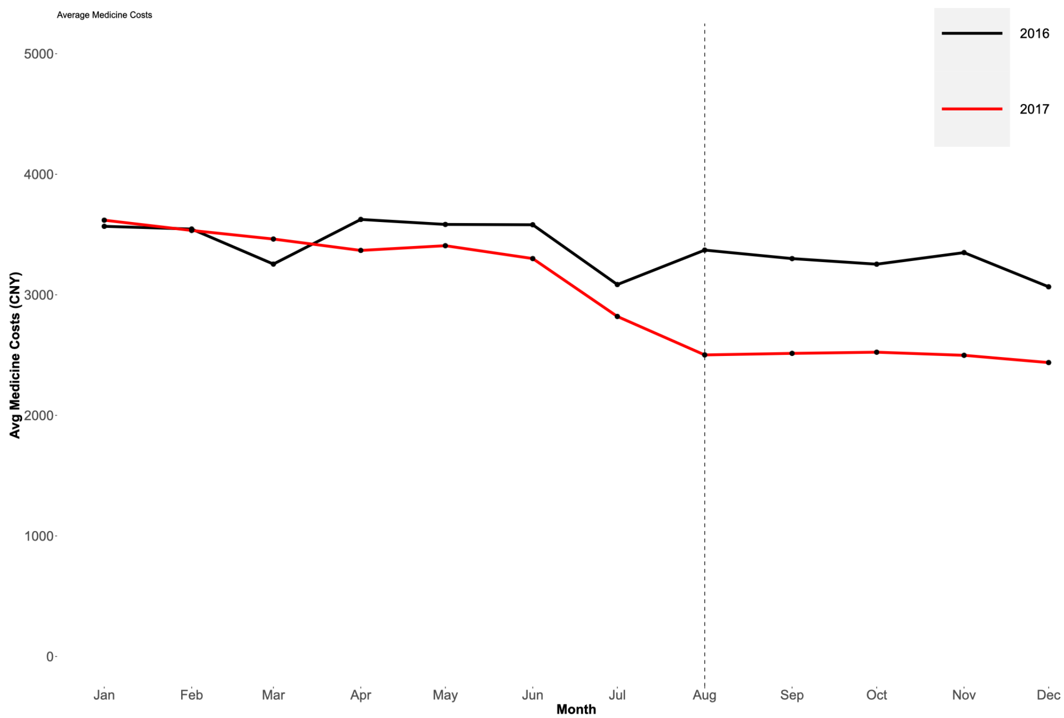

Figure 2.3: Average Medicine Expenditure in Local Currency, Hospital B

Note: This figure depicts trends of the average patient's spending on medicine prescriptions in Hospital B (City No.2 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). Daily trends were calculated as the monthly average of daily records, aggregated across patients in Hospital B. The dashed vertical line on August 1st indicates the date after the official implementation of ZMP. The black (red) line indicates the trend of patient records in the calendar year 2016 (2017).

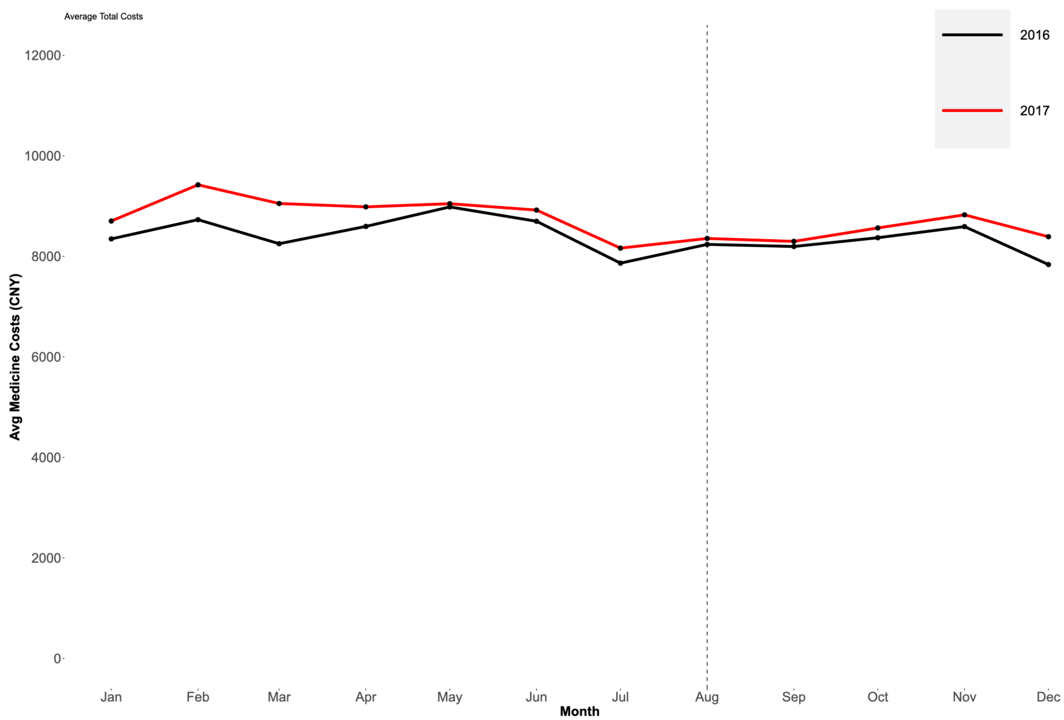

Figure 2.4: Total Medical Expenditure in Local Currency, Hospital B

Note: This figure depicts trends of average patient's total spending in Hospital B (City No.2 People's Hospital). The expenses are denominated in the local currency (Chinese Yuan, CNY). Daily trends were calculated as the monthly average of daily records, aggregated across patients in Hospital B. The dashed vertical line on August 1st indicates the date after the official implementation of ZMP. The black (red) line indicates the trend of patient records in the calendar year 2016 (2017).

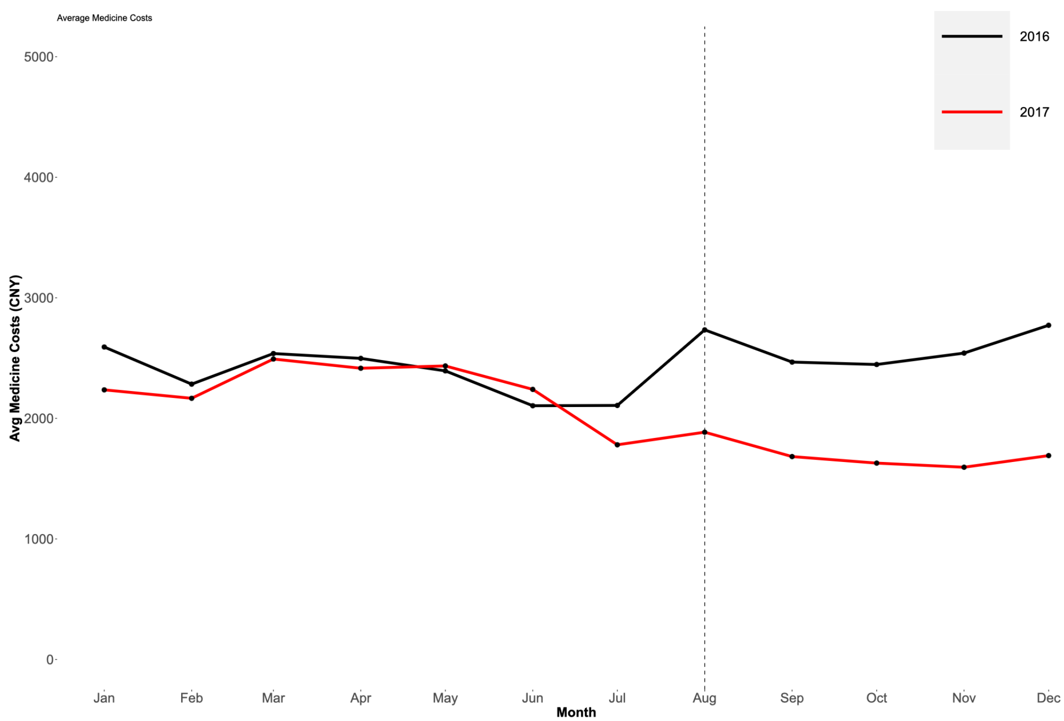

Figure 2.5: Average Medicine Expenditure in Local Currency, Hospital C

This figure depicts trends of average patient's spending on medicine prescriptions in Hospital C (City's Hospital of Traditional Chinese Medicine). The expenses are denominated in the local currency (Chinese Yuan, CNY). Daily trends were calculated as the monthly average of daily records, aggregated across patients in Hospital C. The dashed vertical line on August 1st indicates the date after the official implementation of ZMP. The black (red) line indicates the trend of patient records in the calendar year 2016 (2017).

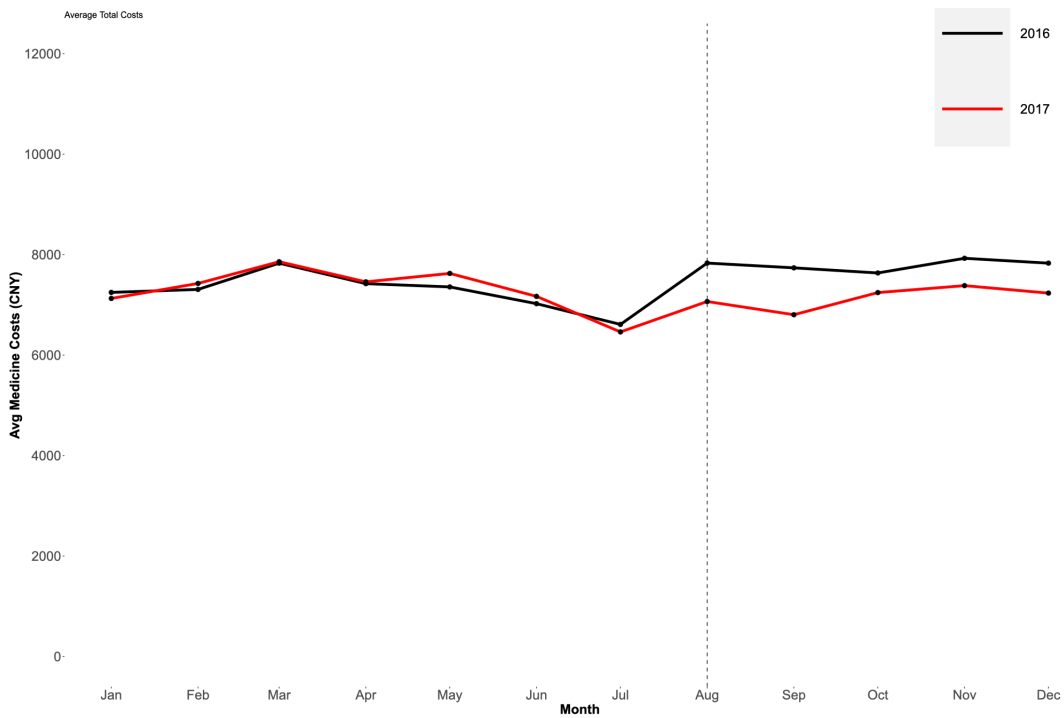

Figure 2.6: Total Medical Expenditure in Local Currency, Hospital C

Note: This figure depicts trends of average patient's total spending in Hospital C (City's Hospital of Traditional Chinese Medicine). The expenses are denominated in the local currency (Chinese Yuan, CNY). Daily trends were calculated as the monthly average of daily records, aggregated across patients in Hospital C. The dashed vertical line on August 1st indicates the date after the official implementation of ZMP. The black (red) line indicates the trend of patient records in the calendar year 2016 (2017).

Event-Study Analysis

The figures below estimate Equation 2.3 on medicine and service expenditure and the utilization rate of different service groups, respectively. The x-axis indicates the calendar months from January to December. The vertical line at x=7 indicates the baseline month of July. It is the month since the ZMP started to take effect and is the baseline group we compare our coefficients. The framework in Equation 2.3 incorporates a set of monthly indicators interacting with the ZMP year. Thus, these coefficients β_1 to β_6 capture the differences in outcome variables between the year 2016 and 2017 for months (January - June) when ZMP was not in effect, and the coefficients β_8 to β_{12} measure the outcome differences relative to July for the months since ZMP (August - December). For the sake of presentation, the outcome variables were all normalized using the mean and the standard deviation of the corresponding outcome costs in 2016. The coefficient estimates, therefore, measure the number of standard deviations that the outcomes deviate from the mean of the previous year.

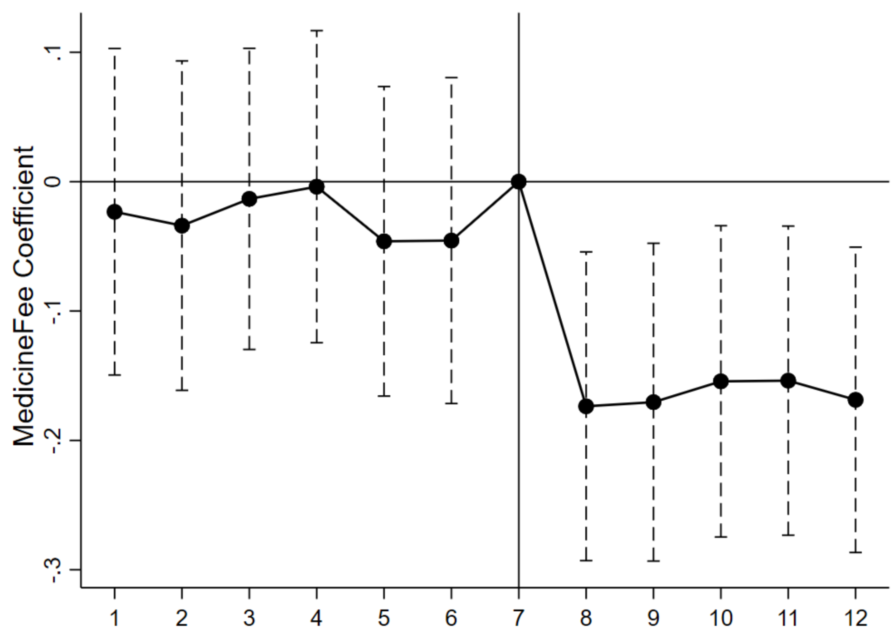

Figure 2.7: Event Study Analysis of Hospital A, ZMP Impacts on Medicine Expenditure

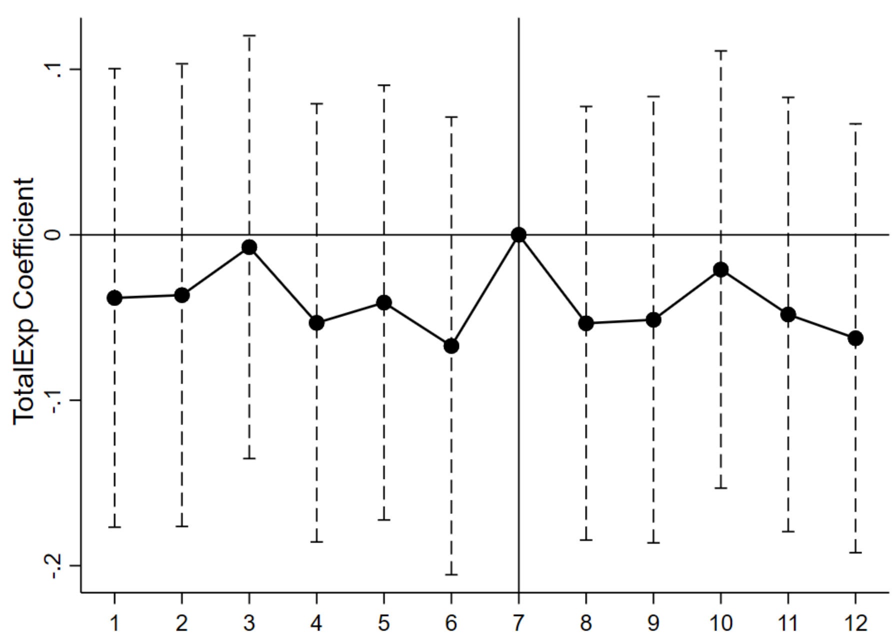

Figure 2.8: Event Study Analysis of Hospital A, ZMP Impacts on Total Expenditure

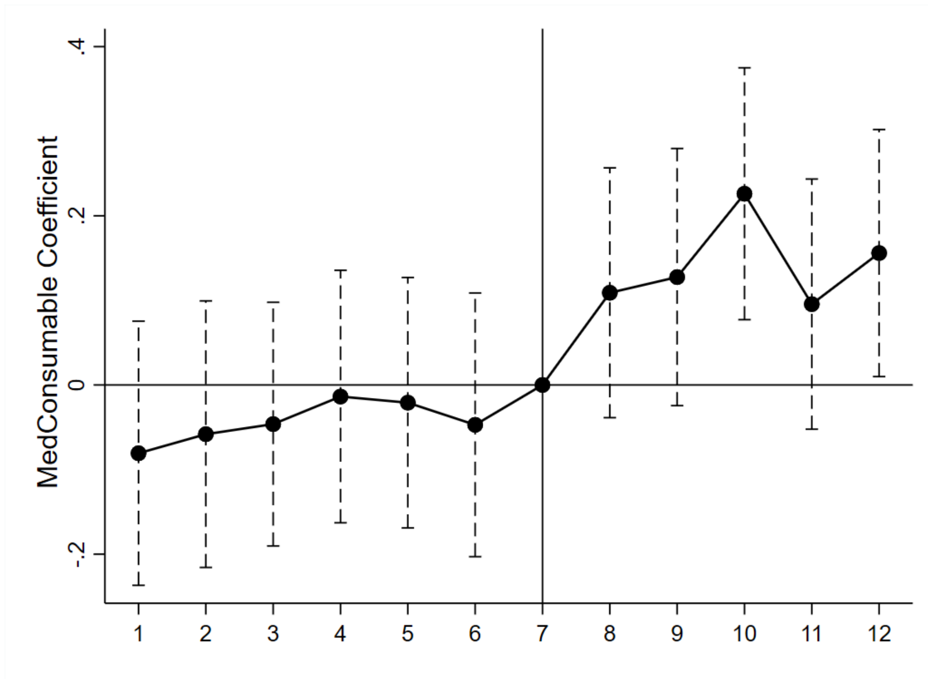

Figure 2.9: Event Study Analysis of Hospital A, ZMP Impacts on Consumable Expenditure

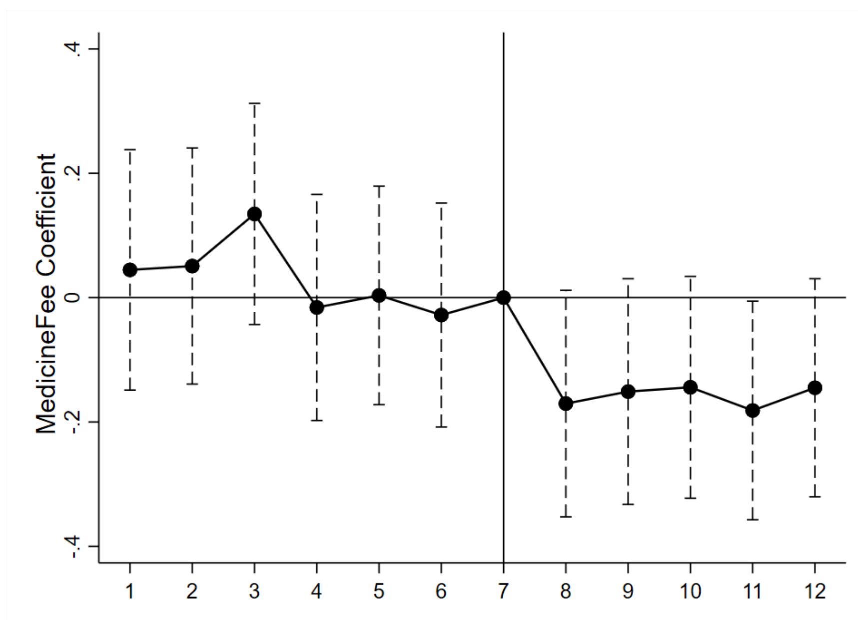

Figure 2.10: Event Study Analysis of Hospital B, ZMP Impacts on Medicine Expenditure

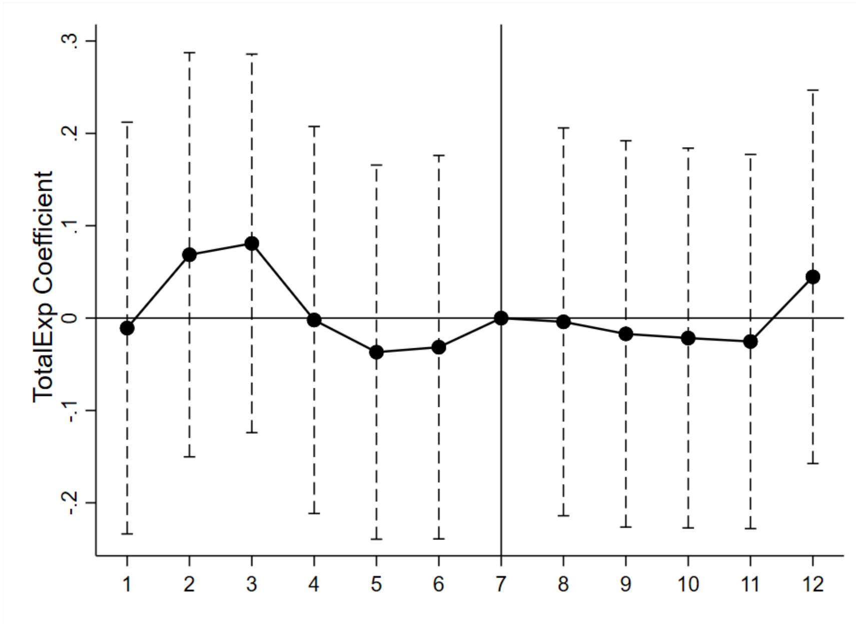

Figure 2.11: Event Study Analysis of Hospital B, ZMP Impacts on Total Expenditure

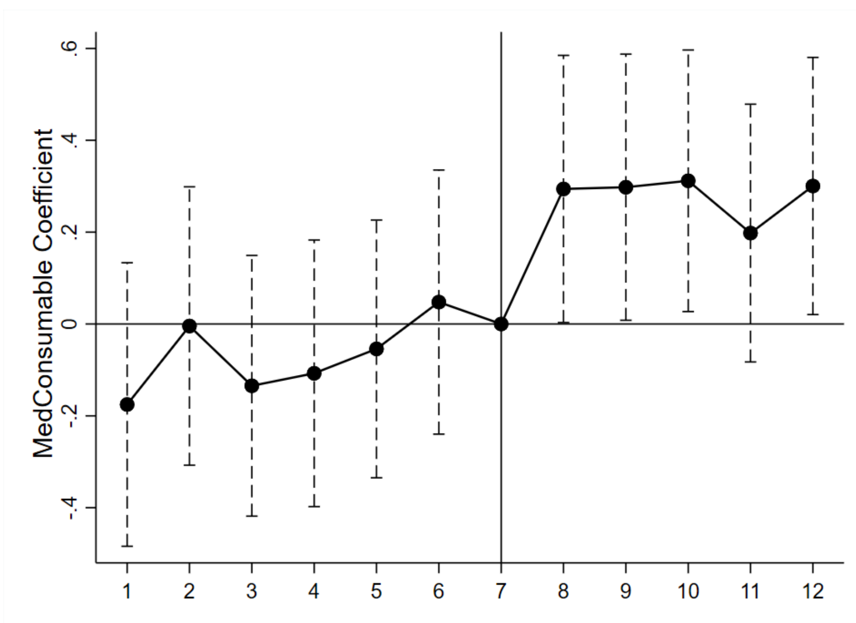

Figure 2.12: Event Study Analysis of Hospital B, ZMP Impacts on Consumable Expenditure

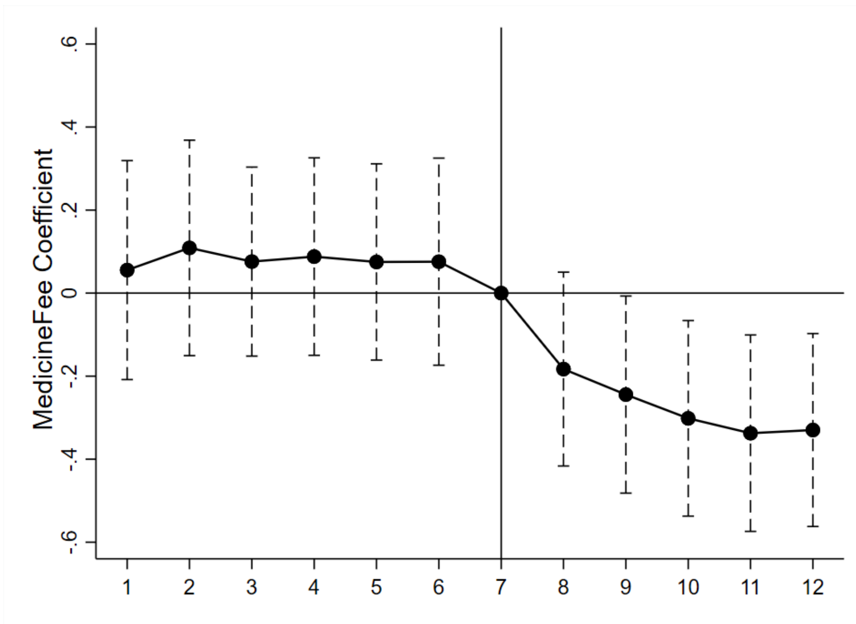

Figure 2.13: Event Study Analysis of Hospital C, ZMP Impacts on Medicine Expenditure

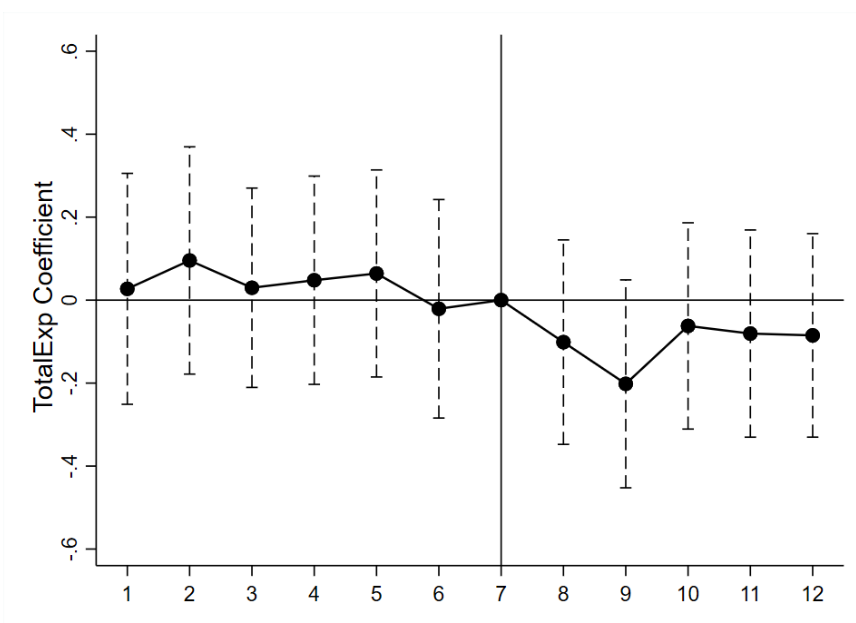

Figure 2.14: Event Study Analysis of Hospital C, ZMP Impacts on Total Expenditure

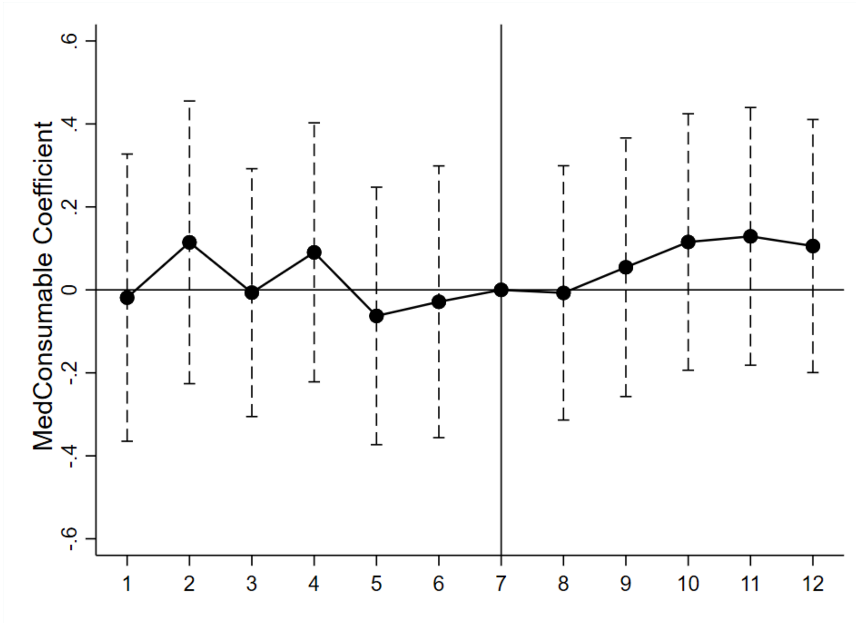

Figure 2.15: Event Study Analysis of Hospital C, ZMP Impacts on Consumable Expenditure

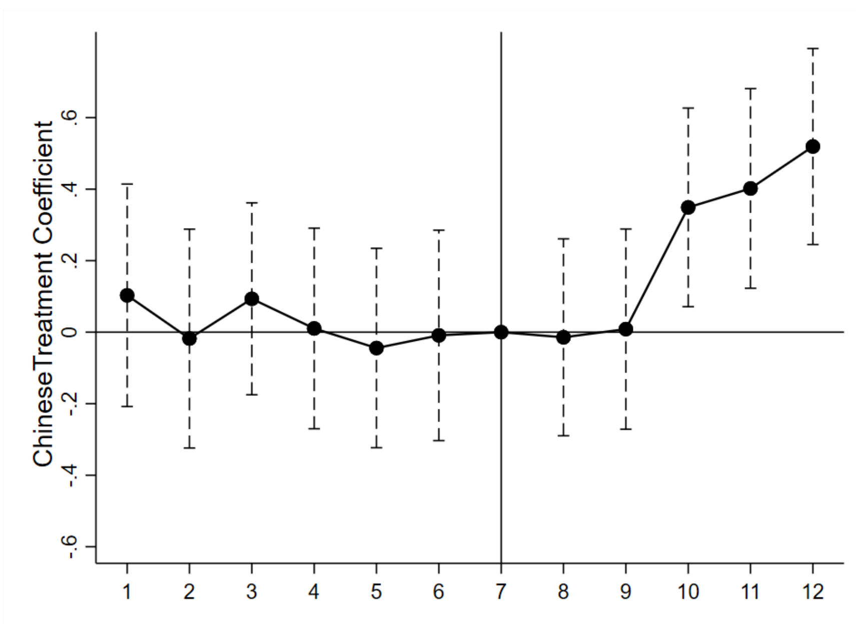

Figure 2.16: Event Study Analysis of Hospital C, ZMP Impacts on Traditional Chinese Treatment Expenditure

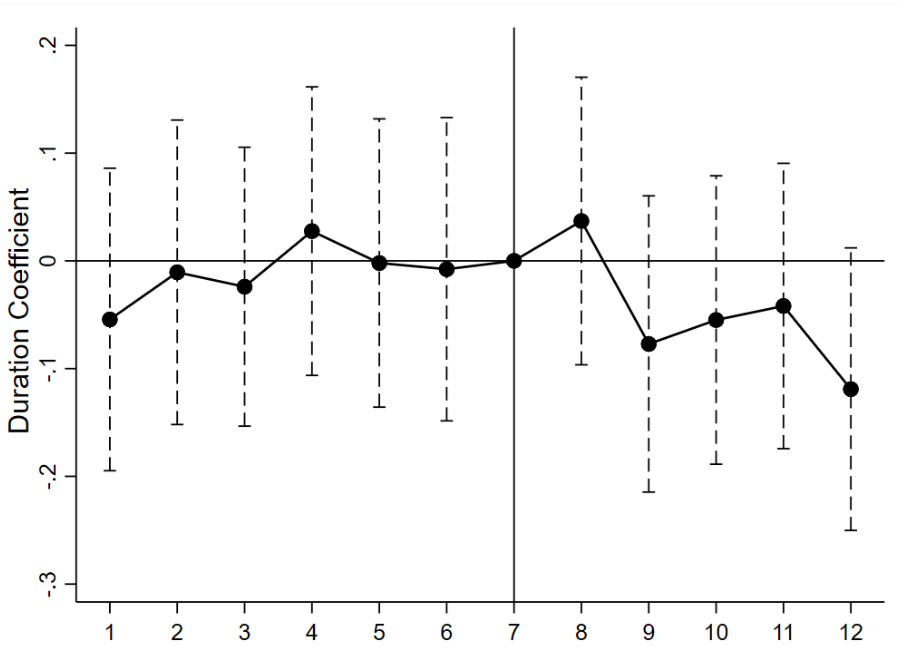

Figure 2.17: Event Study Analysis of Hospital A, ZMP Impacts on the Average Length of Stay

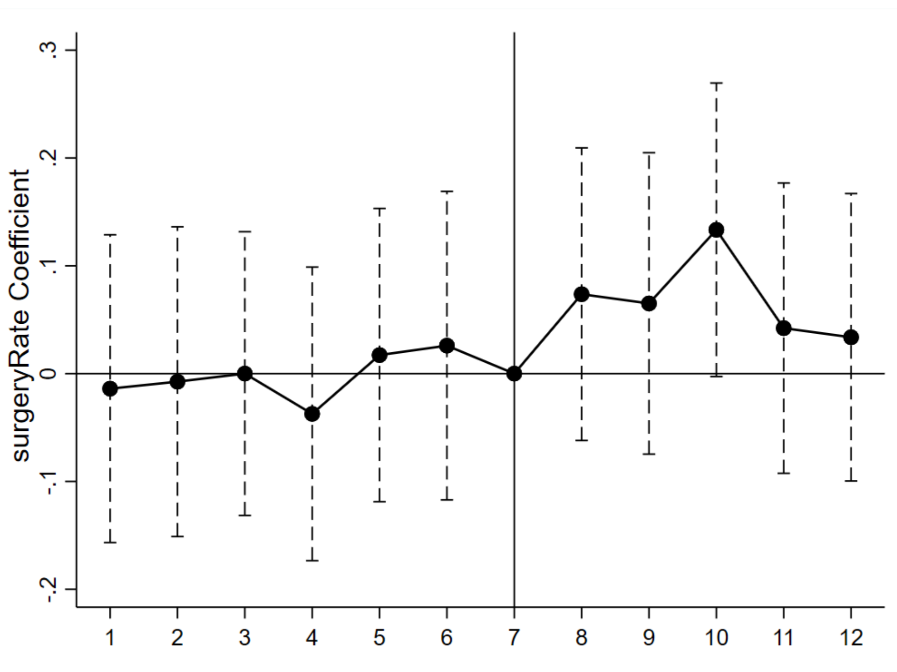

Figure 2.18: Event Study Analysis of Hospital A, ZMP Impacts on the Average Surgery Rate

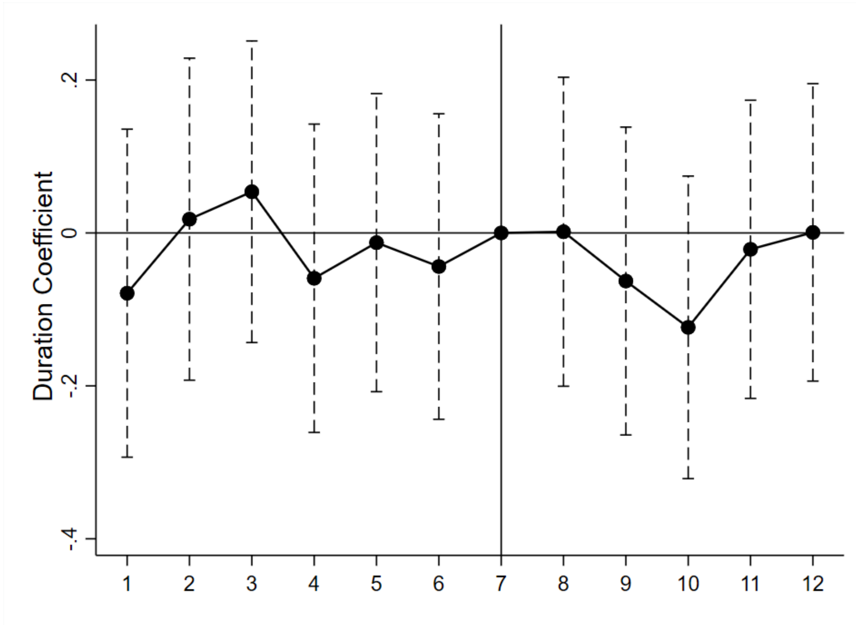

Figure 2.19: Event Study Analysis of Hospital B, ZMP Impacts on the Average Length of Stay

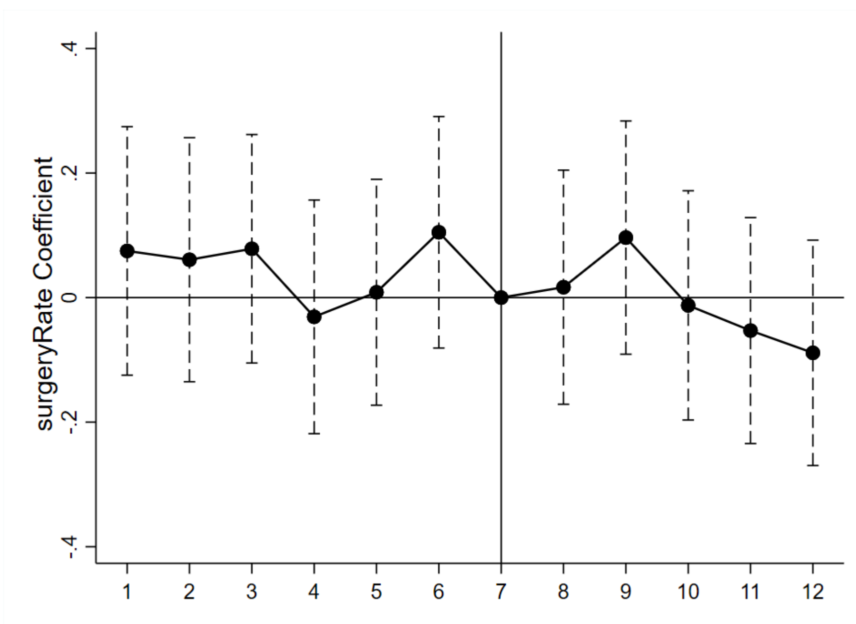

Figure 2.20: Event Study Analysis of Hospital B, ZMP Impacts on the Average Surgery Rate

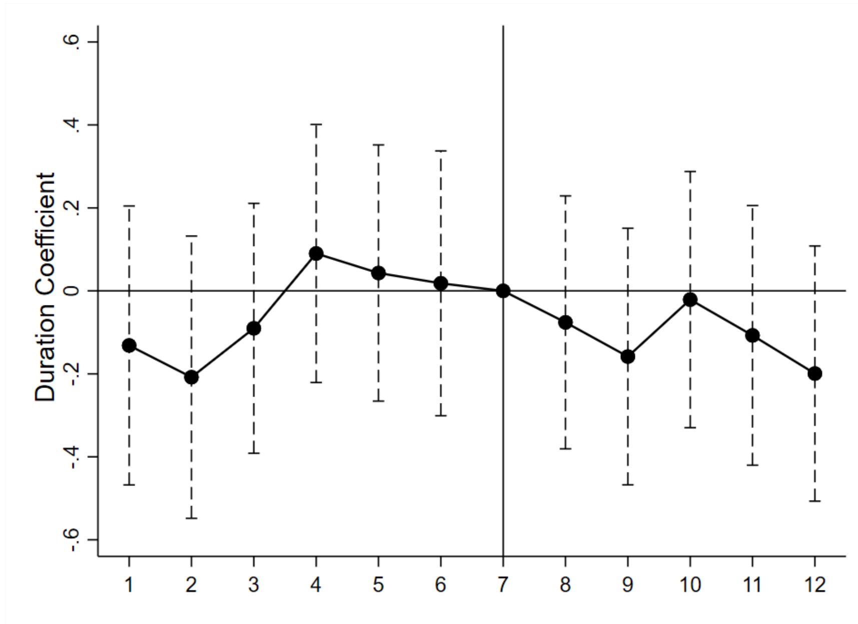

Figure 2.21: Event Study Analysis of Hospital C, ZMP Impacts on the Average Length of Stay

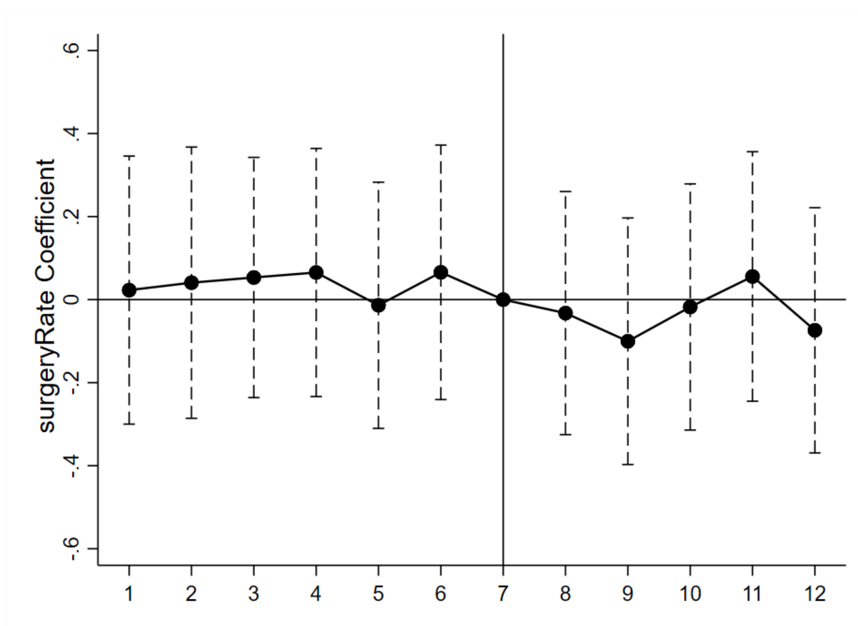

Figure 2.22: Event Study Analysis of Hospital C, ZMP Impacts on the Average Surgery Rate

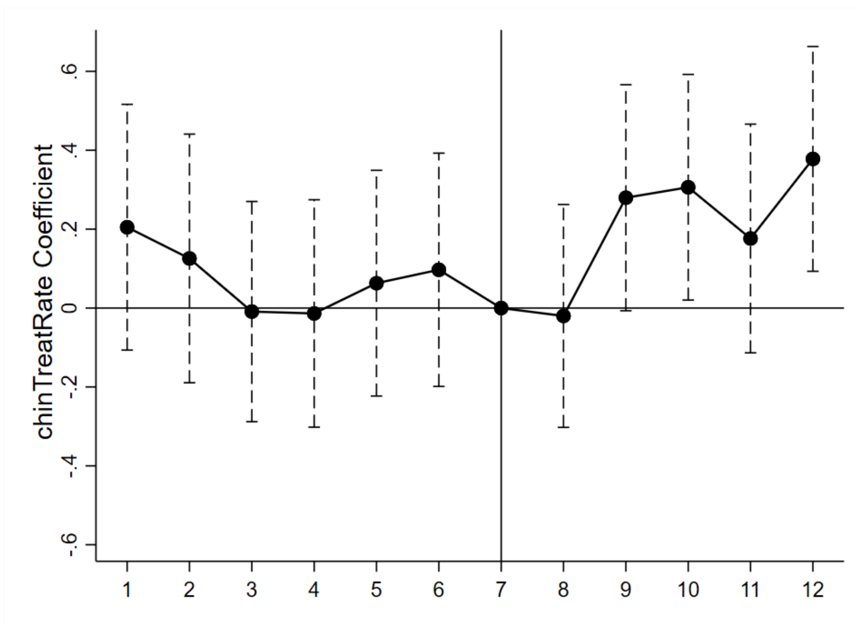

Figure 2.23: Event Study Analysis of Hospital C, ZMP Impacts on the Average Usage Rate of Traditional Chinese Treatment

Appendix

Disease Classification

We map each granular disease name to an ICD-10 sample which contains detailed disease names and their ICD codes consisting of letters and numbers. A code always starts with a letter followed by five digits (e.g., S01.001, S01.101). The first three characters designate the category of the diagnosis. For example, all diseases with an "S" initial represent a diagnosis related to "Injuries, poisoning and certain other consequences of external causes related to single body regions." If the second character is number 6, it indicates that the diagnosis falls into the category of "Injuries to the wrist, hand, and fingers".

Our method of classification is to first match the disease names from our data with the disease names from the ICD-10 sample, with the words described in Chinese Characters. We then group the matched diseases to letter codes of ICD-10. Specifically, for each disease (variable "disease sample") from our sample, we loop over the disease names in the ICD-10 sample (variable "disease ICD") and locate the ones with the most matches of Chinese characters. We choose the disease name with the most number of matched characters to be the matched disease and then link its ICD code to the matched words.

The algorithm, particularly for Chinese characters, enables groupings of the diseases into a harmonized category that can be applied systematically for the related analysis covering the local healthcare system, which information is usually too granular to be compared.

Service Price Ceiling Changes Post- ZMP

In the HSA dataset, the expenditures were recorded under a broader category. Each broader category of medical services in the document contains tens or hundreds of service items, and each may vary by unit. The city government publishes a price guidance document that lists the price ceiling for each granular level service item. For example, during the entire hospitalization experience, there could be multiple times of nursing services provided, whereas there is generally at most one surgery conducted.

The summary table below demonstrates the price ceiling changes of service categories. The guidance document published exhibits a variety of price cap changes across service types.

Table 2.25: Summary of Price Ceiling Guidance for City Hospitals

Service Category	Item coverage	New Price Regulation during ZMP?	Price ceiling change allowed
Inpatient and Post-Surgery Care	Consultation, Inpatient Care and Nursing Fee	Y	unit price * (2-3 times)
Examination	MRI, CT, Blood Tests, Ultrasound etc	Y	- 6% to -15%
Surgical Treatment	Any surgery	Y	20% to 30%
Medical Consumables	Disposable medical supplies, high-value consumables	N	No change
Traditional Chinese Physical Therapy*	Supplementary Chinese treatment: Acupuncture, herbal treatment	Y	15%

Additional Tables

The additional tables below show the ZMP impacts on sub-category service utilization rates by disease groups across hospitals. Specifically, the tables below show the across-disease impacts of ZMP on patients' average medical examination expenses, inpatient and post-surgery care expenses, surgery expenses, length of stay, and surgery rates in the three hospitals, respectively.

Table 2.27: ZMP Impacts on Exam Expense by Disease Groups, Hospital B

Hospital B	(1) CD	(2) I	(3) K	(4) N	(5) S
ZMP	-228.936**	-558.167***	-204.935	-166.536	6.252
	(106.490)	(113.871)	(169.774)	(130.365)	(147.322)
N	3320	3133	1869	1085	1067
Mean	2546.5	3406.8	2856.1	1928.3	2069.5
month FE	Y	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y	Y
wave FE	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y

Table 2.26: ZMP Impacts on Exam Expense by Disease Groups, Hospital A

Hospital A	(1) CD	(2) I	(3) K	(4) M	(5) R
ZMP	-247.138	-397.096***	-375.530***	-186.268	-370.841**
	(154.200)	(90.388)	(133.415)	(167.225)	(173.285)
N	3244	6484	3235	1296	2001
Mean	3072.1	3465	2679.6	2579.2	2831.4
	(6) J	(7) E	(8) H	(9) N	
ZMP	-343.077**	-632.819***	-232.039	225.805	
	(153.500)	(167.160)	(188.847)	(157.675)	
N	2126	1468	1395	2014	
Mean	2443.5	3332	2486.2	2637.1	
month FE	Y	Y	Y	Y	
day-of-week FE	Y	Y	Y	Y	
wave FE	Y	Y	Y	Y	
Demographics	Y	Y	Y	Y	

Table 2.28: ZMP Impacts on Exam Expense by Disease Groups, Hospital C

Hospital C	(1) CD	(2) I	(3) K	(4) M
ZMP	-66.544	-93.780	-284.483**	-590.259***
	(262.210)	(131.805)	(136.907)	(101.413)
N	708	1578	1280	2071
Mean	2073.6	2610.1	2038.1	1976.2
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y

Table 2.29: ZMP Impacts on Inpatient and Post-Surgery Care Expense by Disease Groups, Hospital A

Hospital A	(1) CD	(2) I	(3) K	(4) M	(5) R
ZMP	132.209	588.561***	541.789***	252.581**	299.836**
	(142.563)	(167.330)	(122.884)	(99.743)	(133.834)
N	3244	6484	3235	1296	2001
Mean	1272.1	1356.4	1319.5	992.2	1050.4
	(6) J	(7) E	(8) H	(9) N	
ZMP	258.812	234.338**	205.313**	113.374	
	(278.317)	(101.856)	(86.610)	(129.719)	
N	2126	1468	1395	2014	
Mean	1849.8	1012.99	834.1	1182.2	
month FE	Y	Y	Y	Y	
day-of-week FE	Y	Y	Y	Y	
wave FE	Y	Y	Y	Y	
Demographics	Y	Y	Y	Y	

Table 2.30: ZMP Impacts on Inpatient and Post-Surgery Care Expense by Disease Groups, Hospital B

Hospital B	(1) CD	(2) I	(3) K	(4) N	(5) S
ZMP	308.326*	736.294***	161.160	201.863*	377.921***
	(162.084)	(263.155)	(134.766)	(120.424)	(120.701)
N	3320	3133	1869	1085	1067
Mean	1387	1339.4	1319.4	1094.2	1269.6
month FE	Y	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y	Y
wave FE	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y

Table 2.31: ZMP Impacts on Inpatient and Post-Surgery Care Expense by Disease Groups, Hospital C

Hospital C	(1) CD	(2) I	(3) K	(4) M
ZMP	-125.893	404.369***	66.122	-26.042
	(351.139)	(140.431)	(101.243)	(61.174)
N	708	1578	1280	2071
Mean	1855	1209.4	1124.8	1175.9
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y

Table 2.32: ZMP Impacts on Surgery Expense by Disease Groups, Hospital A

Hospital A	(1) CD	(2) I	(3) K	(4) M	(5) R
ZMP	288.297**	-153.537	219.908	155.838	42.168
	(146.754)	(93.867)	(137.153)	(361.468)	(136.852)
N	3244	6484	3235	1296	2001
Mean	730.3	283.4	1018.7	1125.8	593.1
	(6) J	(7) E	(8) H	(9) N	
ZMP	1.754	334.624*	37.193	357.399**	
	(116.102)	(174.443)	(140.918)	(147.639)	
N	2126	1468	1395	2014	
Mean	449.5	729.6	463.6	870.5	
month FE	Y	Y	Y	Y	
day-of-week FE	Y	Y	Y	Y	
wave FE	Y	Y	Y	Y	
Demographics	Y	Y	Y	Y	

Table 2.33: ZMP Impacts on Surgery Expense by Disease Groups, Hospital B

Hospital B	(1) CD	(2) I	(3) K	(4) N	(5) S
ZMP	-10.607	46.908	-26.451	313.804*	4.535
	(75.031)	(31.785)	(137.251)	(186.084)	(220.756)
N	3320	3133	1869	1085	1067
Mean	846.7	1156.5	1077.2	498.7	706.1
month FE	Y	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y	Y
wave FE	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y

Table 2.34: ZMP Impacts on Surgery Expense by Disease Groups, Hospital C

Hospital C	(1) CD	(2) I	(3) K	(4) M
ZMP	121.080	63.510	272.274***	-13.268
	(132.583)	(51.932)	(92.050)	(102.269)
N	708	1578	1280	2071
Mean	480	126.8	716.9	305.1
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y

Table 2.35: ZMP Impacts on Length of Stay by Disease Groups, Hospital A

Hospital A	(1) CD	(2) I	(3) K	(4) M	(5) R
ZMP	-0.532	-0.591	0.056	-1.661	0.095
	(0.802)	(0.553)	(0.689)	(1.053)	(0.684)
N	3244	6484	3235	1296	2001
Mean	11.9	11.5	11.2	12.2	9.9
	(6) J	(7) E	(8) H	(9) N	
ZMP	0.182	-0.577	-0.860	-1.208*	
	(0.961)	(0.750)	(0.694)	(0.704)	
N	2126	1468	1395	2014	
Mean	13.2	11.5	10	10.8	
month FE	Y	Y	Y	Y	
day-of-week FE	Y	Y	Y	Y	
wave FE	Y	Y	Y	Y	
Demographics	Y	Y	Y	Y	

Table 2.36: ZMP Impacts on Length of Stay by Disease Groups, Hospital B

Hospital B	(1) CD	(2) I	(3) K	(4) N	(5) S
ZMP	-0.669 (0.734)	0.277 (0.625)	-0.914 (0.688)	0.359 (0.781)	0.426 (0.893)
N	3320	3133	1869	1085	1067
Mean	12.1	12.2	11.3	9.6	12.1
month FE	Y	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y	Y
wave FE	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y

Table 2.37: ZMP Impacts on Length of Stay by Disease Groups, Hospital C

Hospital C	(1) CD	(2) I	(3) K	(4) M
ZMP	-5.514* (2.863)	0.453 (0.991)	-1.466 (1.113)	-0.852 (0.781)
N	708	1578	1280	2071
Mean	17.2	14.7	13.4	17.3
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y

Table 2.38: ZMP Impacts on Surgery Rate by Disease Groups, Hospital A

Hospital A	(1) CD	(2) I	(3) K	(4) M	(5) R
ZMP	0.046	0.004	0.052	0.101*	0.033
	(0.035)	(0.023)	(0.035)	(0.054)	(0.045)
N	3244	6484	3235	1296	2001
Mean	0.46	0.30	0.55	0.33	0.41
	J	E	H	N	
ZMP	0.026	0.084*	-0.016	0.014	
	(0.040)	(0.050)	(0.052)	(0.044)	
N	2126	1468	1395	2014	
Mean	0.34	0.35	0.35	0.44	
month FE	Y	Y	Y	Y	
day-of-week FE	Y	Y	Y	Y	
wave FE	Y	Y	Y	Y	
Demographics	Y	Y	Y	Y	

Table 2.40: ZMP Impacts on Surgery Rate by Disease Groups, Hospital C

Hospital C	(1) CD	(2) I	(3) K	(4) M
ZMP	0.033	-0.054	0.034	-0.051
	(0.077)	(0.043)	(0.051)	(0.042)
N	708	1578	1280	2071
Mean	0.47	0.25	0.71	0.32
month FE	Y	Y	Y	Y
day-of-week FE	Y	Y	Y	Y
wave FE	Y	Y	Y	Y
Demographics	Y	Y	Y	Y

Chapter 3

The Impacts of International Students on Local Product Markets: Evidence from College Towns

3.1 Introduction

The United States is the top host nation for international students, with more than 1.09 million students from abroad attending U.S. colleges and universities in 2018 – a notable increase from the 0.54 million recorded in 2000 (Institute of International Education, 2018). When international students study in the U.S., they contribute significantly to the economy by spending money on housing, food, and other living expenses, while also paying tuition to U.S. educational institutions.

The substantial financial impact of this influx of international students on the U.S. economy is widely acknowledged. As per the U.S. Department of Commerce's records, international students contributed an impressive $42.4 billion to the U.S. economy in 2017 through their expenditure on tuition, living costs, and various goods and services (Institute of International Education, 2018). This financial injection not only bolsters educational institutions but also provides support to local businesses and economies in the areas where these students reside. Employing a formula that translates exports into job creation, NAFSA estimated that foreign-born students were responsible for generating approximately 455,000 jobs during the 2017-2018 academic year (NAFSA International Student Economic Value Tool).[1]

While the impact of immigrants on the wages of domestic workers has been a subject of heated debate (e.g., Card, 1990; Bodvarsson, Van den Berg, and Lewer, 2008; Borjas, 2009), research on the effects of international student flows has received comparatively less attention. Despite student visa restrictions, particularly under the F-1 category, that limit employment opportunities for international students, their presence could play a crucial role in driving economic growth in various communities across the United States.

This paper examines the influence of changes in international student flows on local product markets, particularly in food-related industries. The underlying mechanism of how businesses respond to expanding local demand can be elucidated through a straightforward framework. Unlike the inflow of immigrants, which can impact both labor supply and demand, potentially complicating the job creation landscape, the influx of international students primarily stimulates increased demand for local non-traded goods, with the local food industry serving as a compelling example within

[1] http://www.nafsa.org/Policy_and_Advocacy/Policy_Resources/Policy_Trends_and_Data/NAFSA_International_Student_Economic_Value_Tool

this context. Consequently, the demand for labor in pertinent product markets shifts outward, offering potential economic opportunities for local businesses.

However, while the potential economic benefits of the additional demand from international students are evident, the precise quantitative contribution of this phenomenon remains an unanswered question. To address this gap, I leverage two primary data sources. First, I construct a city-year-level dataset of university student enrollments spanning the academic years from 2000 to 2019, sourced from the Integrated Postsecondary Education Data System (IPEDS). This dataset records student enrollment by domestic and international classifications for all degree-granting institutions in the U.S. The analysis specifically focuses on a sample of 272 U.S. cities commonly known as "college towns". These are places where the college population constitutes at least 10% of the city's total population and are distinct from larger metropolitan areas, thereby allowing for a plausible impact of international students on local product markets.Second, I integrate various Census data sources and pair city population estimates with local business patterns specific to these college towns.

I exploit the panel structure of the dataset, which captures the within-city, over-time variation in the proportion of international students relative to the city's total population and their resultant effects on local product markets. Focusing on within-city variations ensures that the estimates of the effects are not influenced by any fixed city characteristics underlying different college towns. Moreover, the analysis confines the sample to the pre-pandemic period. This approach ensures that the regression estimates primarily capture the demand effects of students who were physically present in the U.S., as enrollment records post-2019 might include students residing outside the country and engaging in online coursework. Additionally, the study controls for year-fixed effects to account for the common economic fluctuations affecting all college towns.

The identification of causal channels related to international students' impacts on local economies hinges on the critical assumption that variations in shocks to foreign student flows are unrelated to local economic factors that affect the product market trajectories. However, the disparate distribution of international students across different locations could be associated with city-specific characteristics that influence the evolution of local businesses over time. While part of these concerns is addressed by controlling for unchanging city attributes through the inclusion of city-fixed effects, there may exist some unobserved contemporaneous shocks that affect both the change of the international student body and the number of local food establishments.

To establish a plausible causality between the presence of international students and their impacts, I employ an instrumental variable (IV) strategy. The IV approach constructs the international student counts using the historical distribution of international students across college towns. It leverages the fact that previous city destinations tend to remain attractive to international students, which means that fluctuations in student inflows over time are not correlated with any contemporaneous factors.

The coefficient estimates from IV regressions provide compelling evidence that the arrival of international students contributes to the expansion of the food and beverage sectors. Over the period spanning 2000 to 2019, introducing an extra 1,000 international students into a city with a baseline population of 100,000 led to the emergence of 9 to 13 additional establishments and the creation of 50 to 70 more jobs within smaller-sized establishments. These outcomes are contingent on the presence of domestic students as well. Notably, the significant effects are concentrated within small-sized establishments that employ fewer than 19 individuals. Among these, the most substantial growth was observed in establishments with employee numbers ranging from 1 to 4, as well as those with employee numbers ranging from 5 to 9. On average, establishments falling into these two categories experienced a boost of 2 to 3 additional establishments when the international student population increased by 1,000 in a city with a baseline population of 100,000.

In addition to the potential direct effect on the size of the student population, there may be other reasons why the growth of international students can impact the local product market. Firstly, international students' diverse cultural backgrounds and living customs may lead to a specific demand for certain ethnic products, with restaurants being a prime example. Among the 1.09 million international students in 2017, approximately 75% originated from Asia. Notably, China and India accounted for the largest proportions of incoming students, constituting 36% and 19%, respectively. This significant presence of Asian students can lead to a substantial increase in demand for ethnic goods and services such as restaurants, beverage shops, and haircut services. Secondly, international students are generally ineligible for in-state tuition rates at public universities and are often excluded from financial aid programs. This distinction implies that international students may possess different and potentially higher purchasing power than their in-state counterparts. This increased purchasing power may not necessarily correlate with household income but reflect individual budget constraints. While we lack direct observations of disposable income for each student type, indirect measures of expenditure support this notion. For instance, there is a significant disparity in funding sources between domestic and international students. In the U.S., 70% of college students took out student loans in the class of 2018 with an average debt of $38,000, whereas 82.3% of foreign-born undergraduate students are supported entirely by family funding and are less likely to face a borrowing constraint.[2] These summaries align with other recent evidence that demonstrates positive fiscal externalities from exchange students (López, Fernández, and Incera, 2016) and rising housing prices in communities with expanding international student populations (Mocanu and Tremacoldi-Rossi, 2023).

The rest of the paper is organized as follows. Section 3.2 places this analysis in the context of other research efforts, highlighting the motivation and contribution of these papers. Section 3.3 provides an overview of the sample and a description of the data. The main empirical specifications and results will be discussed in section 3.4 and section 3.5. Section 3.6 concludes.

3.2 Research Context

This paper connects to research efforts in the economics of education and labor economics, which examine the domestic impacts of the flows of foreign-born students and high-skill workers to the U.S. This section begins by presenting the patterns and determinants of international student enrollment at U.S. colleges and universities, with the goal of establishing a plausible exogeneity of the national inflows of the international students and discussing their purchasing power. It then describes the strands of research literature that form the backdrop for this analysis.

[2] https://www.debt.org/students/

3.2.1 Flow of Students to the U.S. Higher Education Market

The trend of rising foreign inflows of students started in the 1950s, with accelerated growth since the 1970s. In the last 20 years, the total number of students from abroad at U.S. colleges and universities almost doubled, from 0.54 million in 2000 to a peak of 1.09 million in 2017 and a slight drop to 0.9 million in 2021. There has been a fluctuation in the growth rates since 2000. The total international enrollment number rose continuously until 2002 and dropped from 586,323 in 2002 to 564,766 in 2005. Since 2006, a consistently positive annual growth rate has been seen, with the peak arriving at a yearly growth rate of 10% in 2014. It has then stayed at more than 1 million since 2015.

Factors Driving the Inflows

The substantial increase in international student enrollment across various program types has generated interest in uncovering the driving forces behind this surge and its implications. Research suggests that the growth can be attributed, in part, to the attractiveness of the U.S. labor market to international students. In support of this idea, a migration model examined by Rosenzweig, Irwin, and Williamson (2006) finds a heightened influx of international students to the United States when the return on education is lower in their home countries. It suggests that students are more inclined to pursue educational opportunities in the U.S. when the potential benefits of obtaining an education are comparatively greater than in their countries of origin. Moreover, an analysis of exogenous policy changes concerning H-1B visas conducted by Shih (2016) reveals the impact of U.S. labor market openness on the enrollment of international students.

Apart from the appeal and the value associated with the U.S. labor market, Bound, Braga, et al. (2016) highlight three additional critical determinants of the supply of undergraduate international students. These factors include the capability to pay tuition, the demand for post-secondary studies, and the supply of higher education in the source country. These factors are closely related to the economic landscape and educational progress within the countries of origin for international students. To illustrate, they argue that the rising income and substantial improvements in secondary education in China have driven more than 90% of all foreign undergraduate increases from 2003 to 2013.[3]

While the overall inflow of international students to the United States is influenced by external determinants occurring outside the country, the specific distribution of international students within the United States may not be random and needs further investigation. Admittedly, certain universities, particularly those renowned for their focus on science and engineering disciplines, may be particularly attractive to international students who have been disproportionately distributed in these fields. However, changes in the inflow of international students to specific locations over time may be more closely related to demand-side shifts occurring in the students' source countries. As underscored by Bound, Turner, and Walsh (2009), individuals from countries with more limited higher education systems might be drawn to departments that are not necessarily ranked at the top. Nevertheless, the changing patterns of student inflows can also possess endogenous characteristics. Bound, Braga, et al.

[3]Some journalistic accounts additionally discussed the role of the elite selection system in China: The intensive competition of the national college entrance exam system may direct some students to opt out of this exam and choose other routes, e.g., study abroad. See LaFraniere, Sharon. 2009. "China's college entry test is an obsession." The New York Times, June 12, 2009. Heinz, Nicholas. 2018. "Failing Grade: How China's All-important Exam is Stunting National Growth" https://bpr.berkeley.edu/2018/12/07/failing-grade-how-chinas-all-important-exam-is-stunting-national-growth/

(2016) propose that public universities, especially those facing reductions in state funding, may have incentives to enroll international students, who, in turn, could generate tuition revenue to offset budget cuts. Consequently, such institutions have emerged as common destinations for international students.

A Growing Purchasing Power

Over the past two decades, a noticeable shift has occurred in the composition of international students studying in the United States. In earlier years, a smaller proportion of foreign-born students were enrolled in undergraduate programs compared to graduate programs. For instance, during the 2003/2004 academic year, only 37.2% of international students were pursuing associate's and bachelor's degrees, while 24.9% were engaged in master's programs, and 17.5% were enrolled in doctoral programs. However, this pattern began to transform in the 2014/2015 academic year when the percentage of students enrolled in undergraduate programs started to surpass that of graduate programs. In 2014/2015, 40.9% of foreign-born students were in undergraduate programs, while 37.2% were in graduate programs (with 20.5% in master's programs and 13% in doctoral programs).

The changing intake of international students over time implies a change in the purchasing power among international students in the U.S. These students typically hold F visas which come with employment restrictions and require them to demonstrate their ability to cover living expenses while studying. According to the 2018 statistics from the Institute of International Education (IIE), the primary source of funds for undergraduate international students is "self-support" or family support, which accounts for 82.3% of all funding sources. This indicates that the majority of undergraduate international students rely on personal or family funds to finance their education and living expenses. In contrast, among graduate students, the proportion relying on self or family support is lower, at 59.9%, and 34.6% of them report receiving funds from university fellowships and research support. Notably, none of the undergraduates and only 0.5% of graduates receive financial assistance from current employment.

3.2.2 Impacts of International Students and Workers on Domestic Labor Markets

This paper aims to isolate and quantify the demand contribution resulting from foreign inflows, with a particular focus on international students. Previous studies have investigated regional demand expansions driven by immigrant workers in the product market, but the presence of both supply and demand effects from immigrant workers has made it challenging to disentangle these channels. For instance, in the analysis of the Mariel Boatlift and the influx of Cuban immigrants, Bodvarsson, Van den Berg, and Lewer (2008) find that the increased labor supply was counteracted by a substantial simultaneous surge in labor demand, resulting in no discernible effect on native workers' wages, as documented by Card (1990). Hong and McLaren (2015) find a rise in local employment associated with immigrant workers, excluding the mechanical supply increase. This paper, however, takes a different approach by focusing on international students. Thus it aims to filter out the supply-side movement in the labor market while pinpointing only the demand-driven changes in the product market.

While a body of literature documents how immigrants benefit domestic labor markets, particularly those with high skills while imposing no harm on native employment (e.g., Kerr and Lincoln, 2010, Stuen, Mobarak, and Maskus, 2012, Peri, Shih, and Sparber, 2015), there is generally limited work on how international students impact the local economy. Considerable debates surround the consequences of an increasing proportion of foreign peers in higher education institutions on the well-being of native students.[4]

Empirical studies investigating the crowd-out effect on native academic outcomes and labor market responses have yielded mixed findings (e.g.,Orrenius and Zavodny, 2015, Ransom and Winters, 2016, Borjas, 2004). For instance, Bound, Braga, et al. (2016) find that the crowd-out on the enrollment of domestic students is unlikely to be one-to-one in some undergraduate programs: they find an association ratio of two international students to one in-state student. The non-existence of complete substitution within institutions implies the importance of also accounting for a change of external demands from domestic students when we attempt to interpret the causality of the international students on local demand expansion.

Recent studies have begun to address the gap in our understanding of the economic contributions of international students. López, Fernández, and Incera (2016) examine the impact of exchange students on the tourism industry in Galicia, Spain. They uncover a positive externality in the tourism sector, as the students contribute to tourism, and their presence leads to increased consumption by visiting family and friends. The expenditure multiplier for this "academic tourism" is 1.43, comparable to traditional inbound tourism. Peri, Shih, and Sparber (2015) highlight the potential

[4]The concerns that admitting more international students for funding issues harms native students are also often seen in journalistic reports: Jordan, Miriam. 2015. "International Students Stream Into U.S. Colleges." Wall Street Journal (March 24)
`http://www.wsj.com/articles/international-students-stream-into-u-s-colleges-1427248801`

> "Brenda Nard of Salem, Ore., said she encountered many out-of-state and international students during her daughter's recent college search. 'You wrestle with it because you want your kids to have the most opportunity, she said, 'I understand the state needs the money, yet I also wonder if it eliminates opportunities for some Oregonians.'"

loss of economic growth and agglomeration resulting from the failure to retain these students in the local economy after they complete their studies. In a related context, Mocanu and Tremacoldi-Rossi (2023) investigate the impact of international student enrollment on rental prices near university campuses.

This paper relates to this strand by examining how international students support local product markets. The presence of diverse tastes among the population can lead to horizontal differentiation, a concept reminiscent of the Hotelling model. The analysis centers its attention on the local food industry, which is closely linked to the ethnic preferences and demand patterns of the international student population. Previous empirical studies also lend support to the idea of immigration-induced product demand. Frattini (2008) for example, finds that the immigrant flows in the UK resulted in higher prices of low-value and everyday grocery goods, possibly related to a relatively larger share of such expenditure for immigrants. Similarly, Mazzolari and Neumark (2012) attribute the positive association between immigration and the diversity of restaurants to the comparative advantage of immigrants in producing food that caters to ethnic demand.

3.3 Data and Sample

Public data sets on higher education enrollment and local economic activity provide the basis for the analysis in this paper. The institution-level student enrollment statistics from the Integrated Postsecondary Education Data System (IPEDS); county- or equivalent city-level population estimates from the U.S. Census Bureau: Incorporated Places and Minor Civil Divisions Datasets, Sub-county Resident Population Estimates; and the annual statistics of the industry-level establishment at each size categories from U.S. Census ZIP Code Business Patterns.

3.3.1 Student Enrollment and City Population

I start by constructing the city-level (U.S. Census places with summary level code 162) counts of the entire college and foreign student groups, with data on the enrollment of students in each academic year from 2000 to 2019.[5][6] I focus on the pre-pandemic period so that the enrollment numbers do not include students who lived outside the U.S. and enrolled online. When there are multiple universities in a city, I compute the city-level counts of international students by aggregating over all universities. I then compare the total college student population to the city population estimates from the Bureau of the Census.

The concept of "college towns" does not have an official definition, but a widely accepted understanding is that they encompass cities where the presence of a college or university significantly influences the community's character and culture (Gumprecht, 2003). In this paper, I define cities as college towns if college students constitute a substantial proportion (at least 10%) of the city's total population. Similar to Mocanu and Tremacoldi-Rossi (2023), I only include the cities with a ratio of college students to city population higher than 10%, at least in one year in the sample period.[7] This selection rule leaves us with an analysis sample of 329 cities.

[5] The IPEDS dataset has a place identifier (city and state) associated with each institution. I use this identifier to match the city names in the Bureau of the Census dataset of population estimates.

[6] The official definition of the variable "non-residential alien" provided by the IPEDS is: "A person who is not a citizen or national of the United States and who is in this country on a visa or temporary basis and does not have the right to remain indefinitely." In this analysis, the total enrollment data contains all students (graduate and undergraduate, part-time and full-time) reported at schools. I choose degree-granting U.S. institutions in my sample.

[7] For example, Melbourn, Florida has a university population accounting for 5.85% of the total city population in 2000, but it meets the population criteria in 2009-2011 and is included in the sample.

I further exclude small towns or big cities whose population is below (or above) the bottom (or top) 10% tails of the distribution of these 329 cities in the base year of 2000. Consequently, cities with less than 8,500 or more than 139,100 residents in 2000 are not in the sample. As a result, big college cities such as Boston and Seattle are not in the final sample. In addition, I drop towns with a total number of international students of fewer than 100 in 2019 because it is less plausible for this amount of students to influence much on the local business.[8]According to the city population estimates in the baseline year, 2000, the smallest city in the final sample is Cheney, WA, and the largest city is Savanah, Georgia. The geographical distribution of college towns and the representation of international students within colleges are demonstrated in figures 3.1 and 3.2.

3.3.2 Census ZIP Codes Business Patterns

Total establishment numbers across sectors are collected from the Census Zip Codes Business Patterns (ZBP). This dataset, which starts from 1994, provides information about the number of establishments during the week of March 12 each year. The data covers a range of business activities and is particularly valuable for analyzing economic activity at a small area level (United Census Bureau, 2018).[9] For each

[8]This is because a single criterion of student proportion threshold inherently includes numerous small towns so that the number of establishments does not change very much, whereas there could be a considerable variation in the foreign student size. For example, a small town such as Longmeadow in Massachusetts (with a city population of around 15,700 and a university student population of approximately 2,000) would have been included based on the population criterion. However, a high variation of international student numbers can also be seen in this town: the total number of international students decreased from 24 in 2003 to 6 in 2012 and then rose to 54 in 2015. Including towns like these would cause imprecise estimates of the effects. The magnitudes of coefficient estimates do not change much using the sample including these cities.

[9]ZBP reports business activities in March, whereas IPEDS presents the Fall enrollment in each academic year. My empirical model is specified to ensure that responses from enterprises are captured over the same period of foreign student net inflows.

NAICS industry category within each Zip code, the ZBP dataset includes data on the total number of establishments and the number of establishments categorized by employment size. Similar to the methodology used for compiling university and international student population estimates, city-level establishment estimates in the sectors of interest are constructed by aggregating the Zip code counts. Given the possibility of new establishments being situated in the same city but in different zip codes, I calculate business counts at a broader geographic level (city) rather than attempting to match Zip-level establishments to the specific Zip codes associated with institutions in the IPEDS data.

3.3.3 Sample Description

Table 3.1 presents the summary statistics of the key variables in the final sample of 272 cities.[10] The cities in the sample have an average population of 43,122 in the year 2000, with notable variation across city population estimates. Figure 3.3 plots the distribution of the city population estimates in 2000, revealing that the majority of the college towns in the sample had an initial population of less than 50,000 in 2000. Given the diverse range of city populations, it is crucial to account for city size in our estimation. Over the period from 2000 to 2019, the average number of international students nearly doubled based on the mean of our pooled sample. Figures 3.1 and 3.2 visually represent the changes in foreign student representation and the total number of college students across the 272 college towns between 2000

[10]Note that the values of the variable "university population to city population" may be higher than one. It may be because the Census estimates do not include all of the students who are temporarily residing in the city. Although very few cities have this issue, it implies that the population estimates in the Census could be under-reported. To avoid measurement error, I keep the samples of cities with a student share of less than 1.5, and the coefficient estimate is robust as only a few cities have this issue.

and 2019. Color intensity reflects shifts in foreign student representation, while circle size indicates the relative magnitude of the total university population within each college town. The figures effectively highlight the substantial variations in foreign student representation both across college towns and over the two decades.

The progression of the total university population and its proportion relative to the city population can be observed in Figure 3.4, where both series exhibit gradual growth throughout the entire sample period. Figure 3.5 displays the average over-time pattern of the student representations within universities in the sample of 272 cities. The line in Figure 3.5 illustrates the pooled sample average of international student representation, which increased from below 4% in 2000 to almost 7% in 2015. Similar to the national-level trend in international enrollment, there has been a decline in the average share of international students within colleges since 2002. However, this ratio shows sustained growth from 2005 onwards, with accelerated expansion after 2010. The bars in figure 3.5 compare the counts of domestic students to the total student population and displays their patterns over time. Overall, there are steady growth in the university population and the domestic students, implying that the growing share of international students represents an expansion of both international inflows and the total student body rather than a substitution of domestic students within the college population.

Figure 3.6 describes how the establishment number across different size categories changes between 2000 and 2020. All statistics of establishments presented in these two figures are from the Census' Food and Drinking Places, NAICS code 722. As larger cities may have more establishments, all the number counts are normalized by the city's baseline population in 2000. The vertical axis shows the number of establishments per 1,000 residents. For example, the blue dot on the top in the year 2000 represents that there are, on average, eight food establishments (lower than four employees) per 10,000 residents.[11] Additionally, small- to medium-sized establishments dominate the food and service industry. The most common type of food and drinking establishment is the one with the smallest size, with a total employee number below 4. The second popular type is the one with 20 - 49 employees. It is also the type that has seen the quickest growth rate from 2000 to 2020.

3.4 Empirical Strategy

This subsection outlines the empirical methodology employed to estimate the causal relationship between international students and local business developments. The hypothesis posits that the expanded customer base of international students should amplify product demand, subsequently stimulating labor demand within non-traded sectors. In this study, the investigation is concentrated on a specific local product market—the food and beverage sector—owing to its close association with the consumption patterns of international students.

The equation used to estimate the effects of international student inflows is specified as follows:

$$\frac{y_{c,t}}{Population_{2000}} = \alpha + \beta_1 \frac{InternationalStudent_{c,t}}{Population_{2000}} + \beta_2 X_{c,t} + \gamma_t + \mu_c + \epsilon_{ct} \qquad (3.1)$$

Here, the notation c represents college towns, and t indicates the year. The two main outcome variables of y_{ct} are the establishment numbers, $Estab_{c,t}$, and total employment level, $Emp_{c,t}$. Specifically, the measure on $Estab_{c,t}$ includes the total

[11] As most of the college towns in the final sample do not have establishments with an employee number of more than 250, I do not include the large establishments in the graph.

number of food and drinking establishments and the detailed establishment numbers across size categories. $Emp_{c,t}$ measures the employment in small-sized establishments, with details explained in the later section. The key independent variable of interest is $InternationalStudent_{c,t}/Population_{c,2000}$, which measures the total enrollment of international students located at a city c relative to the baseline city population estimates in the year of 2000.[12]

To mitigate potential biases and facilitate meaningful comparisons, I standardize all level measures of both dependent and independent variables by the city's initial population in the year 2000, aligning with common practice in studies examining immigrant inflows (e.g., Card, 2001; Peri and Sparber, 2011; and Hong and McLaren, 2015). This normalization serves several purposes. Firstly, it addresses the potential influence of larger cities attracting a greater number of international students, and concurrently, the size of a city often correlates with business growth. Without this adjustment, a direct association between foreign student inflows and local business expansion might lead to an upward bias. Secondly, as highlighted by these papers, is to address the presence of heteroskedasticity. In this case, normalizing both the dependent and independent variables by the initial population size is equivalent to employing a weighted least squares estimator, where heteroscedasticity is modeled as a function of the square of population size Wozniak and Murray (2012).[13]

The identification of causal channels from international students hinges on the

[12]Recall that IPEDS gives us the annual Fall enrollment counts whereas the ZBP reports the economic indicator in early March. Therefore, the annual change of establishment numbers between 2005 March and 2006 March is matched with student enrollment data from the 2004 Fall to 2005 Fall. My sample uses the ZBP reports from 2001 March to 2016 March and IPEDS from 2000 Fall to 2015 Fall.

[13]I test this assumption by first obtaining the residuals from the regression simply between the level (without normalization) of the outcome variables and student enrollment. The square terms of the residuals are then regressed on the city's initial population and its quadratic term. The coefficient significance of the latter variable indicates the need to correct for heteroscedasticity.

crucial assumption that variations in shocks related to foreign student inflows are independent of other factors influencing the growth of establishments. On the one hand, the inflow of students into the U.S. is influenced by various factors such as the attractiveness of the U.S. labor market, the financial capacity of households from abroad to pay tuition, and the availability of higher education in the countries of origin, as discussed in section 3.2. Consequently, the correlation between these determinants of the national level of international enrollment and the local economic performance is relatively less concerning.

On the other hand, the distribution of international students may not occur randomly across different places, which poses challenges for identification in the analysis. The equilibrium of foreign student enrollment could be influenced by both the location preferences of international students (demand endogeneity) and the enrollment policies of universities (supply endogeneity). The term "demand endogeneity" refers to potential unobserved factors (e.g., expectations of better economic growth) that attract international students to certain locations and contribute to the local growth perspective. Conversely, the "supply endogeneity" problem arises when a university's decision to enroll more international students is linked to contemporaneous local economic conditions (Bound, Braga, et al., 2016). Consequently, a simple ordinary least square (OLS) estimation could be prone to positive or negative bias due to uncertainties in the correlation between unobserved local characteristics and the relative size of international student enrollment.

To address the potential endogeneity problem, I propose an approach inspired by a widely-cited instrument based on historical settlement patterns. The goal is to isolate the exogenous variation of local demand from international students, ensuring that the inflows are not correlated with contemporaneous factors driving local business growth. Existing immigration literature highlights the instrument's strength, relying on the assumption that new immigrants tend to locate in areas where previous immigrants from the same countries have settled. This similarity in school choices may also be expected from international students who highly value such social networks. Moreover, the historical enrollment of international students holds significance as it may reflect anticipated admission difficulty for many international students. The exclusion assumption is that past city destinations remain attractive to international students. Consequently, fluctuations in student inflows over time are not correlated with any contemporaneous factors that might influence the development of local food and beverage establishments. Although the assumption may be strong, it remains plausible within our specific context.[14]

The instrument that is used in the current analysis to predict international student enrollment at period t is:

$$Intern\widehat{ational}Student_{c,t} = \frac{InternationalStudent_{c,t_0}}{InternationalStudent_{all,t_0}} * InternationalStudent_{all,t} \tag{3.2}$$

Here, $\frac{InternationalStudentc,t_0}{InternationalStudent_{all,t_0}}$ is the number of all international students who were enrolled at colleges in location c, relative to the total number of college students of all cities in my sample, at a reference time t_0. The reference year t_0 is set as 1995 in this analysis.

[14]One potential challenge to the identification strategy is the presence of local economic factors that historically attracted a larger number of international students, impacting the overall levels and the upward trend of the local food and beverage industry. While it is true that international students who chose to remain in the U.S. after graduation may have also considered their employment prospects and the amenities offered by the city's local market, their career choices were primarily influenced by the national economic outlook and employment opportunities. Moreover, it is important to note that we imposed a few restrictions on our college town sample, excluding cities with an initial population of more than 139,100. This particular subset of cities mitigated the likelihood of admission decisions being influenced by local economic conditions that concurrently affected the growth rates of food industries.

The first term is where the cross-sectional variation of the instrumental variable comes from. In contrast, the second term of the instrumental variable captures the cumulative count of all foreign-born students within the sample from 2000 to 2019. This component drives the time-series variation of local-level international enrollments. As discussed previously, the overall number of students studying in the U.S. at the national level is considered to be exogenous to the local factors affecting business patterns in college towns.

Figure 3.7 plots the relationship between the distribution of international students across cities during the later period between 2000 and 2019 and the historical distribution in the year 1995 (x-axis). The y-axis is the average ratio of the international student number in city c to the total number of international students present in all cities in the sample during the 19 years. It serves to illustrate the IV strength: a steep and upward-sloping pattern indicates that cities with a higher concentration of international students in 1995 also tend to exhibit a greater presence of international students in subsequent periods.

I additionally control for the normalized number of domestic students, as captured by the variable $DomesticStudent_{c,t}/Pop_{c,2000}$, or the normalized number of university students, $UniversityStudent_{c,t}/Pop_{c,2000}$ in in $X_{c,t}$. It is to control for any potential impact brought by an expanding size of domestic students.[15]Controlling for the total university population enables an estimation of the contribution from the composition change within a university. The model further incorporates the city- and time-fixed effects in Equation 3.1 to capture the unique characteristics of cities

[15]Note the scope of this analysis is not to make an acute comparison between the consumption contribution from domestic and international students. It is because the location choices of domestic students could also be endogenous, and one cannot interpret β_2 as any causal relationship. The underlying assumption of the IV specification is that the total number of international students enrolled at a city c is uncorrelated with unobserved factors affecting the establishment numbers in the same city, conditional on the total enrollment of domestic students.

and the national economic impact, such as recession that might concurrently affect local economic performance. The estimation employs a first-differenced panel model. As with the unit fixed-effect model, the first difference exploits only the within-city variation between cities. The estimator is consistent under a weaker assumption than that for the fixed-effect model. Standard errors are clustered at city levels to account for potential autocorrelation over time.

3.5 Results

This section presents the results for the regression equation 3.1 on the impacts on the local food and beverage industry.

Tables 3.2 and 3.3 report the regression results of the equation by exploiting the high-frequency (yearly) changes of international students and outcome variables during the period between the academic school years 2000 and 2019. Table 3.5 shows the estimation results of using the low-frequency data that exploits the longer-span variation (i.e., school years 2000, 2005, 2010, 2015), and to avoid short-term fluctuations.

The Tables 3.2 and 3.3 collectively offer insights into the relationship between international student dynamics and outcomes of interests. Specifically, the dependent variable of the first two columns is the total number of food establishments in each city, the main outcome of interest. Columns 3 and 4 measure the effect on the employment level in small-sized establishments (with the total number of employees below 19). The outcome variables in Columns 5 to 10 are the specific number of establishments across different size categories. The two numbers shown in each column head represent the two endpoints of a size range: for example, the outcome variable in Columns 5 to 6 measures the total number of establishments with a total employee number between 1 and 4. The dependent variables in table 3.3 are the numbers of medium-sized establishments with total employee numbers ranging from 20 to 249. In all the tables below, odd-numbered columns are the OLS estimates, and even-numbered columns are the IV estimates. The last row in the tables presents reasonably strong first-stage F statistics.

Table 3.2 presents estimates of equation 3.1, with the normalized number of domestic college students included as a control. The OLS results are generally positive but statistically insignificant. As discussed, the estimates are likely biased downward because of the supply-side endogeneity of international student flows. Comparing the IV with OLS estimations, all coefficient estimates become larger and more significant, indicating a potentially downward bias from the OLS estimation.

The coefficient of IV estimation in table 3.2 rises from 0.000 to 0.009 when the dependent variable is the total number of food and drinking establishments (Columns 1 to 2). The IV estimate in Column 2 shows that adding international students that account for a one-percent rise relative to the initial city population creates 0.009 percentage points more establishments of the initial city population after we control for the size of domestic students. Given that the average number of international students in the sample increased from 452 in 2000 to 888 in 2019, and the sample average of the baseline population is 43,122, this coefficient translates to an average increase of 4 food and drinking establishments from 436 more international students into a city with an initial size of 43,122.

After examining the growth patterns across different sizes of establishments in Table 3.2, it becomes evident that the business type benefiting the most is the ones with a total employee count between 1 and 9 (Columns 6 and 8). Within this size category, a net increase of 1,000 international students relative to an initial city population of 100,000 results in four more establishments. The coefficient in Column 10 shows that the contribution effect is not statistically significant for relatively larger establishments with more than 10 employees on the payroll.

Table 3.3 evaluates the effects of international students on medium-sized establishments. None of the coefficients has a magnitude that is significantly different from 0. The results in this table demonstrate that international students only contribute to small-sized establishments with a total employee size below 19.

I further evaluate the impacts on the employment level in the food and service industry. Since the Census ZBP dataset does not reveal a continuous measure of employee numbers, I recode the binned variables of establishment size categories using the mid-point value of each range.[16] The employment proxy is computed only across small establishments with total employees below 19 because these places are seeing significant effects on the extensive margin. Furthermore, the approximation at only small-sized establishments raises less concern about measurement errors. The IV estimates of the impact of international students on the employment level in small-sized establishments are shown in Column 4 of table 3.2. The magnitude of 0.05 indicates that an average of 50 more jobs are created in response to 452 more international students in a city with an average initial population of 43,122.

Table 3.4 examines the same dependent variables as Table 3.2, but in this case, it includes the total number of college students as a control. Consequently, the coefficient estimates in this table demonstrate the effect of altering the composition of university students while controlling for the size of university students. The magni-

[16]The census business pattern data includes the number of establishments, employment during the week of March 12, first quarter payroll, and annual payroll. Therefore, part-time employees are also included if they are on the payroll.

tudes and significance of the coefficient estimates are similar, albeit with a slightly larger magnitude compared to Table 3.2. These results suggest that factors beyond pure size expansion might influence the changes in the local non-traded sector, including culinary industries.

The results presented thus far are based on estimations using annual (high-frequency) data. While high-frequency data offers valuable insights, one potential drawback is the increased possibility of measurement errors. To address this concern, I perform a robustness check by focusing on international student variation over a relatively long period, which may better capture more salient variations and enhance accuracy. In this robustness check, only the five-year changes (specifically, we keep school year observations in 2000, 2005, 2010, and 2015) are retained for each panel. The estimation results are displayed in Table 3.5. Encouragingly, I find consistently positive and significant effects similar to those in Table 3.2. The instrumental variable (IV) estimate in Column 2 exhibits a slightly larger magnitude than the one in Table 3.2, and it remains statistically significant.

To further investigate the extent of adjustment of international flows on local business activities, I include the changes of scaled international students from the academic year before and after as additional explanatory variables. By including the previous inflow, the examination extends to observe the speed and extent of the adjustment process. Additionally, the future inflows serve as proxies for expectation effects, revealing whether businesses anticipate changes in customer demand. Similar to the key independent variable, I instrument these additional variables using the same shift-share instruments as discussed. The results are presented in Table 3.6. In summary, the changes in the local establishment number are mainly influenced by the changes in international students in the corresponding academic year, and the coefficients of the flows from other years are statistically insignificant. Given that the student enrollment data are recorded in the Fall of each academic year (t), and the establishment data measures activities in March of the following year (t+1), our main specification may have already accounted for a certain adjustment period of small business owners to accommodate a larger customer base resulting from changes in international student flows.

3.6 Conclusion

The recent expansion in the influx of international students has sparked many debates on potential displacement effects between domestic and international students. Nevertheless, there remains to be a gap in our understanding of the financial implications of international students on the local economy. This paper seeks to contribute to the analysis of international student impacts by expanding the range of examined outcomes, with a specific focus on the perspective of the local product market.

The paper provides support and quantifies the presence of local demand contributions from international students enrolled in U.S. universities. These students encounter constraints on their employment options, setting them apart from immigrant workers. This distinctive attribute enables the study to isolate and precisely measure demand-induced shifts in the product market.

The study targets a group of cities where college students constitute a significant share (10%) of the local population. By leveraging fluctuations in international student enrollments within these cities over two decades, I gauge the degree to which they contribute to the growth of local businesses in the product market. This contribution is likely propelled by an enlargement of the customer base and unique demand attributes. Through IV estimations, the results reveal the economic benefits that international college students bring to the local product market, with a notable influence primarily observed within local small businesses.

In conclusion, the outcomes presented in this paper support the notion of increased local demand. Nonetheless, one improvement upon the current analysis is to consider distinction between different university program types. As international student demographics and enrollment compositions shift over time, so too may their purchasing power. Future research could delve into differentiating between institutions and program types over time. Understanding how international students' preferences and spending habits vary based on their academic pursuits could enable a more nuanced analysis of their impact on local businesses.

3.7 Tables

Table 3.1: Summary Statistics of the City and College Population

		2000				2019			
VARIABLES	N	mean	sd	min	max	mean	sd	min	max
Population estimate	272	43,122	29,901	8,866	132,895	49,813	35,380	9,116	170,243
Number of universities	272	1.265	0.570	1	4	1.227	0.532	1	4
Number of international students	272	452.1	672.6	0	4,627	888.4	1,469	24	9,085
Number of univ. students	272	11,128	8,716	1,143	57,578	14,587	13,857	1,217	115,547
Univ. students to city population	272	0.318	0.244	0.0521	1.379	0.347	0.261	0.0460	1.267
International students to city population	272	0.0118	0.0168	0	0.161	0.0216	0.0458	0.000393	0.587

Note: The left (right) panel shows the summary statistics of key variables in 2000 (2019). *Data source: the IPEDS, the Census population estimate, and author's calculation.*

Table 3.2: Normalized International Students and Small-sized Establishments, High-frequency Variation

Variables	(1) estab	(2) estab	(3) emp (small-sized estab)	(4) emp (small-sized estab)	(5) 1_4	(6) 1_4	(7) 5_9	(8) 5_9	(9) 10_19	(10) 10_19
Normalized no. of international students	0.000	0.009***	0.002	0.050***	-0.000	0.002***	-0.001***	0.002***	0.000*	0.002
	(0.000)	(0.003)	(0.003)	(0.018)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Normalized no. of domestic students	0.000	0.000	0.002*	0.001	-0.000	-0.000	-0.000	-0.000	0.000*	0.000
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Number of college towns	272	272	272	272	272	272	272	272	272	272
city and year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
no. of endogenous variables		1		1		1		1		1
No. obs.	5136	5136	5136	5136	5136	5136	5136	5136	5136	5136
First stage: F statistics		34.889		34.889		34.889		34.889		34.889

Columns of the odd numbers present OLS results whereas columns of the even numbers show IV results. The dependent variable in columns 1 and 2 is the total number in the food and drinking establishment (NAICS 722), scaled by the city population in the year 2000. Columns 3 to 4 measures the change of employment (mid-point proxy) hired in the small-sized establishments (total employee number is less than 19). Columns 5 to 10 present the establishment's number across size categories. The two numbers in the column title present the establishment size (range of employee number) and Column 3 show the number of establishments that has 1 to 4 employees hired. The first row indicates the key variable, the number of nonresidential students, scaled by the 2000 city population. The second row uses the (scaled) total number of domestic students as the control during the same period. Standard errors clustered at the city level in parenthesis.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3.3: Normalized International Students and Medium-sized Establishments, High-frequency Variation

Variables	20_49	20_49	50_99	50_99	100_249	100_249
Normalized no. of international students	0.000	0.001	-0.000	0.000	0.000	0.000
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)
Normalized no. of domestic students	0.000	-0.000	0.000***	0.000**	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Estimation method	OLS	IV	OLS	IV	OLS	IV
Number of college towns	272	272	272	272	272	272
city and year FE	Y	Y	Y	Y	Y	Y
no. of endogenous variables		1		1		1
No. obs.	5136	5136	5136	5136	5136	5136
First stage F statistics		34.889		34.889		34.889

Columns of the odd numbers present OLS results, whereas columns of the even numbers show IV results. Columns 1 to 6 present the establishment's number across size categories. The two numbers in the column title present the establishment size (range of employee number) and Column 1 shows the number of establishments that has 20 to 49 employees hired. The first row indicates the key variable, the number of nonresidential students, scaled by the 2000 city population. The second row uses the (scaled) number of domestic students as the control during the same period. Standard errors clustered at the city level in parenthesis.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3.4: Normalized International Students and Small-sized Establishments, High-frequency Variation

Variables	(1) estab	(2) estab	(3) emp (small-sized estab)	(4) emp (small-sized estab)	(5) 1_4	(6) 1_4	(7) 5_9	(8) 5_9	(9) 10_19	(10) 10_19
Normalized no. of international students	0.000	0.011***	0.002	0.057***	-0.000	0.003**	-0.001***	0.003**	0.000*	0.002
	(0.000)	(0.003)	(0.003)	(0.021)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Normalized no. of university students	0.000	-0.001	0.002*	-0.004	-0.000	-0.000	-0.000	-0.000	0.000*	-0.000
	(0.000)	(0.001)	(0.001)	(0.005)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Number of college towns	272	272	272	272	272	272	272	272	272	272
city and year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
no. of endogenous variables		1		1		1		1		1
No. obs.	5136	5136	5136	5136	5136	5136	5136	5136	5136	5136
First stage F statistics		23.479		23.479		23.479		23.479		23.479

Columns of the odd numbers present OLS results, whereas columns of the even numbers show IV results. Columns 1 to 6 present the establishment's number across size categories. The two numbers in the column title present the establishment size (range of employee number) and Column 1 shows the number of establishments that has 20 to 49 employees hired. The first row indicates the key variable, the number of nonresidential students, scaled by the 2000 city population. The second row uses the (scaled) number of university students as the control during the same period. Standard errors clustered at the city level in parenthesis.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.5: Normalized International Students and Small-sized Establishments, Low-frequency Variation

Variables	(1) estab	(2) estab	(3) emp (small-sized estab)	(4) emp (small-sized estab)	(5) 1_4	(6) 1_4	(7) 5_9	(8) 5_9	(9) 10_19	(10) 10_19
Normalized no. of international students	0.001**	0.013***	0.009	0.070***	0.000	0.005***	-0.000	0.002	0.002	0.003**
	(0.000)	(0.004)	(0.012)	(0.021)	(0.001)	(0.002)	(0.000)	(0.002)	(0.001)	(0.001)
Normalized no. of domestic students	0.001**	0.001***	0.005***	0.005***	-0.000	-0.000	0.000	0.000	0.000***	0.000***
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Number of college towns	272	272	272	272	272	272	272	272	272	272
city and year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
no. of endogenous variables		1		1		1		1		1
No. obs.	815	815	815	815	815	815	815	815	815	815
First stage F statistics		17.75		17.75		17.75		17.75		17.75

Columns of the odd numbers present OLS results whereas columns of the even numbers show IV results. The dependent variable in Columns 1 and 2 is the total number in the food and drinking establishment (NAICS 722), scaled by the city population in the year 2000. Columns 3 to 4 measures the change of employment (mid-point proxy) hired in the small-sized establishments (total employee number is less than 19). Columns 5 to 10 present the establishment's number across size categories. The two numbers in the column title present the establishment size (range of employee number) and Column 3 shows the number of establishments that has 1 to 4 employees hired. The first row indicates the key variable, the number of nonresidential students, scaled by the 2000 city population. The second row uses the number of domestic students as control during the same period. Standard errors clustered at the city level in parenthesis.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3.6: Normalized International Students and Small-sized Establishments, Robustness Check

Variables	(1) estab	(2) estab	(3) emp (small-sized estab)	(4) emp (small-sized estab)	(5) 1_4	(6) 1_4	(7) 5_9	(8) 5_9	(9) 10_19	(10) 10_19
Normalized no. of international students	0.010	0.008**	0.083	0.025	0.006	0.002	0.005	0.002	0.002	0.000
	(0.007)	(0.003)	(0.058)	(0.035)	(0.005)	(0.003)	(0.004)	(0.001)	(0.004)	(0.002)
Normalized no. of international students (Lag)	-0.001		-0.046		-0.004		-0.004		-0.001	
	(0.007)		(0.059)		(0.005)		(0.004)		(0.004)	
Normalized no. of international students (Lead)		0.002		0.041		0.002		0.000		0.002
		(0.004)		(0.035)		(0.003)		(0.002)		(0.002)
Estimation method	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV
Number of college towns	272	272	272	272	272	272	272	272	272	272
city and year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
no. of endogenous variables	2	2	2	2	2	2	2	2	2	2
No. obs.	4862	4862	4862	4862	4862	4862	4862	4862	4862	4862
Partial First stage F statistics										
Normalized no. of international students	32.65	27.76	32.65	27.76	32.65	27.76	32.65	27.76	32.65	27.76
Normalized no. of international students (Lag)	30.64		30.64		30.64		30.64		30.64	
Normalized no. of international students (Lead)		23.93		23.93		23.93		23.93		23.93

Columns of the odd numbers present OLS results whereas columns of the even numbers show IV results. The dependent variable in Columns 1 and 2 is the total number in the food and drinking establishment (NAICS 722), scaled by the city population in the year 2000. Columns 3 to 4 measures the change of employment (mid-point proxy) hired in the small-sized establishments (total employee number is less than 19). Columns 5 to 10 present the establishment's number across size categories. The two numbers in the column title present the establishment size (range of employee number) and Column 3 shows the number of establishments that has 1 to 4 employees hired. The first row indicates the key variable, the number of nonresidential students, scaled by the 2000 city population. The second row indicates the the number of nonresidential students from previous school year, scaled by the 2000 city population. The third row indicates the the number of nonresidential students from the next school year, scaled by the 2000 city population. The first-stage partial F-statistics for each endogenous variable is reported in the last three rows. The models use the number of domestic students as control during the same period. Standard errors clustered at the city level in parenthesis.
*** p<0.01, ** p<0.05, * p<0.1

3.8 Figures

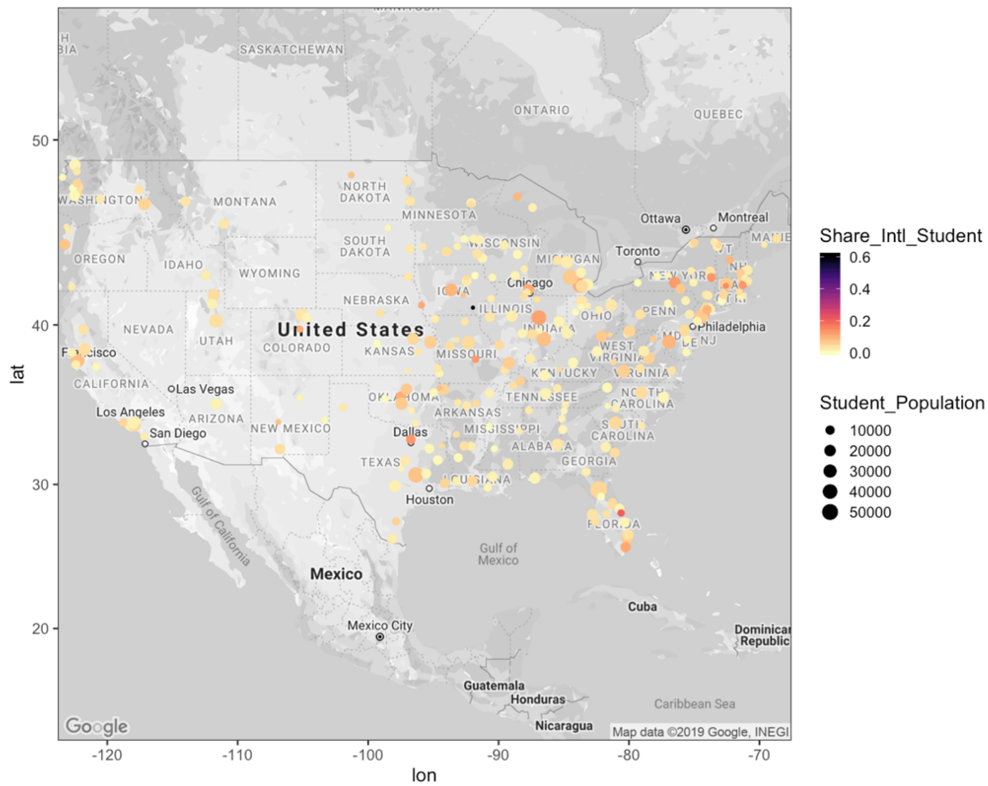

Figure 3.1: College Population and International Student Representation in 2000

Note: Circle sizes show the total number of college students at the city level. Color intensities measure the extent of city-level foreign student representation, total number of foreign students/total college students. *Data source: the IPEDS, the Census population estimate, and author's calculation.*

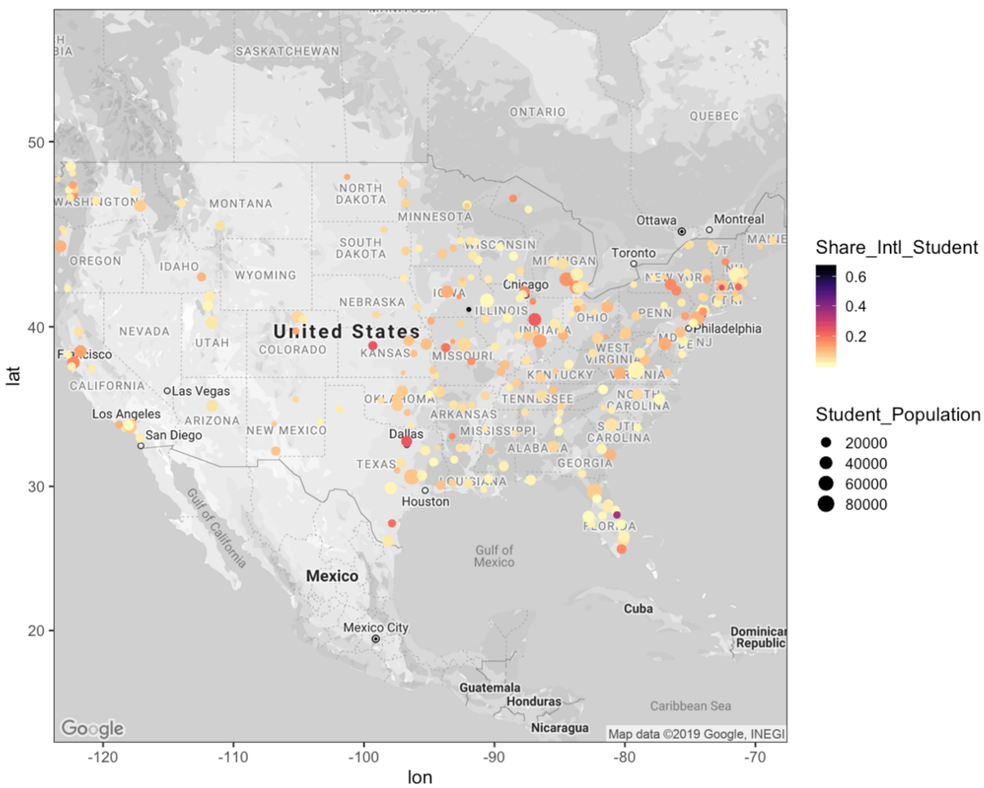

Figure 3.2: College Population and International Student Representation in 2019

Note: Circle sizes show the total number of college students at the city level. Color intensities measure the extent of city-level foreign student representation, total number of foreign students/total college students. *Data source: the IPEDS, the Census population estimate, and author's calculation.*

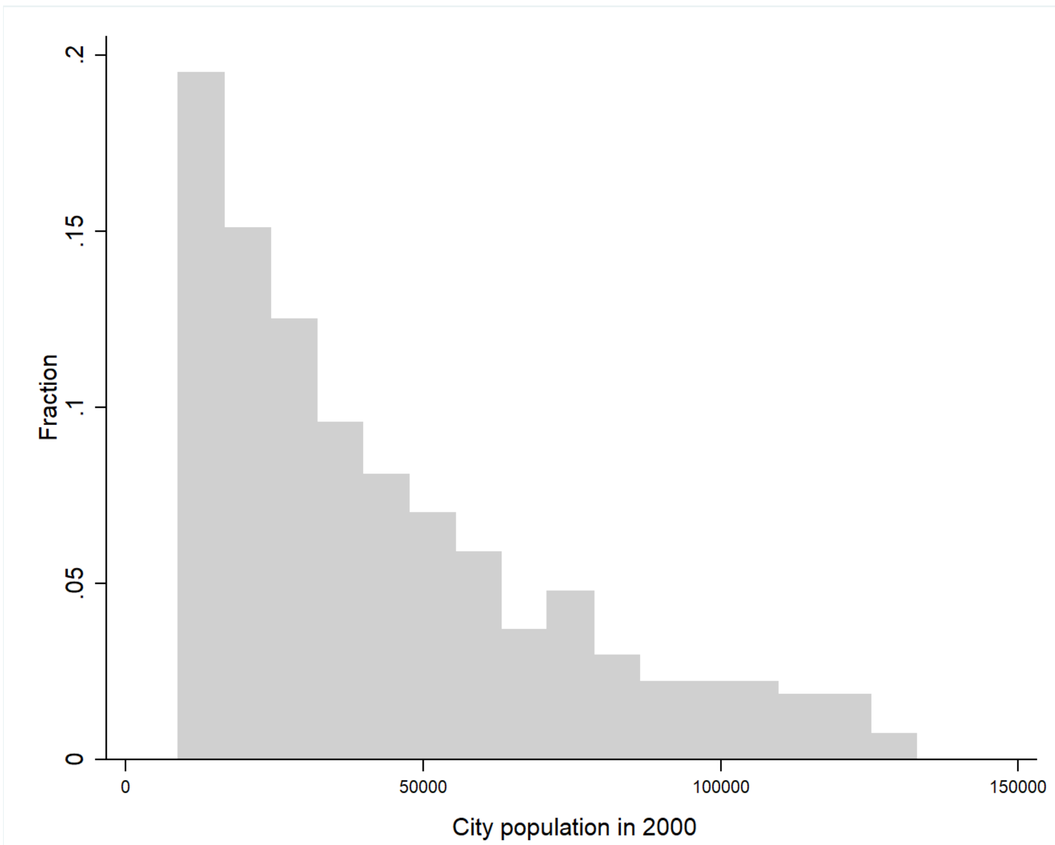

Figure 3.3: Sample Distribution of City Population in 2000

Note: The horizontal axis displays the bins of city population estimates in the year of 2000. The vertical axis shows the fraction of cities within each bin. *Data source: the Census population estimate, and author's tabulation.*

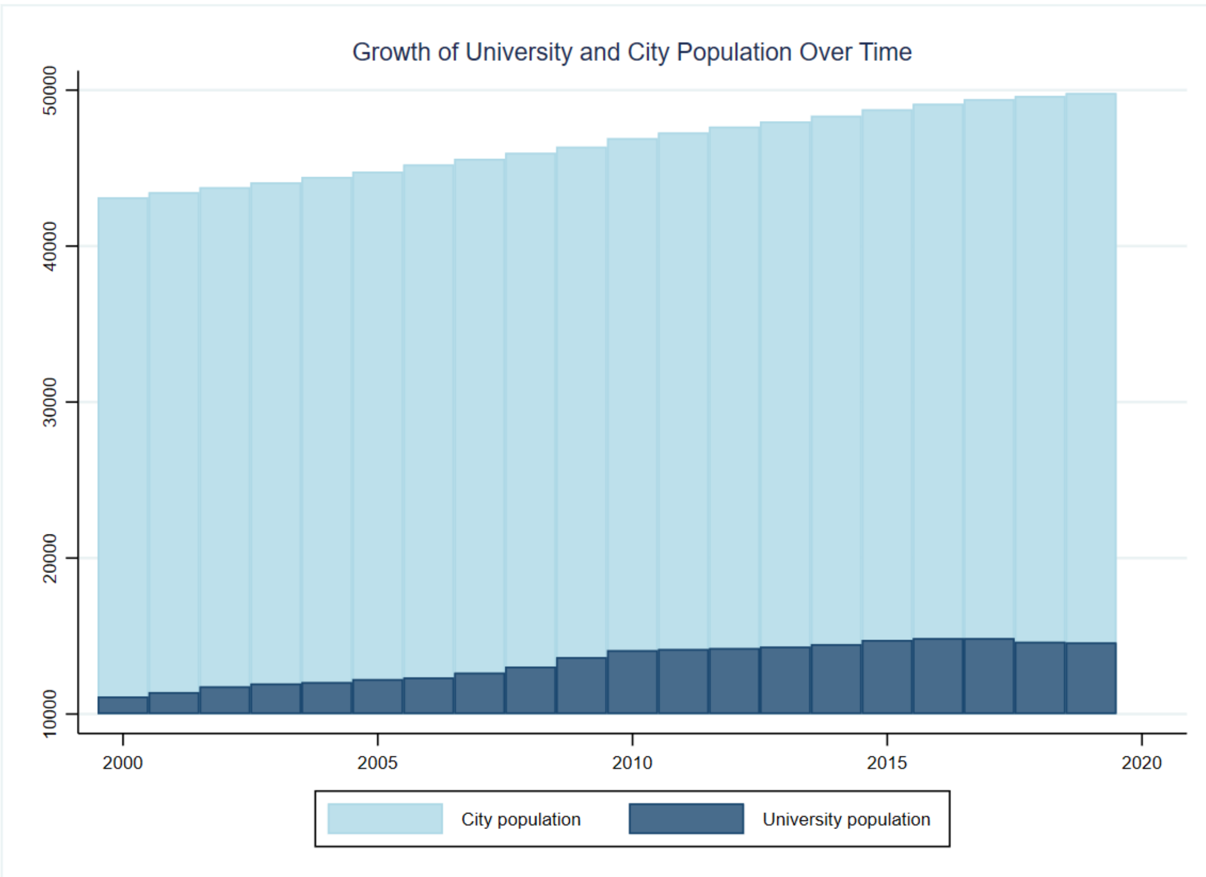

Figure 3.4: University and City Population, 2000 - 2019

Note: The vertical axis shows the average city population estimates and total university student counts at the city levels. *Data source: the IPEDS and the Census population estimate, and author's tabulation.*

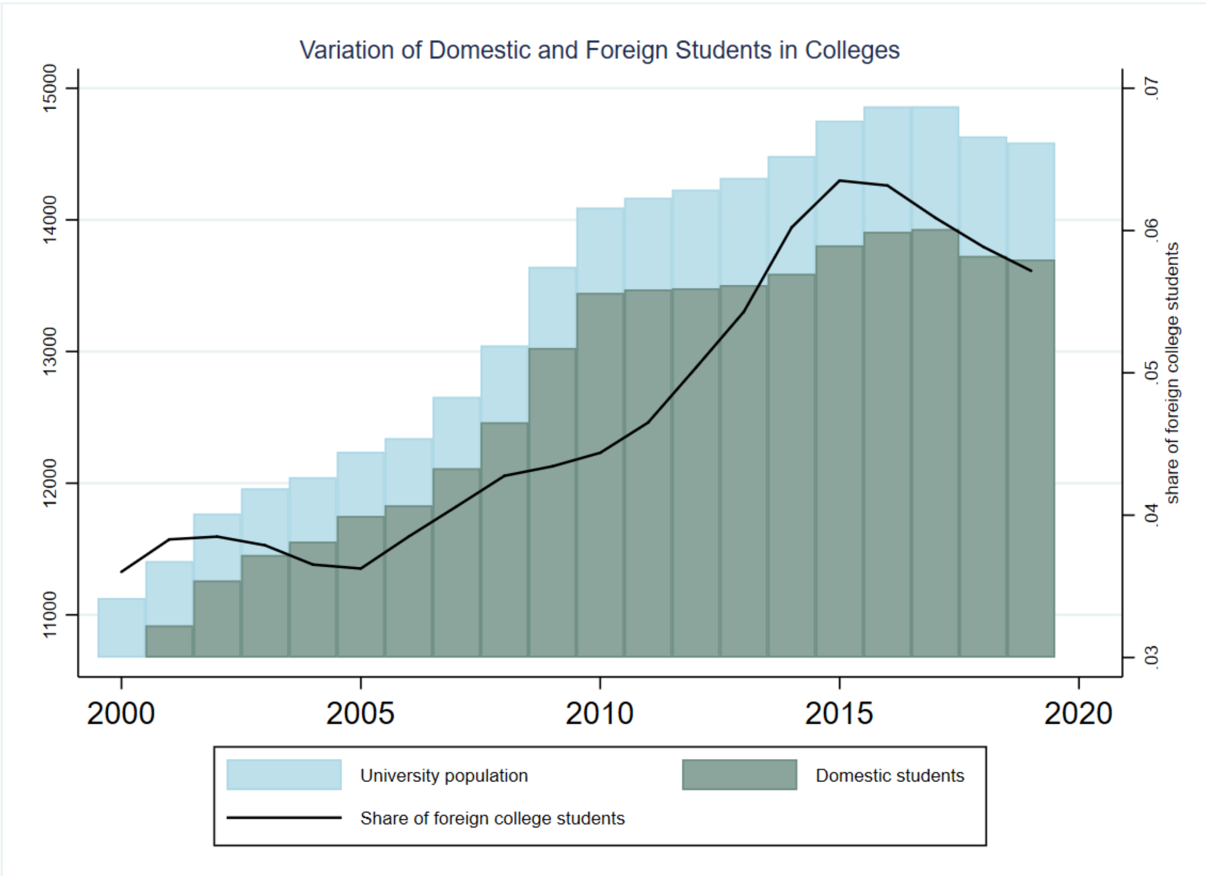

Figure 3.5: University Population and Student Representation, 2000 - 2019

Note: The left axis shows the student counts at the city level. The right axis displays the average share of average foreign students within a university and is the axis for the line graph. *Data source: the IPEDS and the Census population estimate, and author's calculation.*

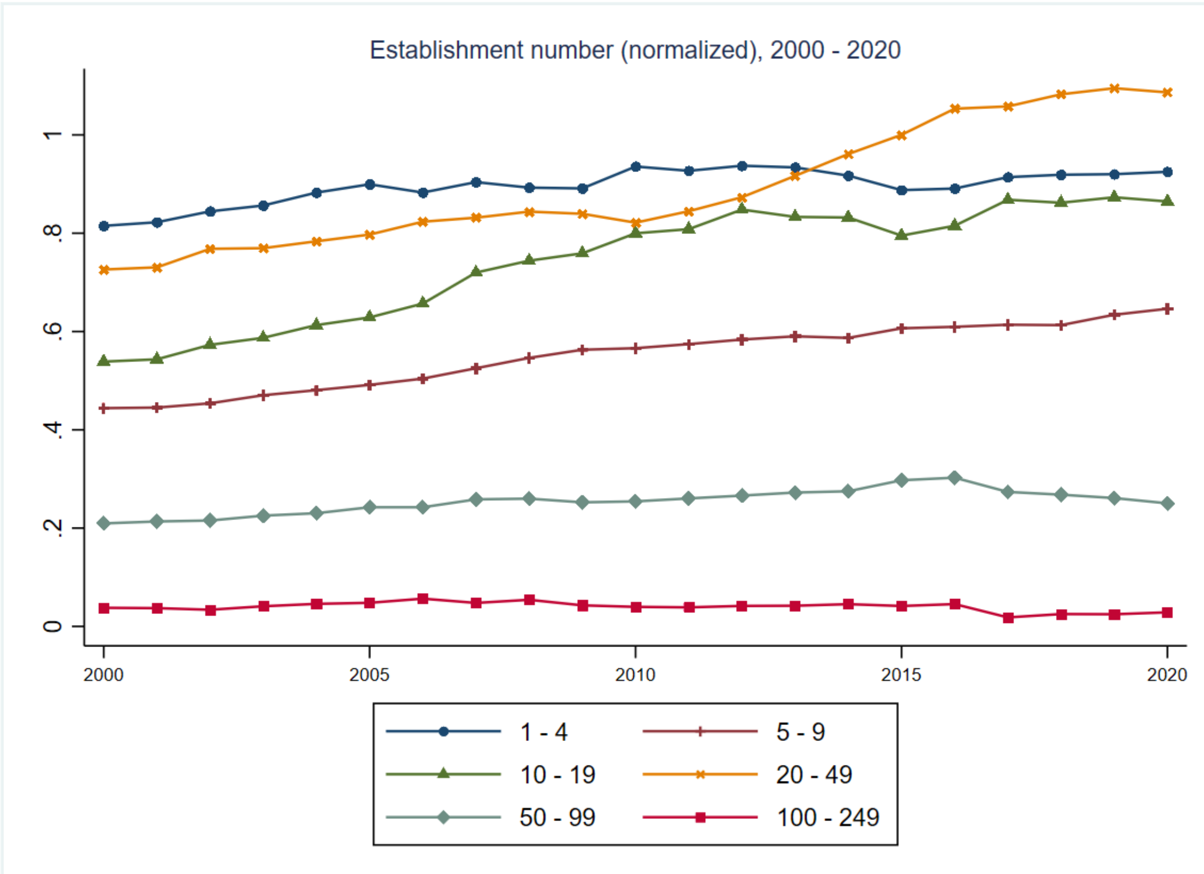

Figure 3.6: Average Normalized Number of Establishments Across Establishment Size Categories, 2000 - 2020

Note: Each line represents a specific size category of establishments, and two numbers in each category indicate the size range. The vertical axis measures the average number of establishments per 1,000 people. The establishment number is normalized by a city's baseline population in 2000. *Data source: Census ZIP Codes Business Patterns, the Census population estimate, and author's calculation.*

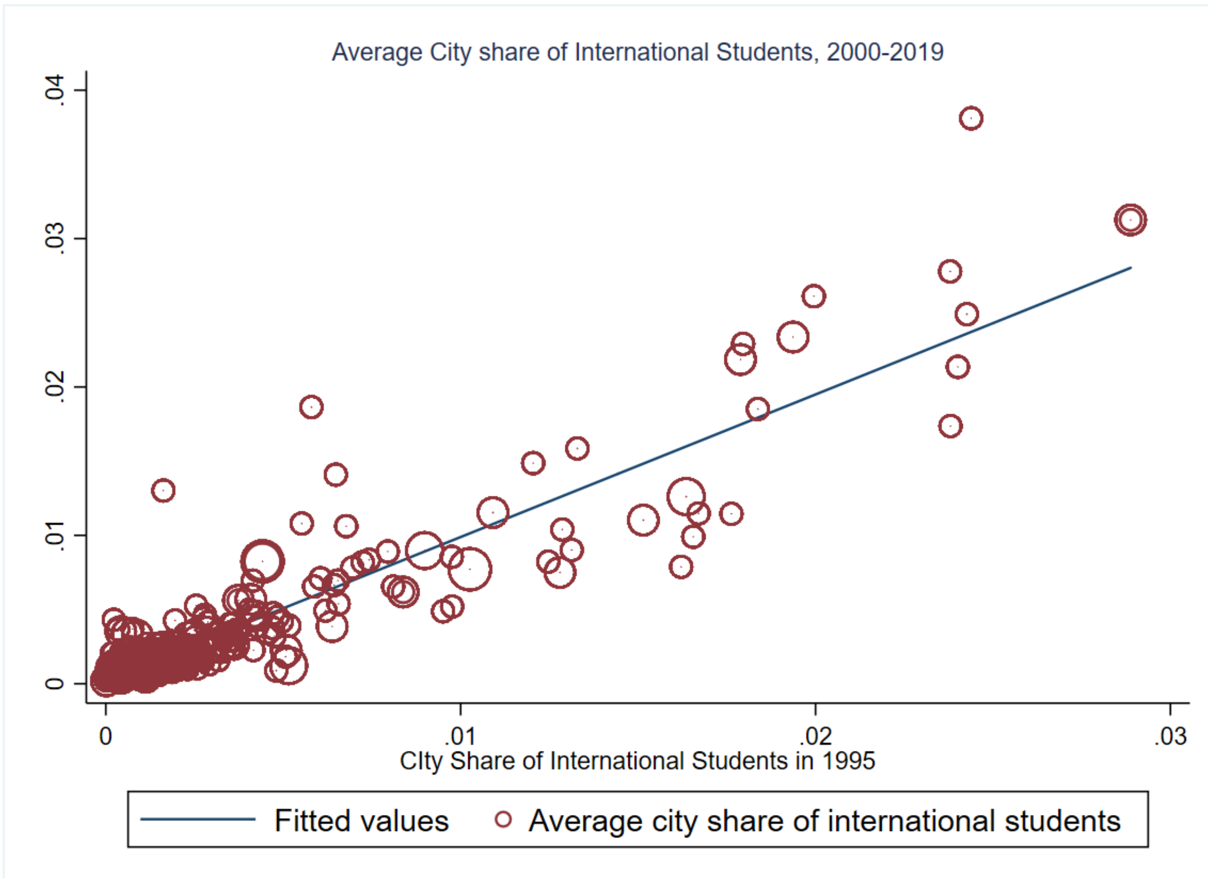

Figure 3.7: Average Share of International Students

Note: The horizontal axis shows the historical distribution of foreign students of all 272 cities in 1995. The vertical axis measures the average distribution of foreign students between 2000 and 2019. The size of hollow circles measures the number of institutions in a city. *Data source: the IPEDS, the Census population estimate, and author's calculation*

Bibliography

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program". In: *Journal of the American statistical Association* 105.490, pp. 493–505.

Acemoglu, Daron and David Autor (2011). "Skills, tasks and technologies: Implications for employment and earnings". In: *Handbook of labor economics*. Vol. 4. Elsevier, pp. 1043–1171.

Acemoglu, Daron and Pascual Restrepo (2020). "Robots and jobs: Evidence from US labor markets". In: *Journal of political economy* 128.6, pp. 2188–2244.

Alexander, Diane (2020). "How do doctors respond to incentives? unintended consequences of paying doctors to reduce costs". In: *Journal of Political Economy* 128.11, pp. 4046–4096.

Arrow, Kenneth J (1965). "Uncertainty and the welfare economics of medical care: reply (the implications of transaction costs and adjustment lags)". In: *The American economic review* 55.1/2, pp. 154–158.

Auerbach, Alan J, Yuriy Gorodnichenko, and Daniel Murphy (2019). *Local fiscal multipliers and fiscal spillovers in the united states*. Tech. rep. National Bureau of Economic Research.

Autor, David H, David Dorn, and Gordon H Hanson (2013). “The China syndrome: Local labor market effects of import competition in the United States”. In: *American economic review* 103.6, pp. 2121–2168.

Autor, David H and Mark G Duggan (2003). “The rise in the disability rolls and the decline in unemployment”. In: *The Quarterly Journal of Economics* 118.1, pp. 157–206.

Barro, Robert J and Charles J Redlick (2011). “Macroeconomic effects from government purchases and taxes”. In: *The Quarterly Journal of Economics* 126.1, pp. 51–102.

Black, Dan, Terra McKinnish, and Seth Sanders (2005). “The economic impact of the coal boom and bust”. In: *The Economic Journal* 115.503, pp. 449–476.

Blanchard, Olivier and Lawrence F Katz (1992). “YRegional Evolutions, *BrookingPapersonEconon*
In: *Economic Studies Program, The Brookings Institution* 23.1, p. 76.

Bodvarsson, Örn B, Hendrik F Van den Berg, and Joshua J Lewer (2008). “Measuring immigration’s effects on labor demand: A reexamination of the Mariel Boatlift”. In: *Labour Economics* 15.4, pp. 560–574.

Borjas, George J (2004). *Do foreign students crowd out native students from graduate programs?* Tech. rep. National Bureau of Economic Research.

— (2009). “Immigration in high-skill labor markets: The impact of foreign students on the earnings of doctorates”. In: *Science and engineering careers in the United States: An analysis of markets and employment*. University of Chicago Press, pp. 131–161.

Bound, John, Breno Braga, et al. (2016). *A passage to America: University funding and international students*. Tech. rep. National Bureau of Economic Research.

Bound, John, Sarah Turner, and Patrick Walsh (2009). *Internationalization of US doctorate education*. Tech. rep. National Bureau of Economic Research.

Brynjolfsson, Erik and Andrew McAfee (2012). *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*. Brynjolfsson and McAfee.

Burga, Ramiro and Sarah Turner (2023). "Does Enrollment Lead to Completion?: Measuring Adjustments in Education to Local Labor Market Shocks". In: *Journal of Human Resources* 58.3, pp. 755–782.

Card, David (1990). "The impact of the Mariel boatlift on the Miami labor market". In: *ILR Review* 43.2, pp. 245–257.

— (2001). "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration". In: *Journal of Labor Economics* 19.1, pp. 22–64.

Chan, Sewin and Ann Huff Stevens (2001). "Job loss and employment patterns of older workers". In: *Journal of labor economics* 19.2, pp. 484–521.

Chen, Chun et al. (2014). "Is the prescribing behavior of Chinese physicians driven by financial incentives?" In: *Social Science & Medicine* 120, pp. 40–48.

Cohen, Lauren, Joshua Coval, and Christopher Malloy (2011). "Do powerful politicians cause corporate downsizing?" In: *Journal of Political Economy* 119.6, pp. 1015–1060.

Coile, Courtney C and Phillip B Levine (2011). "Recessions, retirement, and social security". In: *American Economic Review* 101.3, pp. 23–28.

Conley, Timothy G and Bill Dupor (2013). "The American Recovery and Reinvestment Act: solely a government jobs program?" In: *Journal of monetary Economics* 60.5, pp. 535–549.

Currie, Janet, Wanchuan Lin, and Juanjuan Meng (2014). "Addressing antibiotic abuse in China: an experimental audit study". In: *Journal of development economics* 110, pp. 39–51.

Currie, Janet, Wanchuan Lin, and Wei Zhang (2011). "Patient knowledge and antibiotic abuse: Evidence from an audit study in China". In: *Journal of health economics* 30.5, pp. 933–949.

Dafny, Leemore S (2005). "How do hospitals respond to price changes?" In: *American Economic Review* 95.5, pp. 1525–1547.

Dao, Mai, Davide Furceri, and Prakash Loungani (2017). "Regional labor market adjustment in the United States: trend and cycle". In: *Review of Economics and Statistics* 99.2, pp. 243–257.

Demyanyk, Yuliya, Elena Loutskina, and Daniel Murphy (2019). "Fiscal stimulus and consumer debt". In: *Review of Economics and Statistics* 101.4, pp. 728–741.

Deryugina, Tatyana, Alexander MacKay, and Julian Reif (2020). "The long-run dynamics of electricity demand: Evidence from municipal aggregation". In: *American Economic Journal: Applied Economics* 12.1, pp. 86–114.

Dickens, William T et al. (2007). "How wages change: micro evidence from the International Wage Flexibility Project". In: *Journal of Economic Perspectives* 21.2, pp. 195–214.

Eckert, Fabian et al. (2020). *Imputing Missing Values in the US Census Bureau's County Business Patterns*. Tech. rep. National Bureau of Economic Research.

Fang, Hanming and Qing Gong (2017). "Detecting potential overbilling in medicare reimbursement via hours worked". In: *American Economic Review* 107.2, pp. 562–591.

Fang, Hanming, Xiaoyan Lei, et al. (2021). *Physician-induced demand: Evidence from China's drug price zero-markup policy*. Tech. rep. National Bureau of Economic Research.

Feyrer, James, Erin T Mansur, and Bruce Sacerdote (2017). "Geographic dispersion of economic shocks: Evidence from the fracking revolution". In: *American Economic Review* 107.4, pp. 1313–1334.

Feyrer, James and Bruce Sacerdote (2011). *Did the stimulus stimulate? Real time estimates of the effects of the American Recovery and Reinvestment Act*. Tech. rep. National Bureau of Economic Research.

Finkelstein, Amy, Sarah Taubman, Bill Wright, et al. (2012). "The Oregon health insurance experiment: evidence from the first year". In: *The Quarterly journal of economics* 127.3, pp. 1057–1106.

Finkelstein, Amy N, Sarah L Taubman, Heidi L Allen, et al. (2016). "Effect of Medicaid coverage on ED use—further evidence from Oregon's experiment". In.

Foote, Andrew, Michel Grosz, and Ann Stevens (2019). "Locate Your Nearest Exit: Mass Layoffs and Local Labor Market Response". In: *ILR Review* 72.1, pp. 101–126.

Frattini, Tommaso (2008). "Immigration and Prices in the UK". In: *University College London, mimeo.*

Fu, Hongqiao, Ling Li, and Winnie Yip (2018). "Intended and unintended impacts of price changes for drugs and medical services: evidence from China". In: *Social science & medicine* 211, pp. 114–122.

Ganong, Peter and Daniel Shoag (2017). "Why has regional income convergence in the US declined?" In: *Journal of Urban Economics* 102, pp. 76–90.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (2020). "Bartik instruments: What, when, why, and how". In: *American Economic Review* 110.8, pp. 2586–2624.

Gruber, Jon, John Kim, and Dina Mayzlin (1999). “Physician fees and procedure intensity: the case of cesarean delivery”. In: *Journal of health economics* 18.4, pp. 473–490.

Gruber, Jonathan and Maria Owings (1994). *Physician financial incentives and cesarean section delivery.*

Gumprecht, Blake (2003). “The American college town”. In: *Geographical Review* 93.1, pp. 51–80.

Hakobyan, Shushanik and John McLaren (2016). “Looking for local labor market effects of NAFTA”. In: *Review of Economics and Statistics* 98.4, pp. 728–741.

Hall, Robert E and Charles I Jones (2007). “The value of life and the rise in health spending”. In: *The quarterly journal of economics* 122.1, pp. 39–72.

Haltiwanger, John, Ron Jarmin, and Cornell John Krizan (2010). “Mom-and-pop meet big-box: complements or substitutes?” In: *Journal of Urban Economics* 67.1, pp. 116–134.

Hebous, Shafik (2016). *Can government demand stimulate private investment? Evidence from US Federal procurement.* International Monetary Fund.

Ho, Kate and Ariel Pakes (2014). “Hospital choices, hospital prices, and financial incentives to physicians”. In: *American Economic Review* 104.12, pp. 3841–3884.

Hong, Gihoon and John McLaren (2015). *Are Immigrants a Shot in the Arm for the Local Economy?* Tech. rep. National Bureau of Economic Research.

Hultquist, Andy and Tricia L Petras (2012). “An examination of the local economic impacts of military base closures”. In: *Economic Development Quarterly* 26.2, pp. 151–161.

Kerr, William R and William F Lincoln (2010). “The supply side of innovation: H-1B visa reforms and US ethnic invention”. In: *Journal of Labor Economics* 28.3, pp. 473–508.

Li, Yongbin et al. (2012). "Overprescribing in China, driven by financial incentives, results in very high use of antibiotics, injections, and corticosteroids". In: *Health affairs* 31.5, pp. 1075–1082.

López, Xesús Pereira, Melchor Fernández Fernández, and André Carrascal Incera (2016). "The economic impact of international students in a regional economy from a tourism perspective". In: *Tourism Economics* 22.1, pp. 125–140.

Lu, Fangwen (2014). "Insurance coverage and agency problems in doctor prescriptions: evidence from a field experiment in China". In: *Journal of Development Economics* 106, pp. 156–167.

Manson, Steven et al. (2017). "IPUMS National Historical Geographic Information System: Version 12.0 [Database]". In: *Minneapolis: University of Minnesota* 39.

Mazzolari, Francesca and David Neumark (2012). "Immigration and product diversity". In: *Journal of Population Economics* 25.3, pp. 1107–1137.

Miller, Amalia R, Carmit Segal, and Melissa K Spencer (2022). "Effects of COVID-19 shutdowns on domestic violence in US cities". In: *Journal of urban economics* 131, p. 103476.

Mocanu, Tatiana and Pedro Tremacoldi-Rossi (2023). "The impact of international students on housing markets". In: *Canadian Journal of Economics/Revue canadienne d'économique* 56.2, pp. 647–675.

Molloy, Raven, Christopher L Smith, and Abigail Wozniak (2011). "Internal migration in the United States". In: *Journal of Economic perspectives* 25.3, pp. 173–96.

Moretti, Enrico (2011). "Local labor markets". In: *Handbook of labor economics*. Vol. 4. Elsevier, pp. 1237–1313.

Nakamura, Emi and Jón Steinsson (2014). "Fiscal stimulus in a monetary union: Evidence from US regions". In: *American Economic Review* 104.3, pp. 753–792.

Neumark, David, Junfu Zhang, and Stephen Ciccarella (2008). "The effects of Wal-Mart on local labor markets". In: *Journal of Urban Economics* 63.2, pp. 405–430.

Newhouse, Joseph P., Rand Corporation. Insurance Experiment Group, and Insurance Experiment Group Staff (1993). *Free for all?: lessons from the RAND health insurance experiment*. Harvard University Press.

Olea, José Luis Montiel and Carolin Pflueger (2013). "A robust test for weak instruments". In: *Journal of Business & Economic Statistics* 31.3, pp. 358–369.

Orrenius, Pia M and Madeline Zavodny (2015). "Does immigration affect whether US natives major in science and engineering?" In: *Journal of Labor Economics* 33.S1, S79–S108.

Peri, Giovanni, Kevin Shih, and Chad Sparber (2015). "STEM workers, H-1B visas, and productivity in US cities". In: *Journal of Labor Economics* 33.S1, S225–S255.

Peri, Giovanni and Chad Sparber (2011). "Assessing inherent model bias: An application to native displacement in response to immigration". In: *Journal of Urban Economics* 69.1, pp. 82–91.

Poppert, Patrick E and Henry W Herzog Jr (2003). "Force reduction, base closure, and the indirect effects of military installations on local employment growth". In: *Journal of Regional Science* 43.3, pp. 459–482.

Ransom, Tyler and John V Winters (2016). "Do foreigners crowd natives out of STEM Degrees and occupations? Evidence from the US Immigration Act of 1990". In.

Rosenzweig, Mark R, Douglas A Irwin, and Jaffrey G Williamson (2006). "Global wage differences and international student flows [with comments and discussion]". In: *Brookings trade forum*. JSTOR, pp. 57–96.

Shih, Kevin (2016). “Labor market openness, h-1b visa policy, and the scale of international student enrollment in the United States”. In: *Economic Inquiry* 54.1, pp. 121–138.

Shoag, Daniel and Stan Veuger (2018). “Shops and the city: Evidence on local externalities and local government policy from big-box bankruptcies”. In: *Review of Economics and Statistics* 100.3, pp. 440–453.

Smith, Christopher L (2012). “The impact of low-skilled immigration on the youth labor market”. In: *Journal of Labor Economics* 30.1, pp. 55–89.

Stuen, Eric T, Ahmed Mushfiq Mobarak, and Keith E Maskus (2012). “Skilled immigration and innovation: evidence from enrolment fluctuations in US doctoral programmes”. In: *The Economic Journal* 122.565, pp. 1143–1176.

Taubman, Sarah L et al. (2014). “Medicaid increases emergency-department use: evidence from Oregon’s Health Insurance Experiment”. In: *Science* 343.6168, pp. 263–268.

Wilson, Daniel J (2012). “Fiscal spending jobs multipliers: Evidence from the 2009 American Recovery and Reinvestment Act”. In: *American Economic Journal: Economic Policy* 4.3, pp. 251–82.

Wozniak, Abigail and Thomas J Murray (2012). “Timing is everything: Short-run population impacts of immigration in US cities”. In: *Journal of Urban Economics* 72.1, pp. 60–78.

Xiang, Jia (2021). “Physicians as persuaders: Evidence from hospitals in china”. In: *Available at SSRN 3906909*.

Yip, Winnie (1998). “Physician response to Medicare fee reductions: changes in the volume of coronary artery bypass graft (CABG) surgeries in the Medicare and private sectors”. In: *Journal of health economics* 17.6, pp. 675–699.

Yip, Winnie, William Hsiao, Qingyue Meng, et al. (2010). "Realignment of incentives for health-care providers in China". In: *The Lancet* 375.9720, pp. 1120–1130.

Yip, Winnie and William C Hsiao (2008). "The Chinese health system at a crossroads". In: *Health affairs* 27.2, pp. 460–468.

Yip, Winnie, William C Hsiao, Wen Chen, et al. (2012). "Early appraisal of China's huge and complex health-care reforms". In: *The Lancet* 379.9818, pp. 833–842.

Zou, Ben (2018). "The Local Economic Impacts of Military Personnel". In: *Journal of Labor Economics* 36.3, pp. 589–621.