ABSTRACT

Title of Dissertation: THREE ESSAYS ABOUT ECONOMIC AND BEHAVIORAL RESPONSES TO GOVERNMENT INTERVENTIONS

Mrinmoyee Chatterjee
Doctor of Philosophy, 2023

Dissertation Directed by: Professor Judith Hellerstein
Department of Economics

This dissertation studies three different policy interventions and the subsequent labor market responses across different times and contexts.

In Chapter 1, I study changes in the relative wages for women in manufacturing between 1940 and 1950. World War II saw an unprecedented influx of women onto factory floors. While most previous literature focuses on the effects on female labor supply via geographical variation in military mobilization, Rose (2018) highlights the importance of production demand in driving female wartime employment. Using data on the wartime employment of women from Rose (2018), I revisit the framework in Acemoglu et al. (2004), and estimate the impact on relative wages for women, due to both state-wide and industry-wide changes in production demand during WWII. I find that wages increased for women in 1950 compared to 1940 in Durable Manufacturing by 35.4-35.9% in the industry with the largest change in the relative demand for women during WWII, whereas impacts of state-level changes in demand are not significant.

Impacts on wages in Non-Durable Manufacturing are statistically insignificant. The relative wage gains are highest for women with 12 or more years of education, suggesting that the increased demand during WWII allowed some women a "foot in the door" into prized manufacturing jobs. My work helps to reconcile the prior literature connecting WWII to gains in employment for women, and recent work highlighting the importance of the War Production Effort in increasing female employment, by showing how changes in the relative industrial demand for women during World War II significantly increased relative wages for women. In addition, by focusing my analysis on Manufacturing industries, (which saw the largest changes in wartime demand), I consider finer industry variation nationally than any previous work in this context.

In Chapter 2, (co-authored with Dheeraj Chaudhary), we test if the intra-state deregulation of banks between 1970–2000 had any impacts on fertility rates. U.S. states deregulated their banking sectors in a staggered fashion between 1960-1999, increasing efficiency through competition between banks and boosting economic growth within a state. We find that deregulated states saw a decrease of approximately 4–6% in their average fertility rates (in both state-level as well as individual-level data) using a classic difference-in-differences strategy leveraging the staggered timing of deregulation across states. In updating our results with recent econometric literature to account for differences in treatment timing, we find that our results are robust for the sample of observations before 1989 but not for later years. Women aged 20-44 saw a decrease of approximately 2-3% on average fertility rates post-deregulation (in both state-level and individual-level data) between 1970 and 1988. We test different possible mechanisms and find that a likely mechanism could be the increased opportunity costs of having children in a growing job market for women, especially in non-white and poorer households.

In my third and final chapter, (co-authored with Nolan Pope), we look at the academic

impacts of a recent large-scale AC installation program in Chicago Public Schools. Since growing evidence demonstrates that heat impairs student learning, a potential policy solution is clearly investing in air-conditioning. Making use of the timing of roll-out of AC across schools, in a $135 million AC installation program undertaken by Chicago Public Schools between 2013–2017, we analyze the effects on student outcomes. We find no evidence AC installation improved students' end-of-year test scores or grade retention, and find marginal improvements in attendance. These results indicate that improvements in test scores (or other student outcomes) with AC installation could be region-dependent with the impacts of heat on learning, and considering the returns can help school districts better optimize their often limited budgets when striving to improve student performance.

THREE ESSAYS ABOUT ECONOMIC AND BEHAVIORAL RESPONSES TO GOVERNMENT INTERVENTIONS

by

Mrinmoyee Chatterjee

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Professor Judith Hellerstein, Chair
Professor Ethan Kaplan
Professor Nolan Pope
Professor Melissa Kearney
Professor Reeve Vanneman

Foreword

Chapter 2 of this dissertation contains work that is jointly authored with Dheeraj Chaudhary. Chapter 3 of this dissertation contains work that is jointly authored with Nolan Pope. I contributed substantially to the research in both papers and have followed all university guidelines with respect to the inclusion of jointly authored material in my dissertation.

Preface

My dissertation covers a wide range of labor market and student outcomes, in line with my research interests in gender, labor and public policy.

My main area of research interest is in the labor market outcomes for women across different contexts and times. My first chapter looks at the role of government intervention during World War II in changing the demand for female labor in manufacturing industries via the war production effort. Although World War II's impact on women has been intensively scrutinized in economics, most previous literature focuses on the effects on female labor supply via geographical variation in military mobilization. However, recent work by Rose (2018) and Shatnawi and Fishback (2018) highlights the importance of wartime production demand in driving increases in employment for women. I motivate a simple model where male and female labor act as imperfect substitutes and suggest the large gains of about 35% that I see in the relative wages for women in 1950 in Durable Manufacturing, could come from changes in the actual or perceived relative labor productivity for women. I provide anecdotal evidence of increased investments in training or physical capital specifically to improve women's efficiencies within manufacturing during World War II and revised expectations of female labor productivity among employers. My results provide an interesting new context on relative wage gains for women via industrial composition in a much-studied setting.

I am also interested in how the labor market choices of women impact their marriage and fertility decisions outside the labor force since they often go hand-in-hand. In fact, Chapter 1

of my dissertation was born from a different paper where I followed the impact of World War II in changing the labor supply of married women into future cohorts of young women who were too young to have been directly impacted by World War II. In the same vein, in Chapter 2 (which is co-authored with Dheeraj Chaudhary), we look for the impacts of banking deregulation starting in the 1970s on the fertility rates of women. We find a 2-3% decline in the fertility rates of women between 1970-1988 that is robust to new econometric techniques for models with differential timing of treatment units. We suggest that one likely mechanism driving fertility declines was the increased opportunity costs of having children in a growing job market for women, as a result of bank deregulation improving labor market outcomes, especially for non-white and poorer households(Demyanyk, 2008; Beck et al., 2010; Levine et al., 2012; Popov and Zaharia, 2019).

Chapter 3 is co-authored with Nolan Pope, and discusses the impact of AC installation on student outcomes in an environment where the often-discussed detrimental impacts of heating on learning (Park et al., 2021) could be marginal — Chicago Public Schools. We do in fact find that AC installation in Chicago Public Schools did not see significant impacts on student achievement.

All materials are used with permission from co-authors.

Dedication

To Ma, Baba, and Ruihai

Acknowledgments

This dissertation would have not been possible without my advisor, Judith Hellerstein. I am profoundly grateful for her guidance, advice, and support over the last five years. Starting from the rigorous grounding in labor economics in her field class (which my cohort was lucky enough to access), Judy has always inspired me to think more deeply about economic questions and shared her vast knowledge of prior literature. Her influence shaped the leading questions in my research, reiterated the importance of solid theoretical foundations for my ideas, and helped me to adapt to roadblocks during the course of my research. I am so thankful for her willingness to read many unfinished drafts and persevere in workshopping half-baked ideas with me. Her feedback has always been kind but thorough and incisive and always instrumental in improving my work.

I am thankful to Nolan Pope for the opportunity to work with him for two years. The most efficient tools in my research and writing toolset come from our time spent together. His limitless work ethic and encouragement have kept me motivated and joyful in exploring new research ideas.

I am also grateful to Ethan Kaplan for his generosity with his time, insightful feedback, and concern for my progress and well-being during this Ph.D. Assisting in his class on inequality is a cherished memory, and I am very grateful for the introduction to his Race Research Group in my third year.

Thank you to Melissa Kearney for providing me with excellent advice and for serving on my committee. Thank you also to Reeve Vanneman for serving on my committee and acting as Dean's Representative.

My colleagues and friends in the program have uplifted and motivated me for the last six years with their excellent research, frequent collaborations, and many cups of coffee. Thank you to Dheeraj Chaudhary, Ming Fang, Eugene Oue, Victoria Perez-Zetune, Christopher Roudiez, Elena Ramirez, Anusuya Sivaram, and Maranna Yoder for your friendship. Thank you also to my friends in the cohorts above me who made my journey so much easier with their generous advice and support. Thank you Owen Denoeux, Claire Hou, Palak Suri, and George Zuo.

I must thank the many people in the Econ Office who were gracious and patient with all means of administrative support, without which this degree would have never seen completion - Erin, Evan, Mark, Phylicia, Shanna, and most of all, Vickie.

Finally, thank you to my family. I am forever indebted to my parents for their loving support, even in the face of a challenging and unfamiliar path. Thank you to all my friends and acquaintances who let me pet-sit their cats and dogs, and our fourteen foster cats for all the serotonin during this degree. Lastly and most importantly, thank you to my partner Ruihai for his patience, care-taking, and assistance in every step of this endeavor from the very beginning until the end.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AC	Air Conditioning
AK	Alaska
ASEC	Annual Social and Economic Supplement
ATM	Automated Teller Machine
BEA	Bureau of Economic Analysis
BLS	Bureau of Labor Statistics
CES	Constant Elasticity of Substitution
COVID-19	Coronavirus Disease 2019
CPS	Current Population Survey
CPS	Chicago Public Schools
DC	District of Columbia
DE	Delaware
ELA	English Language Arts
FE	Fixed Effects
FLFP	Female Labor Force Participation
HI	Hawaii
HS	High School
IPUMS	Integrated Public Use Microdata Series
MT	Montana
NLSY79	National Longitudinal Survey of Youth - 1979 Cohort
NOAA	The National Oceanic and Atmospheric Administration
NM	New Mexico
NV	Nevada
NYC	New York City

PCI	Per Capita Income
SD	South Dakota
STAR	Tennessee Student/Teacher Achievement Ratio Project
TFR	Total Fertility Rate
TWFE	Two-Way Fixed Effects
US	United States
USES	United States Employment Services
VT	Vermont
WWII	World War 2
WY	Wyoming

Chapter 1: The WWII Production Effort and Changes in Labor Demand for Women

1.1 Introduction

The 1940s saw the largest proportional change in female labor force participation (FLFP) in the United States since the nineteenth century, with the influx of nearly 7 million women into the labor force during World War II (Rose, 2018; Acemoglu et al., 2004).

Previous papers have extensively studied the role of World War II in increasing female employment (Acemoglu et al., 2004; Goldin, 1991a; Goldin and Olivetti, 2013) and impacting other sectors of a woman's life such as fertility and marriage (Goldin, 1991b; Fernández et al., 2004; Fernández, 2013; Doepke et al., 2015; Larsen et al., 2015). Crucially, however, most prior work interprets the impact of World War II (WWII) on female labor force participation as a story of changing female labor supply rather than changes in the relative labor demand for women. However, there is no clear rationale to do so.

Many cultural and institutional barriers to women's work were lowered during World War II (such as segregation into low-wage occupations, marriage bars that legally prevented women from working in most occupations, and cultural norms that viewed married women working unfavorably (Goldin, 1991b)). The lack of these barriers could translate to women finding it

easier to supply their labor, but it could also make it easier for industries to now employ women.

Furthermore, the monumental inflow of women into the workforce during World War II was due to the rising demand for workers in the American war production effort, and the prime reason for lowering the aforementioned barriers to women's work in the first place. During WWII, American industry provided more than two-thirds of all Allied military equipment, outstripping production levels of not only other nations but also the US's own pre-World War II capacities (Burns, 2007). As the war progressed and the need for workers intensified, nearly a quarter of the prime-age male labor typically used to fill such needs was drafted by the military and taken out of the available workforce. Consequently, women were needed in large numbers to keep up with production demands and urged to join the workforce via a nationwide public campaign (Hartmann, 1982; Milkman, 1987).

A recent string of literature highlights the importance of production demand on increases in FLFP during World War II. Rose (2018) shows that the spatial distribution of female employment in WWII was not impacted by men being drafted out of the workforce for the military as in Acemoglu et al. (2004) but rather by geographical variation in wartime production and the need for labor. Shatnawi and Fishback (2018) focus on Pennsylvania and find that the relative demand for female production workers rose by more than 40% during World War II and hypothesize that part of the increase could come from women learning new skills during the war, adjustments in production technology to maximize women's output and employers revising their perceptions of female productivity.

Not only did the exigencies of the production effort draw millions of women onto factory floors, but it also drastically changed the industrial composition of the female workforce. Less than a quarter of the female workforce worked in manufacturing in 1940, and even within manufac-

turing, approximately 70% of all women worked in Non-Durable Industries, primarily in Food, Textiles, and Apparel (see Figure 1.1). The war production effort rapidly placed women in new industries like Machinery, Electrical Machinery, and Transport Equipment (Warm, 2002), where women only formed a small proportion of the workforce in 1940. The proportion of women in Durable Manufacturing, for instance, went from 8.6% in the 1940 Census to a peak of nearly 25% in October 1944, rising by 188.4%. In comparison, Non-Durable Manufacturing (which also contained war-critical industries like Chemicals and Rubber) went from being 39.5% female in 1940 to a high of 45.3% female in Oct 1944, a change of only 14.7% (Pidgeon, 1947).

Given this background, it is not unreasonable to expect that in industries where the relative wartime demand for women rose most disproportionately, relative wages for women could stay elevated even after World War II ended. A large proportion of war-critical industries did not employ many women before WWII and had to invest in changing physical capital and training to make production jobs easier for women, changes that could permanently increase the relative productivity of women in these industries.

In their seminal paper, Acemoglu et al. (2004) assert that state-wide differences in drafting men into the military during World War II drove increases in female employment and changed their relative wages. Rose (2018) demonstrates instead, that the spatial distribution of wartime production was the main driver of female employment during World War II, and not military mobilization as per Acemoglu et al. (2004) and subsequent literature.[1] However, Rose (2018) stops short of estimating the impact of wartime production on women's relative wages post WWII. Additionally, none of the previous literature linking FLFP and World War II considers

[1] Rose (2018) also points out how the dependent variable of "Weeks Worked" used by Acemoglu et al. (2004) to measure employment might be flawed due to changing definitions between the 1940 and 1950 Censuses that might artificially inflate women's employment in 1950. For more see Section 1.4.

the impacts of industrial variation in female employment during World War II, apart from the recent paper by Shatnawi and Fishback (2018) who focus on Pennsylvania.[2]

In this chapter, I extend the analysis from Rose (2018) to measure the impact of changes in the relative wartime employment of women on their relative earnings between 1940 and 1950. I use the framework outlined in Acemoglu et al. (2004), and incorporate industry variation in wartime production alongside the state-level variation that is used to study effects on female employment by Rose (2018), and test the impacts on relative wages for women compared to men. A relative rise in female earnings, combined with the increase in female employment between 1940 and 1950 found by Rose (2018) in Durable Manufacturing, would indicate a net increase in the relative demand for women post World War II in Durable Manufacturing.

I use new data from Rose (2018) on the placements of job seekers to available openings by the United States Employment Service (USES) between 1944 and 1946 across states and industries, and define two new measures of relative female demand during World War II. The first is a change in the number of women relative to the change in total workers in an industry (or state) during World War II due to USES assignments. The second is a change in the relative proportion of women in an industry (or state) during World War II due to USES assignments. Both measures are intended to capture the combination of high wartime overall labor demand and a drastic rise in the demand for women specifically. I restrict my focus to workers in manufacturing since more than 50% of all wartime labor assignments in the data are in manufacturing, and this allows me to consider finer sub-industry detail within manufacturing industries.

[2]Even though Rose (2018) highlights the importance of production demand in his paper, he uses the geographical variation in military contracts, grouping all industries together for most of his analysis. He does split them in Table 6 (p.25) to look at the lasting impacts of female wartime work on employment in 1950 and finds that state-economic areas with higher female employment during World War II did have increased employment of women in 1950, but only in Durable Manufacturing industries.

In Durable Industries, I find that working in the industry with the highest increase in the relative demand for women (Transport Equipment) during World War II compared to the industry with the lowest increase in the relative demand for women (Fabricated Metal, or Lumber and Wood) increased wages for women (relative to men) between 1940 and 1950 by 35.4–35.9% according to my preferred specification using either measure of relative demand for women. In contrast, in Non-Durables, I do not find statistically significant impacts of changes in the relative demand for women during World War II on relative wages for women. The magnitudes of the impacts, in addition to being statistically insignificant, are also smaller.[3]

Taken in conjunction with the results of Rose (2018), who finds gains in the employment of women in Durable Manufacturing in 1950 (compared to 1940) in state-economic-areas of high wartime demand for women, my results indicate a rise in the relative demand for women in Durable Manufacturing that persisted until 1950 in the form of both employment and relative wage gains for women compared to men, in magnitudes similar to that found in Pennsylvania by Shatnawi and Fishback (2018). The fact that I find significant positive impacts on relative wages for women in Durable Manufacturing industries (where women formed a very small minority of workers before World War II) and not in Non-Durable Manufacturing industries (which held more than 60% of all women in manufacturing in 1940), even after using measures that account for the prior industrial composition of women in the industry, highlights how wartime employment and demand for workers evolved differently across these two manufacturing sectors.

In Non-Durable Manufacturing, war-critical industries such as Chemicals or Rubber, had workforces that were on average, about 20% female before World War II. As production needs

[3]In my preferred specification, they range from an insignificant decrease of 1.6% to an insignificant increase of 4.4% in the relative wages for women working in the industry with the highest rise in the relative demand for women (Leather and Footwear, or Rubber) relative to the industry with the lowest rise in the relative demand for women (Tobacco) in Non-Durable Manufacturing.

mounted, and the available male labor force shrank, new women hired were likely placed into roles that were not drastically different from the roles that women had been in before. For example, women already worked as organic and inorganic chemists before World War II. They could easily and quickly be trained to work as foundry chemists as part of the war effort, according to a report published in 1942 by the United States Employment Service (USES, 1942). In contrast, war-critical industries in Durable Manufacturing, like Transport Equipment, had work-forces that were only on average 10% female in 1940, who worked almost entirely in clerical and communication positions (USES, 1942). The urgent need for riveters, machinists, and welders in these industries had to be filled at least in part by women during World War II, who had to either be trained in new skills or placed into revised production processes where employers made adjustments to physical capital in favor of women (USES, 1942; Joiner and Welner, 1942). Anecdotal evidence points out that this may have even changed or updated employer perceptions of female productivity (Hartmann, 1982; Milkman, 1987; Encyclopedia.com, 2021). Hence, drastic increases in the number of women in Durable Manufacturing industries during World War II could have raised real or perceived female productivity, leading to a permanent rise in relative wages, in contrast to women in Non-Durable Manufacturing.

Interestingly, I find no statistically significant impacts of the overall WWII industrial demand on the wage growth between 1940 and 1950 either for men or women. In addition, changes in the relative female demand or overall demand for workers by state also do not impact male or female wage growth between 1940 and 1950 in a statistically significant way. Even after accounting for possible state-wide variation in military mobilization, only industry-wide variation in the relative wartime demand for women is a significant driver of wage growth for women in manufacturing. This is notable since nearly all previous literature focuses on state-level variations in female

employment during World War II.

Although I do not have data on USES placements by different demographic characteristics, I test if the impact of shocks in the relative demand for all women on the wages of women in 1950 differs by worker characteristics. I find that the largest increases in relative wage growth between 1940 and 1950 are for women with 12 or more years of education (compared to similar men). This is similar to the subgroup of women that Goldin and Olivetti (2013) find to have the largest employment gains as a consequence of World War II, albeit via variation in military mobilization. Consistent with an employer demand story, this group could be likeliest to see relative wages increase due to changes in the "perceived productivity" of women lowering the non-pecuniary Beckerian costs of employers hiring such women (Becker, 1962), or via specific training developed during the war that changed the real relative productivity of women (USES, 1942).

The remainder of this chapter is organized as follows: Section 1.2 reconciles the historical background and previous literature, and Section 1.3 details a simple model to test my theories. I discuss the sources of my data and the construction of samples in Section 1.4. Section 1.5 outlines my empirical strategy, and Section 1.6 discusses the results. I conclude in Section 1.7.

1.2 Background and Motivation

1.2.1 The War Production Effort

When the U.S. officially entered World War II in December 1941, the armed forces took many prime-aged men out of the civilian workforce. At the peak of the war in 1945, more than a third of men aged 18–44 were in active duty service due to a mix of the draft and voluntary

enlistment (Hartmann, 1982). This produced a sudden void in the U.S. labor force.

Around the same time, the war production effort on the home front began in earnest in January 1942 with the establishment of the War Production Board. The American War Production Effort led to an unprecedented rise in the demand for workers to make products considered essential to the war such as tanks, munitions, and airplanes (Warm, 2002). President Roosevelt set staggering production goals for factories across America, and in four years, U.S. production of defense materials, already the largest in the world in 1941, doubled in size (Burns, 2007).

The War Production Board reported in November 1944, "... in 1940, the average monthly production rate of airplanes was 500; in 1941 it rose to 1,600 and, in 1942, reached the 4,000 level. The value of guns and ammunition increased from an annual rate of $170 million in 1940 to $900 million in 1941, and sky rocketed to over $15 billion in 1942. From 250 in 1940, the building of new ships jumped to 2,000 in 1941, and then shot up to 11,500 in 1942. These spectacular achievements were similarly dwarfed when the output data for 1943 end 1944 *(sic)* became known. Aircraft production, for example, reached a monthly average of almost 8,000." (USES, 1948, p.16).

The combination of the War Production Effort and the military draft led to a rapidly tightening labor market where women were recruited in large numbers and were encouraged by employers and the War Manpower Commission to join non-traditional roles in the labor market, as part of their patriotic duty (Hartmann, 1982; Warm, 2002). Several recruiters compared factory labor with household tasks in an effort to entice women to join the labor force, saying for example "Instead of cutting the lines of a dress, this woman cuts the pattern of aircraft parts. Instead of baking cake, this woman is cooking gears to reduce the tension in the gears after use.... They are taking to welding as if the rod were a needle and the metal a length of cloth to be sewn. After

a short apprenticeship, this woman can operate a drill press just as easily as a juice extractor in her own kitchen. And a lathe will hold no more terrors for her than an electric washing machine" (Milkman, 1982, p.341).

Soon, factory floors across the country were filled with female production workers, drawing women from outside the labor force as well from more typically 'female' sectors of the labor market (Hartmann, 1982).

1.2.2 The United States Employment Service

In September 1939 (when the war began in Europe), the U.S. was still dealing with mass unemployment from the vestiges of the Great Depression. Within three years, this turned into an acute labor shortage as the War Production Effort surged (Brennan, 1973). In this new labor market, the United States Employment Service (USES) played an instrumental role in matching jobs to available workers (Rose, 2018).

The USES was created prior to World War I and reinstated in 1933 by the Wagner-Peyser Act to help employers and job seekers in the immediate aftermath of the Great Depression (Brennan, 1973). In April 1941, the national and state employment service offices put together a large-scale registration of occupational skills to measure the availability of key craftsmen and production workers needed for the burgeoning defense production. (USES, 1948)[4] This database helped analyze the suitability of workers at hand for high-demand occupations. It also provided a new strategy of "job dilution" to fill the outstanding demand for skilled workers — dividing skilled occupations into several simpler operations that could be performed by workers with a

[4]Prior to WWII, the USES had also compiled the first edition of the "Dictionary of Occupational Titles" in 1935, which collected information about skills required across different occupations and helped streamline hiring across industries.

minimum of experience and skill (USES, 1948).

As job seekers applied to state and national employment services offices, the USES worked as a national unified labor clearing house. Rose (2018) states that "in 1944 alone, the USES was responsible for filling 11.4 million vacancies, including 3.8 million with women. About 6.8 million of these jobs were in manufacturing industries" (Rose, 2018, p.7).

USES placements were highly concentrated in manufacturing industries, with assignments to manufacturing making up 54.8% of all USES assignments in the available data (see Appendix Figure A.1). Durable Manufacturing alone made up 31.68% of all USES assignments in the data, more than double the fraction of placements in the next largest non-manufacturing industry (Personal Services, 12.12% of all assignments). Therefore, I focus my attention on variations in wartime demand within manufacturing.

Table 1.1 shows the distribution of employment by industry and sex in the 1940 and 1950 Decennial Censuses as well as for the USES job placements between 1944-46 accessed from Rose (2018). The number of USES placements in manufacturing represents a significant proportion of the 1940 workforce. Moreover, the data on USES placements are from the last 2 years of WWII (from the third quarter of 1944 to the second quarter of 1946), implying that the total numbers placed during the entirety of World War II including assignments from 1942 and 1943 are almost definitely higher than the numbers reported in Table 1.1.

The proportion of female workers in the USES assignments represents higher proportions female of manufacturing industries than seen in 1940 or 1950 (consistent with the idea that more women were employed in the war effort across all industries than before the war or after). However, the relative influx of women was more drastic in some industries than others. For example, the USES placed 237,574 women in Food and Beverages between 1944–1946 (see

Table 1.1), which represented a 7% change in the total number of women from 1940. In contrast, the USES placed fewer overall women in Transport Equipment (179,947), but that represented a 15.6% change in the total number of women from 1940, more than twice the change in Food and Beverages (see Table 1.2).

1.2.3 Changes in Female Productivity

In the first half of the twentieth century, professions tended to be fairly segregated by sex (Bergmann and Adelman, 1973). Of all women working in non-agricultural industries in the 1940 Census, nearly half worked in Retail Trade, Professional and Personal Services (Ruggles et al., 2020). When breaking down the proportions of women within manufacturing industries (as seen in Figure 1.1), Food, Textiles, and Apparel contained more than half of all the women employed in manufacturing in 1940.

During World War II, women flooded new industries, where they were often hired for the very first time as production workers. Within Durable Manufacturing, a large number of industries that now employed women as welders, machine operators, riveters, or solderers had previously only employed them in clerical capacities as bookkeepers, typists, receptionists, or clerks (USES, 1942). For the first time, women were allowed a 'foot in the door' into new (and lucrative) factory work.

A possible channel through which the relative productivity of women may have permanently increased could be the investment in physical capital and vocational training for women during World War II. Confronted with a shortage of skilled male workers, the USES released a report in February 1942 redefining several occupations that did not typically hire women into smaller

occupations and officially designated their "suitability" for women, as well as the time it would require to train them in said occupations in an effort to utilize the available numbers of women to fulfill production demands (USES, 1942). In this report, the USES gave an example of such a practice: "For example, it usually takes years of training to become efficient in all the aspects of the occupation of Precision Lens Grinder. Certain phases such as blocking, cementing, inspecting, however, can be taught within a comparatively short time, and women have been found to be very adaptable to these tasks" (USES, 1942, p.VIII). In 1943, The National Industrial Conference Board reported that a large majority of manufacturing plants had broken down complex assembly jobs into simpler operations with lower training required to accommodate female workers (Hartmann, 1982). Opportunities for pre-employment vocational training for women increased, and employers made capital investments into technology to more easily integrate women into the production process (Hartmann, 1982; Joiner and Welner, 1942). All these measures taken to utilize women more effectively within manufacturing may have increased their real relative productivity leading to relative wage growth.

Another channel through which the relative demand for women could have changed is employer perceptions of productivity, in contrast to (or alongside) changes in the real productivity of women. Purportedly, attitudes of some employers did change with regards to the ability of women to perform tasks involving greater motor skills and technical precision (Hartmann, 1982; Encyclopedia.com, 2021). Milkman (1987) quotes the trade journal Automotive War Production saying in their October 1943 issue, "..on certain kinds of operations—the very ones requiring high manipulative skill—women were found to be a whole lot quicker and more efficient than men" (Milkman, 1987, p.59). General Motors president Charles E. Wilson found women "more enthusiastic and showing much better spirit" (Hartmann, 1982, p.63).

1.3 Theory

I start with a standard Cobb-Douglas production function of the following form:

$$Y_{jst} = K_{jst}^{\alpha} L_{jst}^{(1-\alpha)} \tag{1.1}$$

and consider labor input to have male labor and female labor as imperfect substitutes in a constant elasticity of substitution (CES) function, borrowing a workhorse model from Autor et al. (1998), also used by Acemoglu et al. (2004) in their paper.[5]

$$Y_{jst} = K_{jst}^{\alpha} \left(\left((A_{jst}^{M} M_{jst})^{\rho} + (A_{jst}^{F} F_{jst})^{\rho} \right)^{(1/\rho)} \right)^{(1-\alpha)} \tag{1.2}$$

Using this production function, the output produced in an industry j in state s at time t is a function of capital (K_{jst}), male labor (M_{jst}) and female labor (F_{jst}) in a constant elasticity of substitution function, with respective 'labor-augmenting' productivity terms—A_{jst}^{M} and A_{jst}^{F}. Jobs within manufacturing were largely sex-segregated pre-World War II (Bergmann and Adelman, 1973), and it seems plausible that men and women were imperfect substitutes in manufacturing industries, with their relative contribution to output not just impacted by their rate of substitutability ρ, but also the relative gender-specific labor-augmenting productivity terms A_{jst}^{M}, and A_{jst}^{F}.

Changes in the relative labor-augmenting productivity across time could represent several ways in which the effectiveness of gender-specific labor input is impacted during WWII via investments in physical capital, changes in the access and quality of vocational training available,

[5] Although Acemoglu et al. (2004) use a production function with CES between male and female labor, they only consider markets at the state-year (st) level rather than the industry-state-year level (jst) that I do in this chapter.

or the perceptions of employers regarding the group-specific productivity of each gender in an industry or state.

If I assume that capital does not respond to changes in the availability or demand for labor between 1940-1950, then while taking the marginal product of Y_{jst} with respect to labor, the capital term will float outside.[6] Therefore to simplify, I, like Acemoglu et al. (2004), drop capital from my production function and have it as

$$Y_{jst} = \left((A^M_{jst} M_{jst})^\rho + (A^F_{jst} F_{jst})^\rho\right)^{(1/\rho)} \tag{1.3}$$

In a competitive market, factors would be paid their marginal product. Deriving the marginal products of each type of labor gives me:

$$f_{Fjst} = \left((A^M_{jst} M_{jst})^\rho + (A^F_{jst} F_{jst})^\rho\right)^{\frac{1-\rho}{\rho}} \times \left(A^{M\rho}_{jst} M^{\rho-1}_{jst} \frac{\partial M_{jst}}{\partial F_{jst}} + A^{M\rho-1}_{jst} M^{\rho}_{jst} \frac{\partial A^M_{jst}}{\partial F_{jst}} + A^{F\rho}_{jst} F^{\rho-1}_{jst} + A^{F\rho-1}_{jst} F^{\rho}_{jst} \frac{\partial A^F_{jst}}{\partial F_{jst}}\right) \tag{1.4}$$

$$f_{Mjst} = \left((A^M_{jst} M_{jst})^\rho + (A^F_{jst} F_{jst})^\rho\right)^{\frac{1-\rho}{\rho}} \times \left(A^{M\rho}_{jst} M^{\rho-1}_{jst} + A^{M\rho-1}_{jst} M^{\rho}_{jst} \frac{\partial A^M_{jst}}{\partial M_{jst}} + A^{F\rho}_{jst} F^{\rho-1}_{jst} \frac{\partial F_{jst}}{\partial M_{jst}} + A^{F\rho-1}_{jst} F^{\rho}_{jst} \frac{\partial A^F_{jst}}{\partial M_{jst}}\right) \tag{1.5}$$

I follow the lead of Acemoglu et al. (2004) and make some more restrictive simplifying assumptions. First, I assume that in any given period where we are measuring wages, the demand

[6] While this may seem restrictive, the labor-augmenting productivity terms help to capture any changes in capital investment across time that might improve labor output. The possibility I rule out by making the above assumption is simply that capital input is not a function of contemporaneous labor input in the period that I measure my outcomes in (1940-1950).

for male labor does not endogenously vary with the demand for female labor, and vice-versa, which means that $\frac{\partial M_{jst}}{\partial F_{jst}} = \frac{\partial F_{jst}}{\partial M_{jst}} = 0$. Secondly, I assume that in any given time period, the relative labor productivity is not an endogenous function of labor input in that period t, i.e. $\frac{\partial A^M_{jst}}{\partial F_{jst}} = \frac{\partial A^F_{jst}}{\partial F_{jst}} = \frac{\partial A^F_{jst}}{\partial M_{jst}} = \frac{\partial A^M_{jst}}{\partial M_{jst}} = 0$.[7] These assumptions help in simplifying our marginal products with respect to each type of labor as

$$f_{Fjst} = \left((A^M_{jst}M_{jst})^\rho + (A^F_{jst}F_{jst})^\rho\right)^{\frac{1-\rho}{\rho}} \cdot \left(A^{F\rho}_{jst}F^{\rho-1}_{jst}\right) \tag{1.6}$$

$$f_{Mjst} = \left((A^M_{jst}M_{jst})^\rho + (A^F_{jst}F_{jst})^\rho\right)^{\frac{1-\rho}{\rho}} \cdot \left(A^{M\rho}_{jst}M^{\rho-1}_{jst}\right) \tag{1.7}$$

However, wages are not likely to be set competitively in the labor market that I study between 1940 and 1950. Despite the growing demands of war production, and the War Manpower Commission exhorting employers to hire workers from different industries, there were still many roadblocks to a fully competitive national labor market with wage-taking behavior across industries, due to variation in skills, training, sex-typing of certain jobs in industries and gendered cultural norms.

Therefore, I adapt a wage-setting model from Card et al. (2018) to include frictions in wage-setting at the industry level that might lead to upward sloping supply curves. In this model, two types of workers (men M and women W), observe an industry-state-year specific wage pair (w_{Msjt}, w_{Fsjt}) posted across J manufacturing industries in state s and year t, and choose where to work based on their indirect utility function

$$v_{iGjst} = \beta_{Gst}\ln(w_{Gjst}) - \alpha_{Gjt} + \varepsilon_{iGjst} \tag{1.8}$$

[7] This assumption does not mean that I do not allow relative labor-augmenting productivities to vary across time, rather that the elasticities of labor-augmenting productivity with respect to current period employment is zero.

where w_{Gjst} are the wages posted for gender G, by an identical representative firm in an industry j in state s and year t, and α_{Gjt} is an industry-specific preference (or cost) for all workers of gender G and ε_{iGjst} captures idiosyncratic preferences for working in an industry j in a state s in the year t. α_{Gjt} is intended to capture costs that may be incurred by individuals due to their gender for the industry they work in, ranging from more tangible costs (such as say childcare services making an industry a more accessible/less costly for women than an industry b that does not provide them) to less tangible costs (such as the sociocultural norms around working in a certain industry as a woman). β_{Gst} denotes the labor-supply responsiveness of each gender G to the posted wages and can change across state and time (since they could be impacted by rental costs in a state-year, availability of childcare in a state-year, etc.).

Although wages can be posted differently for state-industry-year markets in my model, I only consider the access costs α_{Gjt} to differ across industry-year markets for each gender G. There may still be costs or preferences that prohibit workers from working across different state markets (for example, the costs of migration or living across states), but I assume that they are either captured in the labor-supply responsiveness of the worker (β_{Gst}) or eliminated in my context of World War II, where notably workers did migrate to different states for work in large numbers and wartime rations tried to keep costs of living stationary. Thus, in this model, I account for frictions that might cause upward-sloping supply curves at both the state and the industry-level markets via α_{Gjt} (costs of working in a particular industry j), and β_{Gst} (costs or benefits of working across different states s) incorporated into the indirect utility function of a representative worker.

Then, the labor supply function for each type of worker can be derived from the model as a linear combination of the log effective wages posted by an industry-state, gender-specific

parameters, and industry-by-gender-specific parameters.[8] The cost-minimizing wages posted by industries in the presence of those supply functions look like:

$$w_{Fjst} = \left(\frac{\beta_{Fst}}{1+\beta_{Fst}}\right)\mu_j f_{Fjst} \tag{1.9}$$

$$w_{Mjst} = \left(\frac{\beta_{Mst}}{1+\beta_{Mst}}\right)\mu_j f_{Mjst} \tag{1.10}$$

where f_{Gjst} is the marginal product of labor of gender G and μ_j represents the marginal cost of production.

Substituting the marginal products from our CES production function in Equations (1.6) and (1.7) into the wage equations above and taking log we get the individual wage equations:

$$ln(w^F_{jst}) = ln\left(\frac{\beta_{Fst}}{1+\beta_{Fst}}\right) + ln\mu_j + \left(\frac{1-\rho}{\rho}\right) ln\left((A^M_{jst}M_{jst})^\rho + (A^F_{jst}F_{jst})^\rho\right) + ln\left(A^{F\rho}_{jst}F^{\rho-1}_{jst}\right) \tag{1.11}$$

$$ln(w^M_{jst}) = ln\left(\frac{\beta_{Mst}}{1+\beta_{Mst}}\right) + ln\mu_j + \left(\frac{1-\rho}{\rho}\right) ln\left((A^M_{jst}M_{jst})^\rho + (A^F_{jst}F_{jst})^\rho\right) + ln\left(A^{M\rho}_{jst}M^{\rho-1}_{jst}\right) \tag{1.12}$$

If I convert this into a wage premium equation by subtracting Equation 1.12 from Equation 1.11 (as per Acemoglu et al. (2004)), we get the following:

[8] An additional assumption I need to reach these labor supply functions is that industries are not using strategic pricing models, i.e. their optimal wage setting functions are set to minimize their own costs, and not responsive to wages set by other industries. Given the context of maximizing production during World War II, and the fact that I have a large enough number of industries within manufacturing (20), I feel justified in assuming away strategic price-setting across industries.

$$ln\pi_{jst} = ln\left(\frac{w^F_{jst}}{w^M_{jst}}\right) = ln\left(\frac{\beta_{Fst}}{1+\beta_{Fst}}\right) - ln\left(\frac{\beta_{Mst}}{1+\beta_{Mst}}\right) + +ln\left(A^{F\rho}_t F^{\rho-1}_t\right) - ln\left(A^{M\rho}_t M^{\rho-1}_t\right) \tag{1.13}$$

which can be simplified to

$$ln\pi_{jst} = \underbrace{\rho ln\left(\frac{A^F_{jst}}{A^M_{jst}}\right) + (\rho-1)ln\left(\frac{F_{jst}}{M_{jst}}\right)}_{\substack{\text{Terms from the classical CES wage}\\ \text{premium equation used by AAL (2004)}}} + \underbrace{ln\left(\frac{\frac{\beta_{Fst}}{1+\beta_{Fst}}}{\frac{\beta_{Mst}}{1+\beta_{Mst}}}\right)}_{\substack{\text{Added term from}\\ \text{CCK (2018) model}}} \tag{1.14}$$

where the log wage premium is a linear combination of the log relative labor-augmenting productivity for women and men in a state s, industry j and time t, the log relative employment of women to men in an industry j, state s and year t, and the log of a function of β_{Gst} i.e. the indirect utilities from wages in state s in year t for a gender $G \in \{M, F\}$.

The relative labor-augmenting productivity for women and men are not observable, and past papers have typically used linear time trends to approximate for changes in the relative labor-augmenting productivity across time to calculate the structural parameters from the wage premium equation (Acemoglu et al., 2004; Katz and Murphy, 1992). However, as I discuss in Section 1.2.3, the labor-augmenting productivity during World War II might be changing not only across time but differently across industries and states as well due to the War Production Effort. Thus, I want to explicitly allow for industry-state-level changes in labor augmenting productivity. Due to data restrictions, I assume that it changes additively in state-time and industry-time for each gender, rather than within an industry-state-year.

The unobservable labor supply elasticities β_{Gst}, that change by state s and time t can

be captured by state-by-gender-by-time fixed effects that I plan to incorporate in my empirical specification. These effects will also capture any state-by-gender-by-time changes in wages. Therefore, I can only separately estimate the linear industry-by-time effects of relative productivity change on relative wages in my model.[9]

1.4 Data

1.4.1 Decennial Census Data

I use the Integrated Public Use Microdata Series (IPUMS) of the US Decennial Censuses from 1940 and 1950 (Ruggles et al., 2020). I use the 1% sample for both decennial censuses which is a 1-in-100 random sample of the population.

I use the sample restrictions imposed by Acemoglu et al. (2004), Goldin and Olivetti (2013), and Shatnawi and Fishback (2018) in their papers. I limit the sample to working-age persons (aged 15–64) and for the sake of uniformity across censuses, I drop those who live in group quarters or are foreign-born. I drop those individuals born or residing in Alaska, Hawaii, and D.C. since these states were established after World War II and did not exist in the 1940 and 1950 censuses. I also follow Acemoglu et al. (2004) and Goldin and Olivetti (2013) and drop Nevada since it was by far the least populous state in 1940 and underwent several demographic changes in the period that I focus on. I restrict the sample to individuals working in non-farm occupations. I therefore also drop those who do not have an industry of work defined or are

[9] If I change my assumption regarding labor-supply elasticities such that they only differ by gender and year but not across states (the original Card et al. (2018) model only has them vary by gender), then they can be estimated via gender-by-time fixed effects only, and I can try to estimate the impact of relative demand shifts during World War II impacting relative wages both by state and industry. Other factors could also impact labor supply responsiveness to wages such as education, marital status, age, and number of children. I include these in my final estimation specification, and let their impact on wages differ by gender and year.

classified as ‘not within the labor force’.

I focus on manufacturing industries for my main analysis since they saw by far the largest increase in labor demand during World War II (see Appendix Figure A.1). When I limit my sample to just those working in manufacturing industries, I also drop Montana, New Mexico, and Wyoming since they either do not have any female workers in manufacturing surveyed in 1940, 1950, or both. I end up with a final sample of 44 states and 20 industries within manufacturing.

For the final sample of workers, I consider only those who are actively in the labor force, not unpaid family workers or self-employed or in the armed forces at the time of the Decennial Census survey. Wages are inflation-adjusted in 1950 to be comparable to 1940 U.S. Dollars. I also follow Acemoglu et al. (2004) and drop workers who report implausibly high or implausibly low (compared to the full population) hourly wages and impute top-coded wages as 1.5 times the maximum wage. I use annual wages as my primary dependent variable (due to flaws in the reporting of ”weeks worked” between censuses). I discuss robustness checks with hourly wages in Section 1.6.4.

The labor input measure I use to capture relative log employment ($\left(\frac{F_{jst}}{M_{jst}}\right)$) in Equation 1.14 is the log of all (population-weighted) women who work full-time in an industry-state-year, divided by all (population-weighted) men who work full-time in that industry-state-year. I prefer the relative log worker count specification for measuring employment rather than relative average weeks worked in a year or average hours worked last week due to data discrepancies between the 1940 and 1950 Decennial Censuses.

As Rose (2018) details in his paper, in 1940, the Census asked respondents to report only their full-time equivalent weeks worked whereas in 1950, any week where the respondent worked was considered a full-time week. For example, if a part-time teacher taught two days a week, they

would report having the same number of work weeks in the 1950 Census as another woman who taught full-time (five days a week). This implies that part-time weeks were more likely to be counted as full-time in 1950. Combined with the propensity of women to be working part-time, this error might artificially inflate relative female employment compared to male employment in 1950. The 'hours worked last week' measure might be considered an improvement, but it also has two major problems. First, the question asks respondents to report the number of hours they worked in the past week, which may or may not be the worker's usual hours worked in an average week in the year. Second, the passage of the Fair Labor Standards Act standardizing the 'regular' workweek to 40 hours was passed in 1940 after the Decennial Census was recorded. Workers in 1950 may be more likely to incorrectly report 40-hour workweeks than in 1940, artificially inflating employment in 1950. When I plot this measure in Appendix Figure A.3, I do see bunching at 40 hours in the 1950 data for both men and women as compared to the 1940 data. The combination of measurement error in the hours worked and the weeks worked measures would compound the noise in an hourly wage measure created by dividing annual wages by weeks times hours worked. I discuss this further in Section 1.6.4.

Lastly, the data on USES placements (that I discuss in the next subsection) are only available as the total number of workers in an industry (or state) by gender, and not the weeks or hours worked by a worker. Using my relative employment measure in terms of the total number of workers is therefore symmetric with measuring relative demand shocks using workers from the USES data.

1.4.2 United States Employment Service Data

I use data on the placements of the United States Employment Service (USES), made available by Rose (2018). The data comes from "The Labor Market" reports of the War Manpower Commission's Bureau of Program Planning and Review from 1942–1945, and reports of the same name from the Labor Department after 1945.

The data details the number of women and men placed between 1944–1946 across different industries and states (the numbers within an industry-state are not reported). I use these values as a proxy for measuring the relative wartime employment demand across different industries and states.

The USES was an important agent in matching employers and workers during the wartime production effort, particularly as the needs for war production grew more acute (Rose, 2018). The assignments data shows the industry or state that a job seeker was placed into when applying for a job with the USES itself, or with state or local Employment Offices (which worked in collaboration with the USES). The USES was particularly key in classifying high-priority jobs in war-critical industries into combinations of different occupations that could be performed by workers with the minimum possible amount of experience and training, via "job-dilution and upgrading" programs (USES, 1948). By 1944, the USES was "responsible for filling seven out of ten jobs in manufacturing" (War Manpower Commission, 1944, p. 40).

The available data on the USES assignments are from placements in the last quarter of 1944 to the third quarter of 1946. The timing of these records has both advantages and drawbacks. One major drawback is that I capture demand from the end of World War II when the bulk of the war production effort had slowed down. If this is indeed the case, then any effects I manage to see of

this increase in demand are likely an underestimate of the total effects.

On the other hand, we may be worried about the exogeneity of the USES placements across industries, particularly for women. There could certainly be some industries more ready to hire women or Black workers, who also are more responsive in updating their wages.[10] As the war went on and traditional white male labor kept getting drafted into the military, the tightening labor market improved employment prospects of many minority and non-traditional groups including black men, women, older workers who would have otherwise retired, the physically handicapped and teenage youth (USES, 1945). Thus, wartime industrial demand in the later years of World War II may be less influenced by employer discrimination based on gender (and other factors), and the USES data from 1944-1946 are better estimates of a true demand shock for women.

The distribution of average employment by industry and state across manufacturing industries in 1940, 1950, and between 1944–1946 from USES assignments are shown in Table 1.1. The final numbers for 1940 and 1950 count only full-time employed workers and are weighted by population-level weights aggregated from the Census samples.[11]

1.5 Empirical Strategy

Equation 1.14 that I derive in Section 1.3 to estimate changes in relative wage premia for men and women between 1940 and 1950 establishes that the log relative wages for women (as compared to men) across time are a function of the change in the log unobservable (female

[10] In an anecdote in her paper, Anderson (1982) cites the story of "Samella Banks, a black woman who along with five other white women was told to apply to the Cadillac Motor Company in November 1942 by the USES. She was told that there might be a janitress opening in a day or two while the white women were hired as welder trainees" (Anderson, 1982, p.8).

[11] Full-time workers are defined as those who report working more than 35 hours a week and more than 40 weeks a year in 1940, and more than 40 weeks a year in 1950 since all weeks reported are supposed to be full-time equivalent weeks in 1950.

to male) relative labor-augmenting productivity across time $\Delta_t ln\left(\frac{A_{js}^F}{A_{js}^M}\right)$, the change in the log relative employment of women to men across time $\Delta_t ln\left(\frac{F_{js}}{M_{js}}\right)$ and the change in the log ratio of (female to male) relative labor supply elasticities across time $\Delta_t ln\left(\frac{\frac{\beta_{Fs}}{1+\beta_{Fs}}}{\frac{\beta_{Ms}}{1+\beta_{Ms}}}\right)$. Thus, any equation trying to estimate log individual wages must be consistent with the wage premium equation in Equation 1.14 when taking the difference between women and men across time.

I assume the USES placements represent a plausibly exogenous shock to the relative labor-augmenting productivity for women in the wage premium equation by changing the relative labor demand for women. I utilize the USES placements to construct measures of changes in the industry and state-wide demand and specifically, the relative industry and state-wide demand for women during World War II, and test the impacts on relative wages for women (compared to men) in 1950 (as compared to 1940).

I revisit the empirical specification used by Acemoglu et al. (2004) and allow for industrial

variation using a very saturated difference-in-differences style model of wages as follows:[12]

$$lnw_{ijst} = \alpha_1 f_i + \alpha_2 d_{1950} + \alpha_3(f_i \times d_{1950}) + \alpha_4 \textbf{industry}_j + \alpha_5 \textbf{state}_s + \alpha_6(f_i \times \textbf{industry}_j) +$$
$$\alpha_7(f_i \times \textbf{state}_s) + \alpha_8(\textbf{state}_s \times d_{1950}) + \alpha_9(f_i \times \textbf{state}_s \times d_{1950}) + \mathbf{X}_i' \nu_t^g +$$
$$\chi ln\left(\frac{F_{jst}}{M_{jst}}\right) + \eta\left(f_i \times ln\left(\frac{F_{jst}}{M_{jst}}\right)\right) + \gamma_1(\text{Total Ind. Demand Shock}_j \times d_{1950})$$
$$+ \gamma_2(f_i \times \text{Relative Female Industry Demand Shock}_j \times d_{1950}) + u_{ijst} \quad (1.15)$$

where the dependent variable lnw_{ijst} is the natural log of an individual i's total annual wages in industry group j, in state s in year $t = \{1940, 1950\}$.[13] f_i is a dummy $= 1$ if the individual i is female, d_{1950} is a dummy $= 1$ if the year $= 1950$, and $\textbf{industry}_j$ and $\textbf{state}_s$ are vectors of industry and state fixed effects respectively.[14]

I allow wages to be different by gender-by-state-by-year, as well as industry-by-year. For example, if we believe that military manpower mobilization caused differential impacts on wages

[12]In comparison, the estimating equation used by Acemoglu et al. (2004) is:

$$lnw_{ist} = \alpha_1 f_i + \alpha_2 d_{1950} + \alpha_5 \textbf{state}_s + \mathbf{X}_i' \nu_t^g + \chi ln\left(\frac{F_{st}}{M_{st}}\right) + \eta\left(f_i \times ln\left(\frac{F_{st}}{M_{st}}\right)\right)$$

where they calculate F_{st} as the average weeks worked by a woman in state s and year t, and M_{st} as the average weeks worked by a man in state s and year t.

They instrument for $ln\left(\frac{F_{st}}{M_{st}}\right)$ using the military mobilization rate times a dummy for 1950 (d_{1950}), and for $\left(f_i \times ln\left(\frac{F_{st}}{M_{st}}\right)\right)$ using the military mobilization rate times d_{1950}, interacted with the female dummy (f_i).

Thus, compared to Acemoglu et al. (2004), my specification is more flexible, and allows for variation in wages at the gender-by-industry level, gender-by-year level and gender-by-state-by-year level. It also notably does not use military mobilization as the instrumental variable due to the findings in Rose (2018), and corrects for the flaws in the weeks worked measure by using employment in number of women and men, and wages in annual wages.

[13]I also re-estimate my specifications with the hourly wage rate instead of annual wages, but that measure is also imperfect. Other than the flawed definition of the weeks worked measure, the hours worked may also be flawed since the 40-hour workweek was passed by the Congress in June 1940. I also estimate my results with and without including employment. For more, see subsection 1.6.4

[14]The baseline state is Pennsylvania and the baseline industry is Machinery for Durable Manufacturing and Paper for Non-Durable Manufacturing.

for women compared to men across states between 1940 and 1950 as per Acemoglu et al. (2004), including state-by-year-by-gender fixed effects would account for these changes.

$\mathbf{X}_i'$ is a vector of individual-level controls that include marital status, race, years of completed education, number of children in the household under age 5, and a fourth-degree polynomial in potential experience. The coefficients on these individual controls are allowed to vary by both gender and year.[15]

The coefficient χ on $ln\left(\frac{F_{jst}}{M_{jst}}\right)$ measures the impact of contemporaneous relative employment in an industry-state-year on both male and female wages. The coefficient η on $\left(f_i \times ln\left(\frac{F_{jst}}{M_{jst}}\right)\right)$ measures the additional impact of contemporaneous relative employment in an industry-state-year on female wages relative to male wages. Employment is measured as the natural log of the ratio of the total number of women in an industry j in state s in year t to the total number of men in an industry j in state s in year t.

Although I use a very saturated model to estimate the wage changes in Equation (1.15), we might be worried about simultaneity bias from having contemporaneous relative employment as an explanatory variable in an equation with wages as the dependent variable. To correct for this, I instrument for contemporaneous employment using the relative employment in an industry-state in 1930 as per the 1930 Decennial Census. Since I want to estimate changes in relative labor-augmenting productivity between 1940 and 1950, I want to use instruments for employment that predate 1940. I instrument for contemporaneous log relative employment, $ln\left(\frac{F_{jst}}{M_{jst}}\right)$, with $ln\left(\frac{F_{js,1930}}{M_{js,1930}}\right)$, i.e. the natural log of the ratio of the total number of women in an industry j in state s in 1930 to the total number of men in an industry j in state s in year 1930. I also

[15]The baseline person is a white, married man with no children in the household below the age 5, no potential experience and 12 or more years of education.

control for the interaction of contemporaneous log relative employment and the female dummy, $\left(f_i \times ln\left(\frac{F_{jst}}{M_{jst}}\right)\right)$, with the log relative employment in 1930 interacted with the female dummy $\left(f_i \times ln\left(\frac{F_{js,1930}}{M_{js,1930}}\right)\right)$. This is my preferred final specification, but I also report estimates without including employment at all in Section 1.6.4.

When I take the difference of my specification for log wages in Equation 1.15 between women and men and between 1950 and 1940, I have a resulting wage premium equation that looks like

$$\Delta_{1940-50} ln\left(\frac{w_{js}^F}{w_{js}^M}\right) = \hat{\alpha_3} + \hat{\alpha_9}(\textbf{state}_s) + \mathbf{X}_i'\left(\Delta_{1940-50}\hat{\nu^F} - \Delta_{1940-50}\hat{\nu^M}\right)$$
$$+ \hat{\eta}\Delta_{1940-50}\left(ln\left(\frac{F_{js}}{M_{js}}\right)\right) + \hat{\gamma_2}(\text{Relative Female Industry Demand Shock}_j) \quad (1.16)$$

Comparing this to our wage premium in Equation 1.14, we can see that $\hat{\alpha_3}$ and $\hat{\alpha_9}$ estimate the differential impact on relative wages by gender-across-time and gender-by-state-across-time respectively, accounting for changes in the relative labor supply elasticities β_{st}^G across time due to factors like inter-state migration costs or housing availability that are different by gender. The difference in the coefficients of individual X's $\left(\Delta_{1940-50}\hat{\nu^F} - \Delta_{1940-50}\hat{\nu^M}\right)$ also capture some of the impacts of changing relative labor-supply elasticities.

The coefficient $\hat{\eta}$, if we are assured is estimated accurately, can help us recover the CES rate of substitution ρ as $\rho = 1 + \hat{\eta}$.

Finally, I capture some of the impacts of changes in industry-level relative labor-augmenting productivity across time via $\hat{\gamma_2}$, which is my main coefficient of interest. This means that given my assumption that relative changes in labor augmenting productivity are linearly additive in

state-year and industry-year separately, I am only able to observe the impact on wages via industrial variation.

If I make a stricter assumption restricting the labor-responsiveness of individuals to wages to just differ across gender as β^G (instead of varying across time and state as β^G_{st}), then I only need gender-by-year fixed effects to capture these changes and can use state-level measures of relative and absolute demand changes during World War II to capture the state-level impact of labor-augmenting productivity changes on wages. This would be a slight variation of the first specification as

$$lnw_{ijst} = \alpha_1 f_i + \alpha_2 d_{1950} + \alpha_3(f_i \times d_{1950}) + \alpha_4 \textbf{industry}_j + \alpha_5 \textbf{state}_s + \alpha_6(f_i \times \textbf{industry}_j) +$$
$$\alpha_7(f_i \times \textbf{state}_s) + \mathbf{X}_i' \nu_t^g + \chi ln\left(\frac{F_{jst}}{M_{jst}}\right) + \eta\left(f_i \times ln\left(\frac{F_{jst}}{M_{jst}}\right)\right) +$$
$$\gamma_1(\text{Total Industry Demand Shock}_j \times d_{1950}) + \gamma_2(f_i \times \text{Relative Female Ind. Demand Shock}_j \times d_{1950}) +$$
$$\gamma_3(\text{Total State Demand Shock}_s \times d_{1950}) + \gamma_4(f_i \times \text{Relative Female State Demand Shock}_s \times d_{1950})$$
$$+ u_{ijst} \quad (1.17)$$

which translates to the wage premium equation

$$\Delta_{1940-50} ln\left(\frac{w_{js}^F}{w_{js}^M}\right) = \hat{\alpha_3} + \mathbf{X}_i'\left(\Delta_{1940-50}\hat{\nu^F} - \Delta_{1940-50}\hat{\nu^M}\right) + \hat{\eta}\Delta_{1940-50}\left(ln\left(\frac{F_{js}}{M_{js}}\right)\right)$$
$$+ \hat{\gamma_2}(\text{Relative Female Ind. Demand Shock}_j) + \hat{\gamma_4}(\text{Relative Female State Demand Shock}_s)$$
$$(1.18)$$

This specification allows me to test the impact of relative state-level shocks in the labor

demand for women during World War II on relative wages for women in 1950 via $\hat{\gamma_4}$ and therefore I can capture both the state-level as well as the industry-level evolution in the relative labor augmenting productivities. In this specification, the labor-supply responsiveness is only captured by gender-across-time effects via $\hat{\alpha_3}$ and $\left(\Delta_{1940-50}\hat{\nu^F} - \Delta_{1940-50}\hat{\nu^M}\right)$.

Finally, I can also make my initial specification in Equation 1.15 more flexible, in case there are other industry-wide trends in wages between 1940 and 1950 that are not fully captured by just changes in industrial wartime demand. So, I also estimate the specification

$$\begin{aligned} lnw_{ijst} = \alpha_1 f_i + \alpha_2 d_{1950} + \alpha_3(f_i \times d_{1950}) + \alpha_4 \textbf{industry}_j + \alpha_5 \textbf{state}_s + \alpha_6(f_i \times \textbf{industry}_j) + \\ \alpha_7(f_i \times \textbf{state}_s) + \alpha_8(\textbf{state}_s \times d_{1950}) + \alpha_9(f_i \times \textbf{state}_s \times d_{1950}) + \alpha_{10}(\textbf{industry}_j \times d_{1950}) + \\ \mathbf{X}_i' \nu_t^g + \chi ln\left(\frac{F_{js}}{M_{js}}\right) + \eta\left(f_i \times ln\left(\frac{F_{jst}}{M_{jst}}\right)\right) + \\ \gamma_2(f_i \times \text{Relative Female Ind. Demand Shock}_j \times d_{1950}) + u_{ijst} \end{aligned} \quad (1.19)$$

which gives me the effective wage premium equation

$$\begin{aligned} \Delta_{1940-50} ln\left(\frac{w_{js}^F}{w_{js}^M}\right) = \hat{\alpha_3} + \hat{\alpha_9}(\textbf{state}_s) + \hat{\alpha_{10}}(\textbf{industry}_j) + \mathbf{X}_i'\left(\Delta_{1940-50}\hat{\nu^F} - \Delta_{1940-50}\hat{\nu^M}\right) \\ + \hat{\eta}\Delta_{1940-50}\left(ln\left(\frac{F_{js}}{M_{js}}\right)\right) + \hat{\gamma_2}(\text{Relative Female Ind. Demand Shock}_j) \end{aligned} \quad (1.20)$$

where once again I am only able to isolate the industrial impact of changes in relative labor-augmenting productivities through the relative changes in the demand for women in an industry via $\hat{\gamma_2}$. I present all three specifications in my main tables, and prefer the final most flexible specification outlined above in Equation 1.19 as the most accurate.

I estimate the equations separately for workers in Durable Manufacturing and Non-Durable Manufacturing since past literature shows that they evolved differently during World War II (Shatnawi and Fishback, 2018). Rose (2018) finds gains in employment in Durable Manufacturing industries in 1950 but losses in Non-Durable Manufacturing industries because of increased wartime female employment, suggesting Durables could have experienced a persistent increase in the relative demand for women, while Non-Durables did not.

I define the total shock in the demand for workers as the percent change in total employment in an industry (or state) between 1940 and 1946, assuming that the only increases in employment come from the USES assignments that we are able to observe. Thus, if we denote total workers in an industry (or state) in 1940 as L_{40}, and the USES assignments of workers in an industry (or state) as L_{USES}, then the 'Total Shock' measure is equal to the total number of workers assigned to an industry (or state) by the USES, divided by the number of workers in that industry (or state) in the 1940 Census:

$$\text{Total Shock} = \frac{(L_{46} - L_{40})}{L_{40}} = \frac{L_{USES}}{L_{40}} \tag{1.21}$$

There is no pre-established way to define the 'Relative Demand Shock for Women'. Given that I want to measure the relative change in employment of women induced by the war effort compared to the overall change in labor demand due to the war effort, I define the 'Relative Demand Shock for Women' using two different measures. I want both measures to reflect the impact of a relative change in the demand for women during the war, which affects relative labor-augmenting productivity specifically for women and impacts wages beyond any overall industry-level shift in labor-augmenting productivity for all workers that might come from wartime capital investment, for example. I define both measures explicitly for all manufacturing industries in

Table 1.2.

The first measure I define is the relative percent change in female employment between 1940 and 1946 in an industry (or state), assuming that the only increases in employment come from the USES assignments that we are able to observe. This means that I divide the percent change in female employment in an industry (or state) from USES assignments in WWII with the percent change in overall employment from USES assignments in an industry (or state). Defining the number of women in an industry (or state) in 1940 as W_{40} and the number of female assignments in an industry (or state) in the USES data as W_{USES}, this measure is operationalized as

$$\text{Measure 1} : \%\Delta_{WWII}\text{Relative Number of Women} = \frac{\frac{(W_{46} - W_{40})}{W_{40}}}{\frac{(L_{46} - L_{40})}{L_{40}}} = \frac{\left(\frac{W_{USES}}{W_{40}}\right)}{\left(\frac{L_{USES}}{L_{40}}\right)} \tag{1.22}$$

In words, this translates to the percent change in the number of women due to USES assignments, divided by the percent change in the number of all workers due to USES assignments in an industry (or state) between 1940 and 1946.

This measure utilizes the absolute shock felt by the wartime assignment of women to different industries but moderates it via both the number of women pre-existing in the industry in 1940 as well as the size of the wartime assignment shock of all workers relative to their 1940 level. For example, if we look at Table 1.1, we can see that the most women assigned by USES during WWII are the 237,574 women assigned to Food and Beverages. The third highest number is the 179,947 women assigned to Transport Equipment. However, we expect the impact of these extra women to be very different in Food and Beverages, which already was comprised of nearly

20% women in 1940 versus Transport Equipment, which had previously been less than 10% female in 1940. The size of the relative wartime demand shock seen by Transport Equipment is nearly twice the size of the shock to Food and Beverages, even though the absolute increase in women in Transport and Equipment was marginally smaller than Food and Beverages. When looking at actual values in Table 1.2, we see that the female employment shock only translated to a 7% increase in the number of women in Food and Beverages in column (5), as opposed to the 15.6% increase in the number of women in Transport Equipment. When further moderating by the size of the shock for all workers in column (6), we find that the relative shock in the number of women in Food and Beverages was 1.2 compared to 2.1 for Transport Equipment.

The second measure I use to operationalize the relative demand change for women is the percent change in the proportion of an industry (or state) that is female between 1940 and 1946, assuming that any changes in employment happened due to the observable USES assignments. Defining the proportion of women in an industry (or state) in 1940 and 1946 as

$$P_{40} = \frac{W_{40}}{L_{40}}; P_{46} = \frac{W_{40} + W_{USES}}{L_{40} + L_{USES}}$$

this measure is operationalized as

$$\text{Measure 2} : \%\Delta_{WWII}\text{Relative Proportion of Women} = \frac{(P_{46} - P_{40})}{P_{40}} \tag{1.23}$$

This may be the more relevant relative demand shock if we think that relative wages for women changed (and relative labor-augmenting productivity was updated) through changes in the industry gender composition as a consequence of the war effort. Therefore, it is not about

industries that saw the largest number of women join the industry but rather how the new women fit into the gender composition of an industry. Again, looking at Table 1.2, this means that although Food and Beverages saw the largest number of women assigned to the industry by USES, it only made the industry go from being 19.4% female in 1940 (in column (2)) to 19.7% female in 1946 (in column (4)), a change of 1.1% (in column (7)). In contrast, Transport Equipment, which saw a smaller number of women assigned by USES, went from being 9.9% female in 1940 (in column (2)) to being 15.6% female in 1946 (in column (4)), nearly a seven times larger difference in composition of 7.5% (in column (7)).

I compare both measures in Figure 1.3. Both measures show little variation in the Non-Durable Manufacturing industries. This reflects that while some of the Non-Durable Manufacturing industries saw large absolute numbers of assignments by USES during WWII both in terms of total employment as well as female employment, they did not see significantly different impacts on relative female demand. On the other hand, among Durable industries, there is significant variation in relative female demand. Some industries like Electrical Machinery and Transport Equipment see large increases in the relative demand for women, while others like Fabricated Metals, Instruments, and Lumber see much smaller increases. Notably, both Fabricated Metals and Transport Equipment saw large increases in overall demand due to World War II in Table 1.1, but only Transport Equipment saw a large increase in the relative demand for women. This is indicative of the differences in the utilization of women across different industries even within manufacturing sectors.

1.6 Results

1.6.1 Prior Findings

Before diving into my results, I want to briefly outline the main findings of two papers that most informed my analysis — Acemoglu et al. (2004) and Rose (2018).

Acemoglu et al. (2004) use state-wide variation in the military mobilization of men during World War II to test impacts on female employment and relative wages. Setting up a difference-in-differences model across states between 1940 and 1950, they find that states with higher military mobilization of men saw increased female labor force participation (using weeks worked) in 1950 as compared to 1940, but they do not find any statistically significant impact on the labor force participation of men.

Thus, utilizing military mobilization as an instrument for changes in the relative employment of men and women between 1940 and 1950, they estimate a relative wage equation for a pooled sample of male and female workers in 1940 and 1950, accounting for differences by gender, state, year, and individual characteristics. They find that both male and female wages fell in response to military mobilization, hypothesizing that men and women were imperfect substitutes during this time.

Rose (2018) points out two important flaws in Acemoglu et al. (2004). First, he points out that the main outcome variable used to measure employment effects in Acemoglu et al. (2004) ('weeks worked') changed definitions between the 1940 and 1950 censuses in a way that could artificially inflate employment for women in 1950 compared to 1940. Second, using new data from World War II, he notes that military mobilization does not have any impact on female

employment during the war, thereby casting doubt on the link found in Acemoglu et al. (2004) between military mobilization and the employment of women in 1950. Rose (2018) highlights instead the relevance of state-economic-level World War II production demand intensities in shifting female wartime employment.

Despite having data on the wartime employment of women and men by industry, Rose (2018) only chooses to construct and use state-wide (or state-economic-area wide) measures of wartime employment.[16] Additionally, while he disproves the findings of Acemoglu et al. (2004) on employment of women in 1950, Rose (2018) stops short of estimating the impact of female employment during World War II on relative wages in 1950. However, if the story behind the increase in female employment during World War II is that of an increase in demand in war-critical industries, then we might also expect to find the relative wages rise for women, in contrast to Acemoglu et al. (2004)'s findings.

Thus, in this chapter, I attempt to close the gap between these two works and test for the impacts of World War II demand for women on the relative wages of women in 1950. As well as testing the impacts on relative wages, I also introduce a dimension of demand changes previously unexplored at the national level in the literature — the variation of wartime demand by industry. State markets may not have been as binding during World War II in the face of increased national migration and hiring of workers across states for war-critical industries. Hence, industrial variation in demand for workers may be the more relevant driver of employment (and relative wage) shifts.

[16] In Table 6 of his paper, Rose (2018) delves into a little cross-industry variation. While looking at the persistence of impacts of state-economic-area wide female employment during World War II on female employment in 1950, he finds no statistically significant impacts. However, when testing these same impacts separately by industry, he finds that there are persistent gains in employment for women in Durable Manufacturing in 1950 and declines for women in Non-Durable Manufacturing.

1.6.2 Impacts of Changes in the Relative Demand for Women

1.6.2.1 Using Measure 1: A Change in the Relative Number of Women

In Table 1.4, I present results for the impact of World War II employment shocks on wages for women (and men) in 1950 (compared to 1940) using my first constructed measure of relative female demand—the percent change in the number of women due to USES assignments relative to the percent change in the number of all workers.[17] In columns (1) and (4) of this table, I present the results of Equation 1.17 from Section 1.5 for Durables and Non-Durables respectively. In this specification, I do not allow for industry, state, or state-by-gender trends. Therefore, I can measure the impact of relative demand shocks in the USES data at the state level as well as the industry level. I find that in column (1), there is a statistically significant and positive increase of 0.176 in the wages for women (compared to men) in 1950 in Durable Manufacturing due to a unit increase in the relative female demand measure.

To interpret my point estimates from Table 1.4, I use the range of values this relative demand measure can take, listed in column (6) of Table 1.2. This relative demand measure ranges from 0.783 in Lumber and Wood to 2.062 in Transport Equipment for women in Durable Manufacturing (Panel B). Consequently, moving from the industry with the lowest change in the relative demand for women to the industry with the highest change in the relative demand for

[17] From section 1.5, if we define the number of women in an industry (or state) in year t as W_t, the number of all workers in an industry (or state) in year t as L_t, the number of female assignments in an industry (or state) in the USES data as W_{USES}, and the number of all assignments in an industry (or state) in the USES data as L_{USES}, this measure is equal to

$$\text{Measure 1} : \%\Delta_{WWII}\text{Relative Number of Women} = \frac{\frac{(W_{1946} - W_{1940})}{W_{1940}}}{\frac{(L_{1946} - L_{1940})}{L_{1940}}} = \frac{\left(\frac{W_{USES}}{W_{1940}}\right)}{\left(\frac{L_{USES}}{L_{40}}\right)}$$

women is a shift of 1.279 ($= 2.062 - 0.783$); and a move from the industry with the median change in the relative demand for women (Stone, Clay, and Glass, with a value of 1.440) to the highest change industry, is a shift of 0.622 ($= 2.062 - 1.440$).

Returning to Table 1.4, this implies that moving from the industry with the lowest change in the relative female demand to the industry with the highest change in the relative female demand in Durable Manufacturing increased wages for women (relative to men) by 0.225 ($= 0.176 \times 1.279$), or 22.5% in 1950 according to column (1). Being in the highest change industry compared to the median change industry in Durables saw wages increase by 0.109 ($= 0.176 \times 0.622$), or 10.9% for women in 1950 (relative to men).

The impact of an increase in relative female demand on wages for women in Non-Durable Manufacturing in column (4) on the other hand, is negative and not statistically significant. There are also no statistically significant impacts of changes in the relative demand for women or the absolute demand for all workers at the state level on wages of men or women in Durable or Non-Durable Manufacturing. This is notable since most prior literature focuses on state-wide variation in employment during World War II to measure the impacts on labor market outcomes of women. From my findings, it appears that industrial variation is the more relevant channel when looking at wages as an outcome.

One last thing we want to note from columns (1) and (4) of Table 1.4 is that there are no statistically significant impacts of changes in the overall industry-wide demand for all workers.[18] We might have expected changes to production efficiency in industries that saw large increases

[18]To repeat, this measure is defined as

$$\text{Total Shock} = \frac{(L_{46} - L_{40})}{L_{40}} = \frac{L_{USES}}{L_{40}}$$

in overall labor, and perhaps a subsequent impact on relative wages. But if these did occur, I do not find that the impacts persisted through to relative wages in 1950.

Next in columns (2) and (5), I present results from Equation 1.15 for Durable and Non-Durable industries respectively, in which I include fixed effects by state-by-gender-by-year.

If we believe the results from Acemoglu et al. (2004), and think that military mobilization increased relative female labor supply in highly mobilized states, we want to also account for possible impacts on wages for women. By including state-by-gender effects, I control for any channels affecting female wages differently by state. I now find that the impact of a unit increase in the relative demand for women during World War II increases wages for women by a slightly larger magnitude of 0.191, which is still statistically significant. Therefore, going from the Durable Manufacturing industry with the lowest change in the relative female demand to the industry with the highest change in the relative female demand increases wages for women (relative to men) by 0.244 ($= 0.191 \times 1.279$), or 24.4% in 1950. Being in the highest change industry compared to the median change industry in Durables now sees a rise of 0.119 ($= 0.176 \times 0.622$), or 11.9% in female wages in 1950 (relative to men).

The impact of an increase in the relative female demand on wages for women in Non-Durables in column (5) is negative and now also statistically significant at the 5% level. Wages decrease by 0.262 in 1950 (relative to men) for a unit increase in relative female demand. Referring to Table 1.2 once again, this translates to a drop in the wages of women (relative to men) in 1950 of 0.147 (or 14.7%) when going from the Non-Durables industry with the lowest change in the relative demand for women (Tobacco, with 1.017) to the industry with the highest change in the relative demand for women (Leather and Footwear, with 1.577).

After controlling for state trends, the overall industry-wide demand for all workers during

World War II is still not statistically significant in changing male or female wages in Durables or Non-Durables.

Finally, results from the most flexible specification of my estimating equation 1.19, where I control for state-by-gender trends as well as industry trends, are reported in columns (3) and (6) of Table 1.4 for Durable and Non-Durable Manufacturing industries respectively. In column (3), I find that a unit increase in the relative demand for women increases female wages by 0.277 in 1950. This translates to a change of 0.354 ($= 0.277 \times 1.279$) or 35.4% when going from the Durable Manufacturing industry with the lowest change in relative female demand to the Durable Manufacturing industry with the highest change in relative female demand during World War II. Being in the highest change industry compared to the median change industry in Durables now sees a rise of 0.172 ($= 0.277 \times 0.622$), or 17.2% in female wages in 1950 (relative to men). The impact of changes in relative female demand on female wages in Non-Durable Manufacturing in column (6) is a very small and statistically insignificant point estimate of -0.030.

1.6.2.2 Using Measure 2: A Change in the Proportion of Women

In Table 1.5, I repeat the exercise from the previous subsection using my second constructed measure of relative female demand — the percent change in the proportion of women.[19]

In columns (1) and (4), I present the results of estimating Equation 1.17 from Section 1.5

[19]Defining the proportion of women in an industry (or state) in 1940 and 1946 as

$$P_{40} = \frac{W_{40}}{L_{40}}; P_{46} = \frac{W_{40} + W_{USES}}{L_{40} + L_{USES}}$$

this measure is operationalized as

$$\text{Measure 2} : \%\Delta_{WWII}\text{Relative Proportion of Women} = \frac{(P_{46} - P_{40})}{P_{40}}$$

for Durables and Non-Durables respectively, using this new measure of relative demand. Before discussing the results, I once again refer back to Table 1.2 to make sense of the magnitudes.

In column (7) of Table 1.2, I report the range of changes in the relative proportion of women. In Durable Manufacturing Industries (Panel B), it ranges from -0.011 (or -1.1%) in Fabricated Metal to 0.075 (or 7.5%) in Transport Equipment. Thus, moving from the industry with the lowest change in the relative proportion of women to the industry with the highest change in the relative proportion of women is a shift of 0.086 ($= 0.075 - (-0.011)$); and a move from the median change industry (Machinery with a change of 0.021) to the highest change industry is a shift of 0.054 ($= 0.075 - 0.021$).

Once again returning to our Table 1.5, the statistically significant and positive increase of 3.174 we see in column (1) implies an increase by 0.273 ($= 3.174 \times 0.086$), or 27.3% in the wages for women (compared to men) in 1950 when going from the Durable Manufacturing industry with the lowest change in the relative proportion of women during World War II to the Durable Manufacturing industry with the highest change. Being in the highest change industry compared to the median change industry in Durable Manufacturing has relative wages for women rise by 0.171 ($= 3.174 \times 0.054$), or 17.1% in 1950 (relative to men).

Just like we saw with our previous measure of relative demand, the impact of an increase in relative female demand on wages for women in Non-Durable Manufacturing in column (4) is still negative and statistically insignificant.

In this specification, I do not have industry or state-by-gender trends which means I can test the impact of relative demand shocks in the USES data at the state level as well. As with our previous demand measure, there are no statistically significant impacts of changes in the relative demand for women or the absolute demand for workers at the state level on wages for men or

women in Durable or Non-Durable manufacturing, as well as no statistically significant impacts of changes in the overall industry-wide demand for workers on wages for men or women in either manufacturing sector.

Next, in columns (2) and (5), I present results from estimating Equation 1.15 for Durable and Non-Durable Manufacturing industries respectively, using my second measure of relative demand change. In this specification compared to those in columns (1) and (4), I include fixed effects by state-by-gender-by-year. I now find that the impact of a unit increase in the relative female proportion on the relative wages for women in 1950 in Durable Manufacturing is still statistically significant and positive, while slightly larger, at a magnitude of 3.395; which translates to 0.292 ($= 3.395 \times 0.086$) or a 29.2% increase in the wages for women (relative to men) in 1950 going from the Durable Manufacturing industry with the lowest change in the relative proportion of women during World War II to the industry that saw the highest change. Being in the highest change industry compared to the median change industry in Durables now sees a rise of 0.183 ($= 3.395 \times 0.054$), or 18.3% in female wages in 1950 (relative to men).

The impact of an increase in relative female demand on wages for women in Non-Durables in column (5), is negative and statistically insignificant. Once again, the impact of changes in the overall industry-wide demand during World War II is still not statistically significant for either Durables or Non-Durables in columns (2) and (4) respectively.

Finally, I report results from my preferred specification of the estimating equation 1.19, where I control for state-by-gender trends as well as industry trends, in columns (3) and (6) of Table 1.5 for Durable and Non-Durable Manufacturing industries respectively. In column (3), I now find that a unit change in the relative proportion of women increases female wages by a statistically significant magnitude of 4.178 relative to men in 1950. This translates to an increase

of 0.359 ($= 4.178 \times 0.086$) or 35.9% in relative female wages in 1950 when going from the Durable Manufacturing industry with the lowest change in the relative proportion of women to the industry with the highest change in the relative proportion of women during World War II. Being in the highest change industry compared to the median change industry in Durables now sees a rise of 0.226 ($= 4.178 \times 0.054$), or 22.6% in female wages in 1950 (relative to men). The impact of relative female demand changes on female wages in Non-Durable Manufacturing in column (6) is now positive in magnitude, but still statistically insignificant.

Overall, these results show that changes in the relative demand for women by industry during World War II (as measured by USES assignments) increased wages for women relative to men in 1950 by 35.4–35.9% in my final preferred specification, in Durable Manufacturing, when going from the industry least impacted by relative female demand changes to the industry most changed. In contrast, there seem to be no statistically significant impacts of changes in relative female demand by industry during World War II for women in Non-Durable Manufacturing. Total changes in demand for all workers by state or industry, or state-wide variation in the relative demand for women do not seem to be a significant driver of wage changes.

1.6.2.3 Heterogeneous Impacts By Demographics in Durable Manufacturing Industries

Goldin and Olivetti (2013) conclude in their paper that the largest gains in female employment from World War II military mobilization accrued to white, married women with 12 or more years of education. They hypothesize that less-educated women could have disproportionately worked in manufacturing which saw many jobs disappear post-war, but more-educated women were able

to enter sectors that allowed them to stay in the labor force post World War II. Although they frame this as a change in female labor supply across the two groups, this would not be inconsistent with female demand increasing more for white, married, and more-educated women than less-educated women.

I do not have USES assignments by race, marital status, or level of education, so I cannot check which subgroups of women (and men) were pulled into different manufacturing industries during the war effort. However, I can estimate the impacts of changes in wages for different subgroups of workers in 1950 as a result of relative demand changes for all women during World War II across industries in Durable Manufacturing.

In Table 1.6, I present results of Equation 1.19 (my preferred specification including state-by-gender trends and industry trends) separately for different sub-populations of workers in Durable Manufacturing. In columns (1)–(4), I use my first measure of a change in the relative demand for women: the change in the number of women in an industry relative to all workers, and in columns (5)–(8) I use my second measure: a relative change in the proportion of women.

I find that the largest relative wage gains accrue to women with a high school education or more, irrespective of marital status. In column (1) of Panel A, married women with a high school education or more gain a statistically significant magnitude of 0.436 in relative wages, which translates to a 55.8% ($= 0.436 \times 1.279$) increase in relative wages when going from the lowest change industry to the highest change industry in Durable Manufacturing. Unmarried women with a high school education gain a slightly larger magnitude of 0.487 in relative wages, which translates to an increase of 62.3% ($= 0.487 \times 1.279$) when going from the lowest change industry to the highest change industry in Durable Manufacturing in column (2) of Panel A. While they seem very similar, the magnitudes in columns (1) and (2) of Panel A are statistically significantly

different from one another.

Columns (5) and (6) provide similar estimates using my second measure of relative demand: a change in the relative proportion of women. Married women with a high school education gain 6.511 in wages in column (5), which translates to a 56% (6.511×0.086) increase in wages when going from the lowest demand industry to the highest demand industry in Durable Manufacturing. Unmarried women with a high school education in column (6) gain 6.020 in wages, which translates to an increase of 51.2% in relative wages when going from the lowest demand industry to the highest demand industry in Durable Manufacturing. Once again, albeit similar, the impacts in columns (5) and (6) are statistically significantly distinct from one another.

When looking at the impacts of changing relative demand for women on those with less than a high school education in columns (3) and (4) and (7) and (8), the relative wage gains are much smaller than for women with a high school education and no longer statistically significant using either measure of relative demand.

Similar patterns hold when restricting my sample to white women in Panel B. Gains seen by women with a high school education or higher range from 42.7-43.7% for married women in columns (1) and (5), and from 50.2-59.8% for unmarried women in columns (2) and (6). All of these gains are significantly larger than the gains estimated for the full sample of women in Tables 1.4 and 1.5 of approximately 35%.

Thus, relative wages in 1950 rose more sharply for women with a high school education or higher, as a result of being in high-demand industries during World War II, perhaps suggesting a larger increase in their relative labor-augmenting productivity across industries. Many new jobs for women created within manufacturing industries during World War II required training prospective employees in smaller ”unit skills” to be done in an assembly line rather than invest

in a longer period of training to replace a fully skilled worker (USES, 1942). It is plausible that women with a high school education or higher were easier to train, and thus saw the greatest rise in demand reflected in relative wage gains. It could also be that the Beckerian employer costs of hiring women were lower for more educated women and so employer notions of perceived productivity were revised faster for these women, increasing their wage gains, and changing their relative demand in Durable Manufacturing by giving them a "foot in the door".

On the other hand, in contrast to Goldin and Olivetti (2013), I find that white married women did not see larger gains in wages in 1950 than unmarried white women, irrespective of their education level. This might be because many married women left the workforce soon after the war ended (Goldin, 1991a), and possible relative wage gains did not persist until 1950. Marriage bars were legally abolished during World War II for employers to avail themselves of a larger available female workforce for the war effort but may have been informally re-instated post-war(Goldin, 1991b). When the war ended, cultural norms encouraging married women to return to the household may have driven married women out of the workforce permanently post-war (Doepke et al., 2015; Larsen et al., 2015; Campbell, 1984).

These results help to reframe the results from Goldin and Olivetti (2013) who find that white, married women with a high school education or higher saw the strongest persistent gains in employment as a result of World War II military manpower mobilization, and attribute it to a change in the female labor supply of these women. By finding that wages changed more significantly for high-school-educated women, I think that it is likely a shift in the relative labor demand for these women that increased, particularly within the Durable Manufacturing industries.

Apart from the differences in the persistence of changes in the relative demand for women

during World War II until 1950 between Durable and Non-Durable Manufacturing, there are possibly other differences in the way that the two sectors developed between 1940 and 1950. One advantage of estimating a saturated wage growth equation like I do in Section 1.5 is that I can use it to note other interesting deviations in the evolution of wages of women and men between 1940 and 1950 across Durable and Non-Durable Manufacturing industries.

1.6.3 Differences Between Durable and Non-Durable Manufacturing Industries

Before looking at the wage growth equations once again, I note some important differences between Durable and Non-Durable Manufacturing Industries in the raw data. In Table 1.3, we see very few distinct differences between workers in Durable and Non-Durable manufacturing when pooling workers from 1940 and 1950. Workers in Durable Manufacturing tend to have marginally higher wages (whether measured annually or hourly) but strikingly, the one noticeable difference is that half as many women are working in Durables compared to Non-Durable Manufacturing industries.

In Figure A.2, I plot the changes in log average annual wages in an industry between 1940 and 1950 for women (by diamonds) and men (by circles) on the y-axis, plotted against the differences in the log total number of women employed (by the diamonds) and men employed (by the circles) by industry on the y-axis between 1940 and 1950, with the size of each diamond and dot weighted by the number of workers of that gender in that industry across the two years. The top panel represents Durable Manufacturing Industries, and the bottom figure represents Non-Durable Manufacturing Industries. We see even in the raw data that the Durable and Non-Durable Manufacturing industries changed very differently between 1940 and 1950. The first thing we can

note is industries within Durable Manufacturing experienced much wider changes in employment (movements along the x-axis), for both men and women between 1940 and 1950 than industries in Non-Durable Manufacturing. Although some industries in Non-Durable Manufacturing see relative wages for women rise by more than for men (the diamonds lie above the circles of the same color), these industries do not also see gains in relative employment for women significantly outstrip that of men. On the other hand, within Durable Manufacturing, in Transport Equipment (in gold) and Electrical Machinery (in light brown), the relative gains in both employment and wages are larger for women than for men (the diamonds lie clearly to the top-right of the dots of the same color), strongly suggesting that these industries saw a significant change in the relative demand for women. Finally, Fabricated Metals (in brown) sees the relative demand increase for men, even though both men and women see gains in employment and wages, suggesting variation even within Durable Manufacturing industries in relative wage gains for women versus men. Stricter sex-typing of jobs within different industries in Durable Manufacturing, and sharp increases in demand in ”masculine” industries during World War II could perhaps explain the variability of industrial change within Durable Manufacturing in the growth of employment and wages for men and women (Milkman, 1982). Therefore, by not allowing for industry-level variation, an important part of the wage premium puzzle is missed. This also reassures us about the separate specification of the wage premia equations across Durable and Non-Durable Manufacturing industries.

In Tables 1.4 and 1.5, I focus on my strictest specification 1.19 in columns (3) and (6) across both tables, where I control for state-by-gender trends as well as industry trends between 1940 and 1950.[20]

[20]The coefficients on the parameters I discuss do not vary widely across the other specification columns in Table

To start, I note the difference in the coefficient on the female dummy variable f_i in the wage estimation equations between columns (4) and (6) of Tables 1.4 and 1.5. I find that the small proportion of women who worked in Durable Manufacturing industries saw a much larger 'female penalty' to their wages at baseline in 1940 (ranging from 60-62.35%) as compared to women in Non-Durable Manufacturing, where the women saw a smaller penalty to wages of 40.6–40.8%. Hence, women started off at a more disadvantaged position in Durables in 1940, not simply in terms of the relative proportion of women employed, but also in terms of their relative wages. The α_{Gjt} that I include to account for industry and sex-specific barriers to working in different industries in my model in Section 1.3, is likely to be much higher for women in Durable Manufacturing industries compared to Non-Durable Manufacturing industries prior to World War II. According to my results from Section 1.6.2, being in the industry that sees the greatest increase in demand within Durable Manufacturing during WWII (relative to the industry that sees the lowest change), increases women's wages relative to men in 1950 by 35.4–35.9%. This decreases their initial 'female penalty' in 1950 to nearly half the magnitude of 1940.

A quick back-of-the-envelope calculation shows that although women who worked in Durable Manufacturing started off at a higher disadvantage in 1940 than women who worked in Non-Durable Manufacturing in terms of relative wages, working in Durable Manufacturing industries with the largest increases in demand during World War II such as Transport Equipment or Electrical Machinery in 1950 could see their wages converge to (or even surpass) the relative wages for women in Non-Durable manufacturing.

Next, we see that although neither industry sees a statistically significantly different change in wages for women between 1940 and 1950 (the coefficient on the female time trend: $f_i \times d_{1950}$

1.4 or Table 1.5.

is not statistically significant), both men and women in Durables see their wages increase by 29.9–30.7%, nearly twice the increase seen by workers in Non-Durables of 14.8–14.9%, which is also only marginally statistically significant at the 10% level.

1.6.4 Robustness

1.6.4.1 Robustness to Other Industry-Wide Trends

Even in the strictest specification I use in Equation 1.19 (including industry and state-by-gender trends), we might still worry that the wage impacts of changes in relative female demand are picking up a spurious correlation in industry-level changes in relative wages for women. This is especially true since I do not have USES assignments data that differs at the state-industry cell level to allow me to control for industry-by-gender trends in wages.

To check if this is the case, I rerun my specification in Equation 1.19 to allow for differential trends in the wages of women by different 1940 industry-level characteristics and present these results in Table A.1. I calculate the share of workers in the industry that have 12 or more years of education in 1940, the share of workers that are non-white in an industry in 1940 and the share of female workers in an industry that are married in 1940, and allow them to impact female wages differently in 1950.

I find that within Durable Manufacturing, my estimates of the increase in relative wages for women in 1950 as a result of an increase in the relative female demand by industry during World War II are robust to the inclusion of other industry-level trends for women. The coefficients for the impact of changes in relative female demand on wages are very stable across columns (1)-(5) using either of the measures of constructed demand in Panel A and Panel B.

Once again in columns (5) - (8), we see that the impacts of relative female demand changes during World War II on the relative wages for women are limited to Durable Manufacturing industries. Within Non-Durables, my estimates are much noisier although largely not statistically significant at the 10% level.

1.6.4.2 Robustness With and Without Employment

I re-estimate my results with and without employment and report the results in Table A.2 for Durable Manufacturing and in Table A.3 for Non-Durable Manufacturing. Employment is measured as the log relative total number of all full-time female workers in an industry-state-year divided by the total number of all full-time male workers in an industry-state-year. I do not include employment in columns (1), (3), and (5); and I include and instrument employment with 1930 levels in columns (2), (4), and (6). The latter columns are the same as the ones presented in my main tables 1.4 and 1.5.

I present the results for specification 1.17 in columns (1) and (2) of both tables, and the results for specification 1.15 in columns (3) and (4) of both tables, where I include state trends by gender; and for specification 1.19 in columns (5) and (6) of both Tables, where I also include industry trends.

For Durables in Table A.2, I find that the increases in relative female wages in 1950 of changing relative demand are lower when not including employment across each of the three specifications, but they are not significantly different upon including employment.[21]

In Table A.3, where I present estimates for Non-Durables, I find that not accounting for

[21] I also estimate these specifications including employment but not instrumenting for it. The results are not presented here but my estimates of impacts on wages due to relative female demand changes are similar to the other columns presented here.

employment makes the impact of relative female demand change on wages larger and more negative on wages for women (compared to men) in 1950. In Panel A, the impacts of changes in the relative number of women employed by USES on relative female wages are negative and statistically significant in columns (1) and (3) when not including employment, but become smaller and less significant upon accounting for employment. In Panel B, the same pattern is true but the results are not statistically significant at the 10% level even when not accounting for employment.

1.6.4.3 Robustness Using Hourly Wages

In setting up my motivating model for estimation in Section 1.3, my main dependent variable is meant to capture returns to marginal productivity in a market with frictions. Thus typically, this would be better measured by the wage rate rather than the annual wage measure I use for my results.

I do this because of several data constraints mentioned earlier. The hourly wage rate is calculated by dividing annual wages by the product of weeks an individual reports working, and the hours they report working in the last week. The change in the definition of weeks worked between 1940 and 1950 (weeks worked were only expected to be full-time in 1940), could lead to over-reporting of weeks worked in 1950 by part-time workers. If we think that women or specifically women in manufacturing may be more likely to work part-time, then this would artificially deflate hourly wage rates for women compared to men by inflating the denominator of weeks worked.

The hours worked variable also asks individuals to report their hours worked in the last

week before the Census is carried out, and may not always be representative of the average hours worked by an individual in a typical week, thus papers estimating employment changes choose to focus on the weeks worked measure. Dividing my main dependent variable by two uncertain measures could increase measurement error in the hourly wage rates. Also, due to the changed definition of workweeks between the 1940 and 1950 censuses, to get the hourly wage rate in 1940, we would divide the annual wages by the weeks reported to be worked times 35 hours (which was considered full-time equivalent). In 1950, to get the hourly wage rate, we divide by weeks worked times hours reported worked last week. As previously discussed, the 40-hour week was legislated after the 1940 Census and may have primed respondents to list their hours worked as 40 in the 1950 Census rather than their real hours worked.

Lastly, as discussed in Section 1.4, the measures for relative demand I construct also use USES data on the number of workers, I do not have hours or weeks reported for them. Therefore the USES measures are more symmetric when I estimate employment in terms of the total number of workers rather than in weeks worked.

With all the caveats out of the way, I do estimate my main specifications with hourly wage rates (constructed as mentioned above) as the main dependent variable, and the log of the average weeks worked per woman in an industry-state-year divided by the average weeks worked per man in an industry-state-year as my measure of labor input, which I instrument for with 1930 levels in Tables A.4 and A.5.[22]

Across both measures of relative female demand I use, I do find a marginally significant positive impact on female wages in 1950 (compared to male wages) in Durable Manufacturing

[22]The 1930 Census does not have weeks worked reported for workers. Therefore, I can only instrument with the actual number of workers in an industry-state-year relative by gender as I do for my other specifications.

in column (3) with my preferred specification after accounting for industry and state-by-gender trends, but the magnitude of the impact is much smaller and less statistically significant than my estimates using annual wages and the number of workers, which is consistent with measurement error attenuating estimates towards zero. In Non-Durable Manufacturing in column (6) using my preferred specification, I find negative (and statistically insignificant) impacts on wages using the change in the relative number of women (Measure 1 in Table A.4), and positive and statistically significant impacts using the change in the relative proportion of women (Measure 2 in Table A.5). The noisiness of the estimates of impacts on Non-Durables leads me to be skeptical of all results in Tables A.4 and A.5 (including those on Durables).

1.7 Conclusion

Given the new findings and data available from Rose (2018), I revisit the framework of Acemoglu et al. (2004) and introduce a new dimension along which we should measure the differential evolution of relative wages for women—industry. This is an important change since despite plenty of anecdotal evidence on the boom in workers employed in shipbuilding and airplane building or the inundation of women workers into munitions factories, this is the first study, to my knowledge, to measure the impacts of the wartime production effort as changes in the relative industrial demand for women at the national level. Rose (2018) only exploits variations in wartime demand at the state-economic-area level in his paper, and while Shatnawi and Fishback (2018) do explore industry-level variation in relative wages for men and women, they restrict their analysis to Pennsylvania.

Utilizing a simple model of wage setting with labor-supply frictions in a market with

imperfect labor substitutes, I show how the relevant channel through which relative wages for women increase could be the changes in relative female-male labor-augmenting productivities. I use the USES assignments data across states and industries to operationalize two measures of the relative demand for women—the change in the number of women relative to the change in all labor in an industry, and a change in the relative proportion of women in an industry. I measure their impacts on wages for women relative to men, in 1940 and 1950 using a saturated model that accounts for industry trends and state-by-gender trends. Thus, even if we believe that military mobilization might have had differential impacts on wages for women across states (due to changes in their employment as per Acemoglu et al. (2004)), I account for these changes in my specification.

I find that within Durable Manufacturing, the impact of increased relative wartime demand for women persisted after the war, raising women's relative wages. Scaling the results I find, this translates to a 35.4–35.9% increase in the wages of women in 1950 as compared to men if they worked in the industry that saw the highest change in the relative demand for women (Transport and Equipment) compared to the industry that saw the lowest change in the relative demand for women (Fabricated Metal, or Stone, Clay and Glass) during World War II. Non-Durable Manufacturing, on the other hand, did not see gains in wages accrue to women.

In addition, wages did not increase for either men or women in response to total wartime employment shocks, either by state or industry. These results both enrich and reframe the literature on the impact of World War II on women in several ways:

First, it partly answers the question posed by Campbell (1984) on whether World War II caused a "watershed" change in the employment of women. While prior literature has said it does not (Goldin, 1991a; Goldin and Olivetti, 2013), I find that there were large increases in the

relative wages for women in Durable Manufacturing, which when reconciled with prior literature on increases in employment in Durable Manufacturing (Rose, 2018; Shatnawi and Fishback, 2018), suggests a substantial increase in the relative demand for women that persists post-war until 1950, despite a substantial exodus of women from the workforce in the immediate post-war period.

Second, nearly all previous literature has exploited state-level shocks in female wartime employment (or manpower mobilization) to find impacts on future female employment. While Rose (2018) does break down workers into different industry groups while comparing employment in 1950 to wartime employment, he still uses state-economic-level variation to find that there are increases for women in Durable Manufacturing and declines for women in Non-Durable Manufacturing. However, I find that when accounting for both state-level and industry-level variation in wartime demand, only industry-level variation is significant in increasing the relative wages for women compared to men in 1950, and it only impacts women in Durable Manufacturing. Despite many descriptive accounts of the drastic industrial composition shift of women during World War II (Hartmann, 1982; Milkman, 1987; Pidgeon, 1947), this is the first addition to the quantitative literature in explicitly measuring the impacts of changing relative industrial demand for women at the national level. Combined with the employment increases in Rose (2018), my finding of relative wage increase in Durable Manufacturing suggests a net national increase in the relative demand for women in Durable Manufacturing that is comparable to the magnitude of the relative increase in the demand for women found by Shatnawi and Fishback (2018) in Pennsylvania for the same period.

Third, my results suggest that the impacts on female relative wages are higher for women with 12 or more years of education, a subgroup that Goldin and Olivetti (2013) also find to have

the most significant positive impacts of manpower mobilization on employment in 1950 (and in 1960) compared to 1940. While they attribute this to a change in the labor supply for these groups of women, I show that this group also experiences a rise in their demand for Durable Manufacturing (as reflected by an increase in their wage premium). Thus, the benefits to educated women seem to have come from increased employer demand relative to shifting supply, perhaps via a change in the pecuniary costs of employing women.

I conclude that the cataclysmic changes in the industrial composition of manufacturing industries during World War II did increase the relative wages for women in Durable Manufacturing industries even after the war ended in 1950, despite the rapid exodus of women from the workforce post-war (Rose, 2018; Goldin, 1991a). These results help reinforce the importance of industrial demand as a mechanism for changing female labor force outcomes post World War II.

1.8 Figures

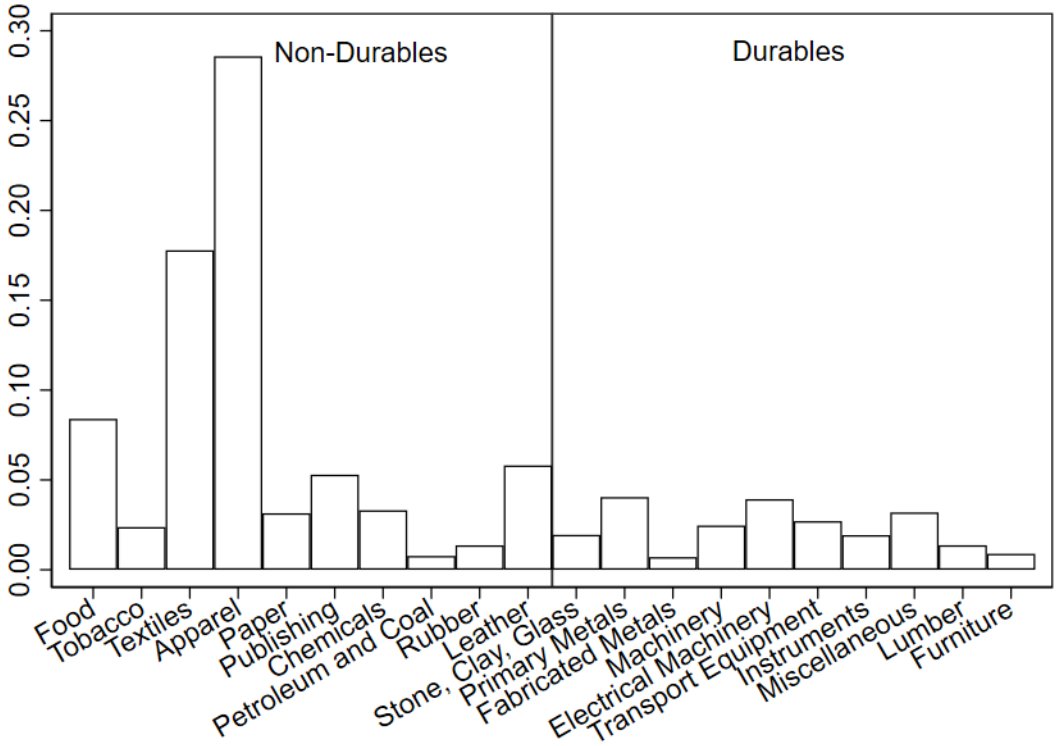


Figure 1.1: **Distribution of Women Within Manufacturing in 1940**

Notes: This figure shows the composition of female workers within Manufacturing in the 1940 Decennial Census. This is defined as the population-weighted number of women in each industry in the 1940 US Decennial Census, divided by the population-weighted number of women across all Manufacturing ($= 34,772,100$ women) in the 1940 US Decennial Census. Of the twenty industries that comprise Manufacturing, industries on the left up until Leather represent Non-Durables Manufacturing, and industries on the right, starting with Stone, Clay, and Glass represent Durable Manufacturing according to the CPS `ind1950` variable.

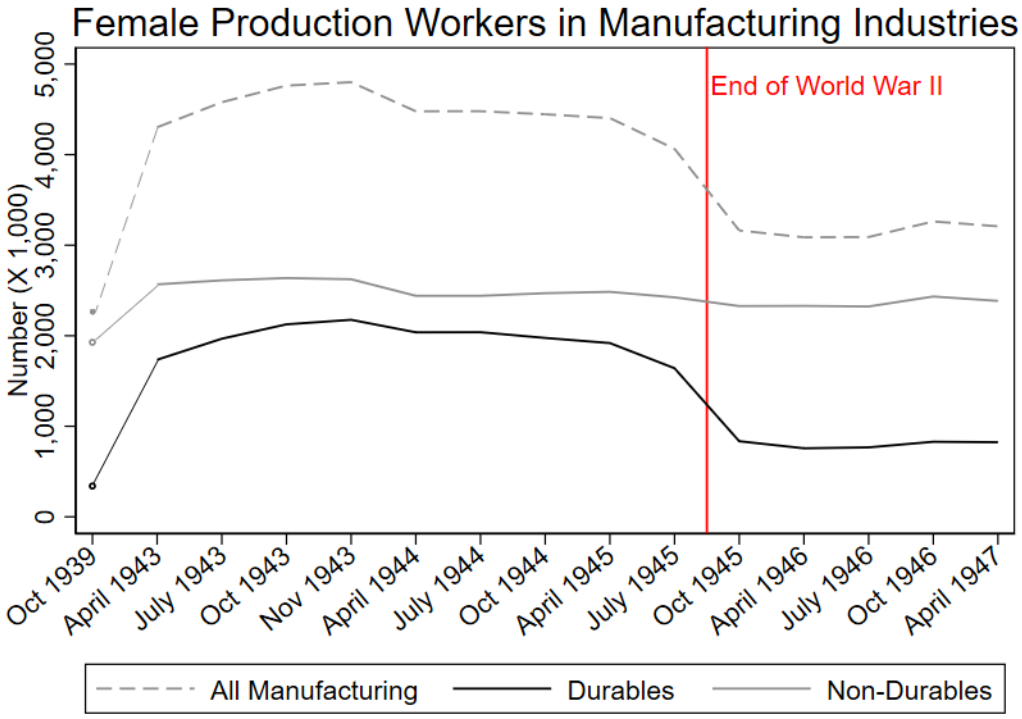


Figure 1.2: **Evolution of Female Employment in Manufacturing During World War II**
Notes: This figure shows the number of female workers in Manufacturing and by the two main manufacturing sub-industries during World War II from Pidgeon (1947), compiled from Monthly Labor Reviews posted by the Bureau of Labor Statistics. The values for 1939 come from the 1940 US Decennial Census.

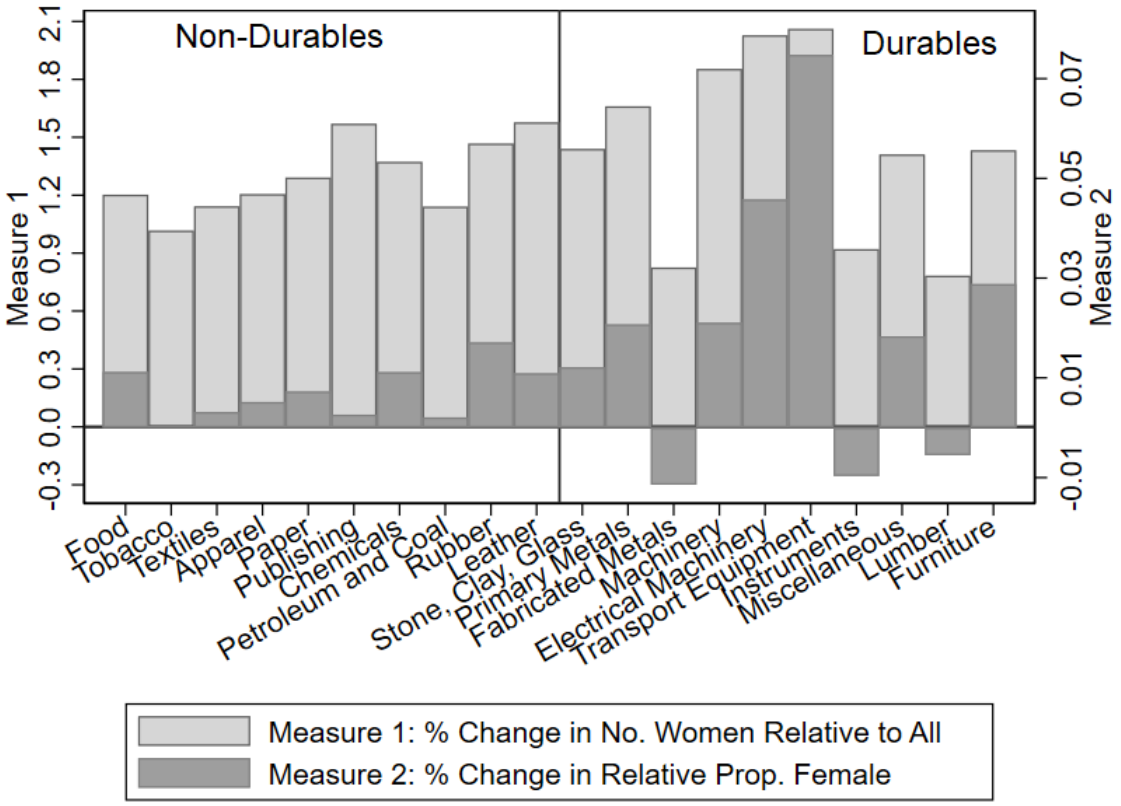


Figure 1.3: **Comparing Two Measures of Relative Female Demand**

Notes: This figure plots the two measures that I use to capture the industry-specific increases in the relative demand for women during WWII using assignments made by the United States Employment Service (USES) between 1944-46 and the 1940 US Decennial Census across all Manufacturing Industries. The first measure "Percent Change in Number of Women, Relative to All" is the percent change in the number of women between 1940 and 1946 divided by the percent change in all labor between 1940 and 1946. The second measure "Percent Change in Relative Proportion Female" measures the percent change in the proportion of an industry that is female between 1940 and 1946. To calculate values for 1946, I add the numbers from 1940 and USES assignments (hence, making the assumption that the only changes in employment in 1946 from the 1940 Census are the added numbers from the USES assignments). The process is also detailed in Table 1.2 and in Section 1.5.

1.9 Tables

Table 1.1: **Employment by Year and Industry**

	1940		1950		1944-1946	
	No. Total	No. Women	No. Total	No. Women	No. Total	No. Women
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Non-Durable Manufacturing						
Food & Beverages	17,375,900	3,378,800	26,392,100	5,541,700	1,015,840	237,574
Tobacco	1,418,200	774,000	1,638,900	832,000	28, 639	15,892
Textiles	14,195,800	5,690,600	18,724,800	6,981,800	318,725	146,016
Apparel	9,387,200	6,053,800	12,654,900	8,882,500	241,196	187,584
Paper	5,697,300	1,289,100	10,104,900	2,553,000	145,529	42,524
Printing	11,842,200	3,018,700	15,588,200	4,596,200	54,586	21,853
Chemicals	7,903,300	1,540,500	12,446,900	2,823,800	243,581	65,193
Petroleum & Coal	3,922,000	400,000	6,300,100	926,600	58,489	6,807
Rubber	2,536,500	575,400	4,185,900	1,091,300	96,245	32,044
Leather & Footwear	4,592,700	1,566,900	5,677,000	2,573,000	88,442	47,598
Total	78,871,100	24,287,800	113,713,700	36,801,900	2,291,272	803,085
Panel B: Durable Manufacturing						
Lumber and Wood	5,691,800	566,300	9,279,200	688,600	148,214	11,550
Furniture	2,145,100	350,000	4,709,500	1,055,600	152,985	35,766
Stone, Clay & Glass	4,504,800	828,700	6,948,000	1,438,300	127,914	33,884
Primary Metal	19,952,700	2,234,100	18,579,600	1,677,300	646,185	120,149
Fabricated Metal	1,302,800	302,800	16,059,800	3,139,900	91,606	17,582
Machinery	10,930,200	1,270,000	24,473,100	3,804,800	276,369	59,541
Electrical Machinery	6,081,300	1,653,400	16,210,900	5,627,000	283,278	156,236
Transport Equipment	11,593,800	1,152,100	25,917,600	3,859,300	878,269	179,947
Instruments	2,762,300	1,085,300	4,292,800	1,686,600	384,898	139,238
Miscellaneous	3,095,700	1,041,600	7,319,400	2,847,100	144,367	68,563
Total	58,060,500	10,484,300	133,789,900	25,824,500	3,134,085	822,456

Notes: The population-weighted total number of people employed in each manufacturing industry by year. The numbers for 1940 and 1950 are population-weighted, full-time workers from the 1940 and 1950 US Decennial Censuses respectively. 'Full-time employed' workers are those who worked more than 40 weeks in a year and more than 35 hours a week in 1950, and those who worked more than 40 weeks in a year in 1940, since all reported weeks are considered to be full-time equivalent in 1940. Numbers from 1944-1946 are all the assignments made by the USES (United States Employment Service) to the different manufacturing industries from 'The Labor Market' reports of the War Manpower Commission's Bureau of Program Planning and Review prior to 1945, and 'The Labor Market' reports from the Labor Department since 1945, made available by Rose (2018). AK, HI, NV DC, NM, MT and WY are dropped.

Table 1.2: **Measures of Relative Demand for Women During World War II**

	1940 Levels		1946 Levels		Percent Change from 1940 to 1946		
	No. Women	Prop. Female	No. Women	Prop. Female	No. Women	No. Women/No. Total	Prop. Female
	(W_{40})	$P_{40} = \left(\frac{W_{40}}{L_{40}}\right)$	(W_{46})	$P_{46} = \left(\frac{W_{46}}{L_{46}}\right)$	$\left(\frac{W_{46}-W_{40}}{W_{40}}\right)$	$\left(\frac{W_{46}-W_{40}}{W_{40}}\right) / \left(\frac{L_{46}-L_{40}}{L_{40}}\right)$	$\left(\frac{P_{46}-P_{40}}{P_{40}}\right)$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Non-Durable Manufacturing							
Food & Beverages	3,378,800	0.194	3,616,374	0.197	0.070	1.203	0.011
Tobacco	774,000	0.546	789,892	0.546	0.021	1.017	0.000
Textiles	5,690,600	0.401	5,836,616	0.402	0.026	1.143	0.003
Apparel	6,053,800	0.645	6,241,384	0.648	0.031	1.206	0.005
Paper	1,289,100	0.226	1,331,624	0.228	0.033	1.291	0.007
Printing	3,018,700	0.255	3,040,553	0.256	0.007	1.571	0.003
Chemicals	1,540,500	0.195	1,605,693	0.197	0.042	1.373	0.011
Petroleum & Coal	400,000	0.102	406,807	0.102	0.017	1.141	0.002
Rubber	575,400	0.227	607,444	0.231	0.056	1.468	0.017
Leather & Footwear	1,566,900	0.341	1,614,498	0.345	0.030	1.577	0.011
Panel B: Durable Manufacturing							
Lumber & Wood	566,300	0.099	577,850	0.099	0.020	0.783	-0.005
Furniture	350,000	0.163	385,766	0.168	0.102	1.433	0.029
Stone, Clay & Glass	828,700	0.184	862,584	0.186	0.041	1.440	0.012
Primary Metal	2,234,100	0.112	2,354,249	0.114	0.054	1.661	0.021
Fabricated Metal	302,800	0.232	320,382	0.230	0.058	0.823	-0.011
Machinery	1,270,000	0.116	1,329,541	0.119	0.047	1.854	0.021
Electrical Machinery	1,653,400	0.272	1,809,636	0.284	0.094	2.029	0.046
Transport Equipment	1,152,100	0.099	1,332,047	0.107	0.156	2.062	0.075
Instruments	1,085,300	0.393	1,224,538	0.389	0.128	0.921	-0.010
Miscellaneous	1,041,600	0.336	1,110,163	0.343	0.066	1.411	0.018

Notes: This table details manufacturing industry-specific increases in the relative demand for women during WWII using USES assignments made from 1944-46 and the 1940 US Decennial Census per measures calculated in Section 1.5. 1940 numbers are the number of population-weighted, full-time women (or all workers for L_{40}) from the 1940 US Decennial Census. 'Full-time employed' workers are those who worked more than 40 weeks in a year in 1940. To calculate numbers for 1946 in columns (3) and (4), I add the numbers from 1940 to USES assignments (making the assumption that the only changes in employment in 1946 from the 1940 Census are the added USES assignments).

Table 1.3: **Summary Statistics by Industry Group**

	Durables	Non-Durables
	(1)	(2)
Panel A: Employment Characteristics		
Log Annual Wage	7.091	6.898
	(0.770)	(0.853)
Log Hourly Wage	-0.290	-0.444
	(0.574)	(0.625)
Weeks Worked Last Year	44.704	44.511
	(11.590)	(12.077)
Hours Worked Last Week	39.388	37.881
	(9.805)	(12.520)
Production Workers	0.756	0.738
	(0.430)	(0.440)
Professional/Clerical Workers	0.210	0.202
	(0.407)	(0.401)
Panel B: Demographic Characteristics		
Female	0.171	0.364
	(0.376)	(0.481)
Married	0.720	0.669
	(0.449)	(0.471)
12 or More Years of Education	0.351	0.325
	(0.477)	(0.469)
Non-White	0.068	0.058
	(0.252)	(0.234)
No. of Children under Age 5	0.289	0.229
	(0.611)	(0.546)
N	27,745	27,642

Notes: Population-weighted means (and standard deviations in parentheses) for demographic and employment characteristics of individuals from the 1% samples of the 1940 and 1950 US Decennial Censuses. The final sample contains US-born, full-time workers employed in manufacturing industries, aged 15-64. AK, HI, and DC are dropped from my sample since they did not exist in 1945. MT, NV, NM, and WY are dropped from my final sample due to the unavailability of data on manufacturing workers in 1940 and 1950.

Table 1.4: **Impact of Industry-Level and State-Level WWII Employment Shocks on Wages in Manufacturing 1940-1950, Using Measure 1**

	Durables			Non-Durables		
	(1)	(2)	(3)	(4)	(5)	(6)
%Δ No. Women by Industry$\times d_{1950} \times f_i$	0.176***	0.191***	0.277***	-0.180	-0.262**	-0.030
	(0.059)	(0.064)	(0.093)	(0.130)	(0.131)	(0.136)
%Δ Total Demand by Industry$\times d_{1950}$	0.059	0.480		0.085	0.218	
	(0.483)	(0.434)		(0.716)	(0.536)	
%Δ No. Women by State$\times d_{1950} \times f_i$	-0.094			0.045		
	(0.085)			(0.063)		
%Δ Total Demand by State$\times d_{1950}$	-0.002			0.000		
	(0.003)			(0.000)		
$ln(\frac{F_{sjt}}{M_{sjt}})$	0.143**	0.162**	0.173*	-0.037	-0.039	-0.035
	(0.070)	(0.083)	(0.090)	(0.042)	(0.041)	(0.041)
$ln(\frac{F_{sjt}}{M_{sjt}}) \times f_i$	-0.157	-0.190	-0.199	-0.048	-0.034	-0.035
	(0.104)	(0.116)	(0.121)	(0.058)	(0.056)	(0.056)
d_{1950}	0.296***	0.344***	0.299***	0.086	0.112	0.149*
	(0.071)	(0.067)	(0.084)	(0.083)	(0.091)	(0.082)
f_i	-0.929***	-0.951***	-0.977***	-0.486***	-0.427***	-0.406***
	(0.279)	(0.308)	(0.323)	(0.118)	(0.116)	(0.117)
$f_i \times d_{1950}$	0.190	-0.037	-0.178	0.340	0.421**	0.082
	(0.185)	(0.155)	(0.179)	(0.219)	(0.215)	(0.215)
Gender-by-State-by-Year FE		✓	✓		✓	✓
Industry-by-Year FE			✓			✓
Individual-Level Controls	✓	✓	✓	✓	✓	✓
N	26,676	26,676	26,676	25,984	25,990	25,990
R^2	0.36	0.36	0.36	0.43	0.43	0.43
Cragg-Donald Wald F-Stat	650.25	560.76	528.91	1457.69	1695.69	1703.64

Notes: Results reported from estimating Equation 1.17 in columns (1) and (4), Equation 1.15 in columns (2) and (5), and Equation 1.19 in columns (3) and (6) for Durable and Non-Durable Manufacturing industries respectively. The main dependent variable is the log individual annual wages from a pooled 1% sample of U.S.-born full-time workers aged 15-64 from the 1940 and 1950 US Decennial Censuses. 'Measure 1' is the percent change in the number of women relative to the percent change in all workers during WWII defined in Section 1.5. All columns include gender-by-year, gender-by-state, and gender-by-industry fixed effects, as well as the log of relative female-male employment in the industry-state-year and its interaction with the female dummy, instrumented with the log levels of relative female-male employment in an industry-state-year in 1930 and its interaction with the female dummy. Columns (2) and (5) include gender-by-state-by-year fixed effects, and columns (3) and (6) include those as well as industry-by-year fixed effects. Individual covariates included are being married, number of children under age 5 in the household, 12 or more years of education, race, and a 4th-degree polynomial in potential experience. Robust standard errors are clustered by state and industry, and sample weights are used. Final sample excludes AK, HI, DC, MT, NV, NM, and WY. Stock-Yogo weak ID test critical values for the Cragg-Donald Wald F-Statistic: 7.03 (10% IV size), 4.58 (15% IV size), 3.93 (20% IV size) and 3.63 (25% IV size) (Stock and Yogo, 2002).

Table 1.5: **Impact of Industry-Level and State-Level WWII Employment Shocks on Wages in Manufacturing 1940-1950, Using Measure 2**

	Durables			Non-Durables		
	(1)	(2)	(3)	(4)	(5)	(6)
$\%\Delta$ Female Prop. of Industry$\times d_{1950} \times f_i$	3.174***	3.395***	4.178***	-1.274	-3.138	2.586
	(0.989)	(1.040)	(1.439)	(5.112)	(5.141)	(5.092)
$\%\Delta$ Total Demand by Industry$\times d_{1950}$	-0.078	0.323		0.307	0.611	
	(0.469)	(0.412)		(0.759)	(0.576)	
$\%\Delta$ Female Prop. of State$\times d_{1950} \times f_i$	-0.113			0.044		
	(0.090)			(0.068)		
$\%\Delta$ Total Demand by State$\times d_{1950}$	-0.002			0.000		
	(0.003)			(0.001)		
$ln(\frac{F_{sjt}}{M_{sjt}})$	0.143**	0.163**	0.174*	-0.037	-0.039	-0.035
	(0.070)	(0.083)	(0.090)	(0.042)	(0.041)	(0.041)
$ln(\frac{F_{sjt}}{M_{sjt}}) \times f_i$	-0.161	-0.193*	-0.199*	-0.049	-0.036	-0.036
	(0.104)	(0.116)	(0.120)	(0.058)	(0.056)	(0.056)
d_{1950}	0.302***	0.351***	0.307***	0.080	0.102	0.148*
	(0.071)	(0.067)	(0.084)	(0.083)	(0.091)	(0.082)
f_i	-0.891***	-0.912***	-0.913***	-0.486***	-0.432***	-0.408***
	(0.274)	(0.302)	(0.307)	(0.118)	(0.116)	(0.117)
$f_i \times d_{1950}$	0.299**	0.189	0.167	0.161	0.109	0.027
	(0.122)	(0.129)	(0.129)	(0.107)	(0.118)	(0.117)
Gender-by-State-by-Year FE		✓	✓		✓	✓
Industry-by-Year FE			✓			✓
Individual-Level Controls	✓	✓	✓	✓	✓	✓
N	26,676	26,676	26,676	25,984	25,990	25,990
R^2	0.36	0.36	0.36	0.43	0.43	0.43
Cragg-Donald Wald F-Stat	650.50	560.92	528.31	1457.69	1695.80	1703.28

Notes: Results reported from estimating Equation 1.17 in columns (1) and (4), Equation 1.15 in columns (2) and (5), and Equation 1.19 in columns (3) and (6) for Durable and Non-Durable Manufacturing industries respectively. The main dependent variable is the log individual annual wages from a pooled 1% sample of U.S.-born full-time workers aged 15-64 from the 1940 and 1950 US Decennial Censuses. 'Measure 2' is the percent change in the relative proportion of women in an industry (or state) during WWII defined in Section 1.5. All columns include gender-by-year, gender-by-state, and gender-by-industry fixed effects, as well as the log of relative female-male employment in the industry-state-year and its interaction with the female dummy, instrumented with the log levels of relative female-male employment in an industry-state-year in 1930 and its interaction with the female dummy. Columns (2) and (5) include gender-by-state-by-year fixed effects, and columns (3) and (6) include those as well as industry-by-year fixed effects. Individual covariates included are being married, number of children under age 5 in the household, 12 or more years of education, race, and a 4th-degree polynomial in potential experience. Robust standard errors are clustered by state and industry, and sample weights are used. Final sample excludes AK, HI, DC, MT, NV, NM, and WY. Stock-Yogo weak ID test critical values for the Cragg-Donald Wald F-Statistic: 7.03 (10% IV size), 4.58 (15% IV size), 3.93 (20% IV size) and 3.63 (25% IV size) (Stock and Yogo, 2002).

Table 1.6: **Heterogeneity in the Impact of the Industry-Level Employment Shock of WWII on Wages in Durable Manufacturing**

	Relative Demand Measure 1				Relative Demand Measure 2			
	High School or Higher		Less Than High School		High School or Higher		Less Than High School	
	Married	Unmarried	Married	Unmarried	Married	Unmarried	Married	Unmarried
Panel A: All Women								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Female Demand by Industry $\times d_{1950} \times f_i$	0.436**	0.487***	0.013	0.276	6.511**	6.020***	1.999	2.292
	(0.220)	(0.168)	(0.117)	(0.173)	(3.230)	(2.323)	(2.071)	(2.724)
d_{1950}	0.132	0.337**	0.246	0.108	0.138	0.375**	0.247	0.122
	(0.146)	(0.170)	(0.168)	(0.213)	(0.146)	(0.164)	(0.168)	(0.208)
f_i	-2.170***	0.107	-0.740	-0.679	-2.089***	0.198	-0.750	-0.596
	(0.740)	(0.394)	(0.614)	(0.687)	(0.716)	(0.372)	(0.611)	(0.657)
$f_i \times d_{1950}$	0.769	-0.861**	-0.144	0.101	1.353***	-0.212	-0.173	0.478
	(0.513)	(0.372)	(0.511)	(0.392)	(0.383)	(0.263)	(0.447)	(0.330)
N	6,343	3,124	13,014	4,195	6,343	3,124	13,014	4,195
R^2	0.26	0.34	0.33	0.35	0.26	0.33	0.33	0.35
Panel B: White Women								
Δ Female Demand by Industry $\times d_{1950} \times f_i$	0.334*	0.468***	-0.020	0.256	5.087*	5.837**	1.335	1.736
	(0.200)	(0.169)	(0.113)	(0.168)	(2.897)	(2.344)	(2.019)	(2.677)
d_{1950}	0.110	0.356**	0.147	0.139	0.115	0.393**	0.148	0.151
	(0.153)	(0.173)	(0.178)	(0.212)	(0.153)	(0.166)	(0.178)	(0.209)
f_i	-1.787***	0.185	-0.628	-0.617	-1.726***	0.272	-0.640	-0.548
	(0.659)	(0.397)	(0.630)	(0.653)	(0.639)	(0.375)	(0.630)	(0.631)
$f_i \times d_{1950}$	0.937*	-0.848**	-0.003	0.081	1.383***	-0.227	-0.069	0.440
	(0.507)	(0.371)	(0.520)	(0.390)	(0.393)	(0.262)	(0.457)	(0.327)
N	6,194	3,055	11,821	3,774	6,194	3,055	11,821	3,774
R^2	0.26	0.33	0.28	0.33	0.26	0.33	0.28	0.33

Notes: Results reported from estimating Equation 1.19 in Section 1.5. Relative Demand Measures 1 and 2 are detailed in section 1.5. The main dependent variable is the log individual annual wages from a pooled 1% sample of U.S.-born full-time workers aged 15-64 from the 1940 and 1950 US Decennial Censuses. All specifications include gender-by-state-by-year, gender-by-industry, and industry-by-year fixed effects, as well as the log levels of relative female-male employment in their industry-state-year and its interaction with the female dummy, instrumented with log levels of relative female-male employment in an industry-state-year in 1930 and its interaction with the female dummy. Individual covariates include being married, number of children under age 5 in the household, 12 or more years of education, race, and a 4th-degree polynomial in potential experience. Robust standard errors are clustered by state and industry, and sample weights are used.

Chapter 2: Putting Maternity on Ice: The Impacts of Financial Deregulation on Fertility

2.1 Introduction

The COVID-19 pandemic renewed interest in declining fertility rates in the US, with articles written in the first two years of the pandemic highlighting its negative impact on fertility (AP, 2021; Tavernise, 2021; Vinopal, 2021). A report published by the National Center for Health Statistics at the start of May 2021 reported that the U.S. fertility rate dropped by 4% over the course of 2020 compared to 2019 levels (Hamilton et al., 2021), and other reports predicted an even larger drop by the end of the pandemic (Kearney and Levine, 2020). Newer work has shown the decline in fertility over the last two years spread unevenly across states and socioeconomic groups, with older married women with higher education even experiencing an increase in births nationally (Adelman et al., 2023; Bailey et al., 2022).

A large decrease in fertility raises policy concerns about the sustainability of social welfare programs for the elderly that rely on contributions from a younger population. This is not the first time such concerns have arisen in the U.S. The U.S. Total Fertility Rate (TFR) for women aged 15–44 currently stands at 1.7, much lower than the 'replacement rate' of 2.1, i.e. the required TFR for a country to sustain its population (McPhillips, 2023).[1] The US fertility rate

[1]The TFR measures the hypothetical lifetime births per woman, based on the age-specific birth rate in a given

has been below replacement since 1971, and despite a slight boom in the early 2000s and the aforementioned evidence of a selective baby bump reversal during the pandemic for some subgroups of women, the TFR has been consistently decreasing in America since 2007, and been a subject of widespread scrutiny (Bailey et al., 2022).

The 1960s saw a nationwide sharp decline in fertility after the baby boom of the 1950s. The decline continued into the 1970s when the TFR dropped below the replacement rate. Several mechanisms have been studied to explain this decrease in fertility, such as a change in cultural values (Fernandez and Fogli, 2009), increasing economic opportunities for women (Willis, 1973; Butz and Ward, 1979), and the increased control over their own fertility for women with the introduction of oral contraception (Goldin and Katz, 2002) and legalization of abortion following Roe v. Wade (Levine et al., 1999).

In this chapter, we test if state-level banking deregulation (that began around this same time) had any significant impacts on fertility. Between 1960 and 1999, states in the U.S. began lifting restrictions on banks, allowing them to expand their branches within and across states, increasing competition between banks (Amel and Liang, 1992) and leading to greater availability of credit for previously excluded groups (Demyanyk, 2008; Beck et al., 2010; Jayaratne and Strahan, 1998).

At first glance, the impact of increased access to credit on fertility is ambiguous. If we think of children as normal goods, and increased access to credit and economic growth as raising household incomes, then we should expect fertility rates to rise in deregulated states. On the other hand, if the increased credit and economic growth lead to better labor market opportunities, then the opportunity cost of having children could rise significantly, and households might want

year.

to decrease their fertility.

We exploit the staggered timing of banking deregulation across states to estimate impacts on fertility using a difference-in-differences approach. Using both state-level data on fertility rates from Bailey et al. (2016), as well as individual-level data from the March supplement of the Current Population Survey (Flood et al., 2020), we find that the deregulation of a state leads to an average decrease in the fertility of women aged 20-44 by about 4.3% at the state level and by 6.5% at the individual level. Fertility declines gradually in the years after the deregulation in a state, with the strongest impacts felt two years post-deregulation onward.

New econometric literature suggests that classic differences in differences can imperfectly estimate two-way fixed effects (TWFE) when there is a difference in the timing of units being treated (Sun and Abraham, 2021; Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021). We find upon re-estimating our main tables using specifications robust to the staggered timing of treatment devised by Borusyak et al. (2023), our results are no longer statistically significant and decrease substantially in magnitude. However, when we break down our sample into before 1989 (when three-quarters of our 47 states had deregulated) and after, we find that our classic TWFE estimates are similar in magnitude to the TWFE estimates robust to the timing of treatment for the sample of years before 1989, and noisier and more disparate in the sample in the years including 1989 and after at both the state and the individual-levels. Since we look at a sample where states deregulate over a long period of time, it is natural to expect that impacts might be more significant for the earlier states in our sample.[2] This is confirmed when we plot our fertility decline impacts separately by year and find estimates after 1988 to be extremely noisy in both our state and

[2]States deregulated over 39 years as seen in Table 2.1. The first states to deregulate did so in 1960, whereas the last to deregulate was Iowa in 1999.

individual data samples in Appendix Figures B.1 and B.2 respectively. Thus, for the rest of the analysis, we restrict our sample to years before 1989. For these earlier years, we find that the impacts of state deregulation on fertility range from declines of 1.7% in the state-level data to 2.7% in the individual-level data.

When looking at possible mechanisms for fertility decline, we find suggestive evidence that the declines in fertility were largest for non-white women and poorer households. This suggests that one of the main drivers of the decrease in fertility could be an increase in labor market opportunities for previously excluded groups, raising their opportunity cost of having children. In line with these findings, we also see that women have fewer total children, and our estimates imply that state deregulation led to marginally decreasing marriage rates and increasing divorce rates, although the results are not statistically significant.

Section 2.2 provides details on bank deregulation and offers some motivation on why we could expect to see the impacts of bank deregulation on fertility. We explain our data in Section 2.3 and outline our difference-in-differences methodology in Section 2.4. Our results are discussed in Section 2.5 and we conclude in Section 2.6.

2.2 Background and Motivation

2.2.1 Bank Deregulation

Prior to the 1970s, banking was highly regulated in the U.S. Most states had restrictions on the number of branches that banks were allowed to operate both within and across states, with some states only allowing one branch per state (referred to as 'unit banking'). This allowed for the formation of local banking monopolies and restricted individuals from accessing credit in

'under-banked' regions.

However, from the 1960s to 1990s, states began deregulating their banking sectors in a staggered fashion (see Table 2.1 and Figures 2.1 and 2.2). The last states to deregulate their banks did so after the 1994 passage of the Riegle-Neal Interstate Banking and Branching Efficiency Act. In deregulated states, banks were allowed to expand both across and within states, either by opening new branches (called '*de novo* branching') or via mergers and acquisitions with other local institutions.[3]

As a consequence of bank reform, there was significant new entry into local banking markets (Amel and Liang, 1992) which improved bank efficiency by increasing competition between banks (Jayaratne and Strahan, 1998, 1996). Deregulation-induced competition between banks and the opening of new branches in remote areas of states led to higher lending, and allowed previously excluded households access to banking (Dick and Lehnert, 2010).

Intra-state banking deregulation significantly increased economic dynamism within the local state economies. Previous literature has found highly positive and significant impacts of banking deregulation on various economic outcomes, such as economic growth (Jayaratne and Strahan, 1996), entrepreneurship (Black and Strahan, 2002; Kerr and Nanda, 2009), small business lending (Rice and Strahan, 2010; Demyanyk et al., 2007), and boosting local business cycles (Morgan et al., 2004). States enacting banking deregulation also saw increases in college enrollment rates (Sun and Yannelis, 2016), as well as decreases in income inequality (Beck et al., 2010) and the gender gap in labor force participation (Popov and Zaharia, 2019).

[3]Previous literature on banking reform has found that the removal of intra-state banking restrictions was far more significant in changing the banking structure and real economy (Beck et al., 2010; Jayaratne and Strahan, 1996). Thus, even though some states may have had their intra-state and inter-state restrictions lifted in different years, we only focus on the year of intra-state reform in our analysis. Recomputing our results including both types of reform directly shows us that inter-state banking reforms did not significantly impact fertility rates.

2.2.2 How Could the Bank Reforms Impact Fertility

As discussed, bank deregulation increased business dynamism and job availability within states (Kerr and Nanda, 2009). In particular, the banking reforms gave lower-income households and marginalized groups to have increased access to credit, lowering income inequality (Beck et al., 2010) as well as racial and gender gaps in labor force participation (Levine et al., 2012; Popov and Zaharia, 2019). Thus, the opportunity cost of being out of the labor force, either due to childbirth or childcare significantly increased, and women (or households) could choose to delay or decrease their fertility. This would be particularly true for groups that saw the largest returns to bank deregulation in the labor market, such as non-white and lower-income households. Access to credit (and related labor market opportunities) could also raise the outside options of women and see them decrease marriage rates and initiate divorces because they are no longer dependent on their families, thus reducing their fertility.

On the other hand, if rising labor market opportunities meant that household incomes rose, and children are considered normal goods, then this could lead to an increase in birth rates.

Bank deregulation increased house prices (Favara and Imbs, 2015), and thus home equity for existing homeowners. Dettling and Kearney (2014) and Lovenheim and Mumford (2013) show that fertility often increases with housing wealth, and this means we could see homeowners having more children in deregulated states. At the same time, if housing prices increase and housing stock is positively correlated with fertility, we could find that renting families see a decline in fertility since their path to homeownership gets more expensive.

2.3 Data

We use two different data sets to explore the effect of bank deregulation on fertility rates at the state level and at the individual level.[4]

First, for our state-level analysis, we use Vital Statistics data created by Bailey et al. (2016).[5] They provide fertility data by county and race from 1915 to 2007. We aggregate their data up to the state level to create a state-year panel and use the observations from 1970-2000. We define the fertility rate as the number of births in a state per 1000 women aged 20-44.

Second, we use the Current Population Survey (CPS) for individual-level data from the March Supplement (ASEC) from 1977–2000 (Flood et al., 2020).[6] We are also able to use the CPS data to control for educational attainment, race, age, and state of residence, as well as test for heterogeneity in results along other individual and household characteristics.

Our main independent variable is created using the data on the year of bank deregulation by state from Beck et al. (2010) and Kroszner and Strahan (1999). Following other literature on bank deregulation, Delaware and South Dakota are not included in our analysis because laws in these states facilitated the growth of the credit card industry, changing the structure of their banking systems. We also exclude Alaska, Hawaii, Puerto Rico, and other territories, but do include D.C. A full list of all 47 states in our analysis and the years they underwent bank reform

[4]We had also hoped to use the National Longitudinal Survey of Youth 1979 (NLSY) (Bureau of Labor Statistics, 2019) from 1979-2000 to augment our individual-level results and check for mechanisms related to credit access. However, we did not find time-varying measures of credit access or banking in the time period we wanted to study (i.e. 1979 to 1989). Further, the small sample size of our restricted sample meant that we lost statistical power to draw meaningful conclusions from the results in the NLSY79 data. Thus, we do not report the results from the NLSY79 in this chapter.

[5]available at https://www.openicpsr.org/openicpsr/project/100229/version/V4/view

[6]We have to limit the CPS data to start in 1977 because only nine states are well-defined between 1968-1976. Other papers that look at the impact of bank deregulation on individual outcomes using CPS data also make the same sample restriction (Beck et al., 2010; Popov and Zaharia, 2019).

are detailed in Table 2.1.

Annual state-level income, population, and employment data, which we use as state-level controls in our specifications, are downloaded from the Bureau of Economic Analysis's online database (Bureau of Economic Analysis, 2020).[7]

2.4 Empirical Strategy

The staggered timing of banking deregulation across states leads us to a difference-in-differences methodology for our empirical strategy. Jayaratne and Strahan (1996) argue that states did not deregulate their banking industry in anticipation of future growth prospects or in response to local economic conditions. Kroszner and Strahan (1999) show that technological innovations such as the invention of automatic teller machines (ATMs), the introduction of checkable money market mutual funds, and improvements in communications technology reduced the monopoly power of local banks, and weakened their ability to fight against deregulation. They show deregulation occurred later in states where politically powerful groups considered competition from large banks as a potential threat to their interests. Using a Weibull hazard model to predict the years remaining to deregulation, we establish in Appendix Tables B.1 and B.2 that fertility rates before deregulation have no impact on a state's time to intrastate deregulation, suggesting that the deregulation at the state level is exogenous to fertility rates, and driven by political economy factors and state banking sector characteristics listed by Kroszner and Strahan (1999).

While this provides some support for the plausible exogeneity of the timing of banking reforms, for our difference-in-differences strategy to be valid, the identifying assumption is that

[7] available at https://apps.bea.gov/regional/downloadzip.cfm

states that deregulated would have had their trajectory of fertility rates move in parallel with the states that did not deregulate, in the absence of bank deregulation. While the counterfactual parallel trends assumption cannot be tested, we can test for parallel pre-trends in fertility between deregulated and non-deregulated states. To do so, we plot the coefficient estimates and 95% confidence intervals from the following equation:

$$Y_{s,t} = \beta_0 + \beta_1 D_{s,t}^{-10} + \beta_2 D_{s,t}^{-9} + ... + \beta_{26} D_{s,t}^{+15} + \delta_s + \phi_t + \epsilon_{s,t} \tag{2.1}$$

where $Y_{s,t}$ is the state s's fertility rate in year t, defined as the number of births per 1000 women aged 20-44. $D_{s,t}^{-\tau}$ equals one for the state s in the τth year before deregulation and zero otherwise. $D_{s,t}^{+\tau}$ is also defined analogously. The coefficient on the dummy variable for the year prior to deregulation is normalized at zero. Figure 2.3 plots the β_τ coefficients to show evidence of zero pre-trends, i.e. that the coefficients on the deregulation dummy variables are insignificantly different from zero for all years before deregulation, and they show a sharp gradual decrease in the years after deregulation.[8]

Thus, we move forward with our difference-in-differences strategy. Our first empirical specification uses state-level data and is as follows:

$$Y_{s,t} = \beta_0 + \beta_1 D_{s,t} + \mathbf{X}_{s,t}\Gamma + \delta_s + \phi_t + \epsilon_{s,t} \tag{2.2}$$

where $Y_{s,t}$ measures the fertility rate in state s at time t, measured as the number of births per thousand women of childbearing age (ages 20–44). $D_{s,t}$ is a dummy variable that equals one

[8]When we include state contemporaneous characteristics in this estimation, our pre-trends become even flatter and closer to zero.

in the years after state s deregulates and zero otherwise. $\mathbf{X}_{s,t}$ is a vector of contemporaneous controls at the state level, which includes the log of population, the proportion of the population that is employed, and the proportion of population that is white. The coefficient of interest is β_1, which measures the impact of branch deregulation on the fertility rate, averaged across all the years post deregulation.

We also include state fixed effects, δ_s and time fixed effects, ϕ_t. State fixed effects control for time-invariant unobserved state characteristics that might affect fertility rates for reasons unrelated to banking market structure. The time fixed effects control for the secular trends in fertility rates at the national level and for any national-level cyclical variation in fertility rates. We do not add linear interactions between the year dummies and indicators for different states, since we do not see any evidence of pre-trends in Figure 2.3.[9] We report standard errors clustered by state.

Next, we move on to our difference-in-differences using individual-level data from the CPS. We use the following specification:

$$Y_{i,s,t} = \beta_0 + \beta_1 D_{s,t} + \mathbf{X}_{s,t}\Gamma + \mathbf{Z}_{i,s,t}\Psi + \delta_s + \phi_t + \epsilon_{i,s,t} \tag{2.3}$$

where $Y_{i,s,t}$ measures whether a woman i, aged between 20–44 in state s in year t has given birth within the last year.[10] $D_{s,t}$ is a dummy variable that equals one in the years after state s deregulates and zero otherwise. $\mathbf{Z}_{i,s,t}$ is a vector of individual controls for a woman i aged between 20-44 in state s in year t, including race, whether a woman has high school (HS) education or higher, and a 4th degree polynomial in age. The coefficient of interest is again

[9] When re-doing our main tables with a linear state time trend, our results stay consistent.

[10] We leave out teenage births here since they follow a very different evolution in the US (Hamilton et al., 2021).

β_1, which now measures the impact of bank deregulation on the likelihood that a woman aged 20-44 has given birth in the last year, averaged across all the years post deregulation in a state.

Again, we include the vector of contemporaneous state-level controls $\mathbf{X}_{s,t}$, state fixed effects, δ_s and time fixed effects, ϕ_t as in the previous specification. We report standard errors clustered by state and year.

2.5 Results

2.5.1 Preliminary Results

We first report estimates from equation 2.2 using our full sample of state-level observations in Table 2.2. Our preferred final estimates in column (3) that include state and year fixed effects as well as state-level controls see the average impact of bank reform on all years post-deregulation as a decrease in the state fertility rate of about 4.32% (or a decline of 3.08 births per 1000 on a mean fertility rate of 71.34 births per 1000 women aged 20–44). The magnitude of the impacts do not change significantly before and after including state contemporaneous controls between columns (2) and (3) of Panel A.

Similar estimates from equation 2.3 using our full sample of individual-level observations from the CPS are reported in Panel A of Table 2.3. According to our preferred specification in column (4) (which includes state and year fixed effects as well as state-level and individual-level controls), women aged 20–44 in states that undergo bank reform are 0.44 percentage points less likely to give birth in the past year, which is a decrease of 6.57% on the mean of 0.07. Again, the magnitude of the impact does not change much upon including state or individual controls between columns (2), (3), and (4).

2.5.2 Robustness to Staggered Timing of Treatment

Recent econometric literature points out the flaws in estimating a two-way fixed effects (TWFE) model in contexts with differential treatment timing across units (Sun and Abraham, 2021; Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021). Since in our setting, we do have different states deregulating at different times, our classic two-way difference-in-differences estimator used in the previous section could over-weight the impacts from the states that were deregulated earlier in our sample.

To correct for this, we re-estimate our main state-level and individual-level fertility results using both the classic TWFE model as well as differences-in-differences robust to staggered treatment timing devised by Borusyak et al. (2023).[11] The results are presented side-by-side in Panel A of Appendix Tables B.3 and B.4. Unfortunately, we find that our results are not robust to this new estimator when using state and year fixed effects between columns (2) and (4) in Panel A of both Tables. The magnitudes and statistical significance sharply shrink when using TWFE estimators robust to the timing of treatment.

We think that this might be because the impacts on fertility are driven by early states in our sample, especially since we cover such a wide time horizon in our analysis. To check this, we plot our treatment coefficients separately by year for both our state-level and individual-level estimates using the CPS in Appendix Figures B.1 and B.2 respectively. We see that the estimates for the impacts of state deregulation on fertility become noisier and more positive halfway through the time period we measure. The year, when 75% of all our sample states had deregulated, is 1989, so we choose to split our sample and present results using both classical TWFE and Robust TWFE

[11] The Github code for this estimator is available at https://github.com/borusyak/did_imputation or can be installed on Stata via the command "ssc install did_imputation".

across these two periods in Panels B and C of Appendix Tables B.3 and B.4.[12]

The Borusyak et al. (2023) estimator calculates similar magnitudes of impacts (although not exactly the same) of deregulation on state-level fertility before 1989 in columns (2) and (4) of Panel B of Appendix Table B.3. This translates to a decline of 4.8-5.5% decrease in state-level fertility after a state deregulates (or a decline of about 3.5-4.1 births per 1000 in a state for all women aged 20-44, on a base of 74.03 births per 1000). However, when looking at the data between 1989-2000, the impacts of state deregulation on fertility are no longer statistically significant at the 10% level using either the classical or the robust TWFE estimators. Additionally, they estimate impacts of opposing magnitudes, which suggests that these estimates are less trustworthy and perhaps indicative of noise rather than a true impact on fertility.

When we repeat the exercise with our individual-level CPS data in Appendix Table B.4, we find similar results. In Panel B, when looking at data from 1977 until 1988, we find similar (but not identical) magnitudes of the impact of state deregulation on fertility of -3.7-4.7% (or a decline of 0.3 percentage points in the probability of having given birth in the last year). The estimates in column (2) and (4) are statistically significantly different from each other when looking at data in 1989 and beyond in Panel C, and the estimates using the robust TWFE estimator shrinks to nearly zero in column (4).

Given these findings for the rest of the analysis, we use the classical TWFE estimators but restrict our sample to observations before 1989. Therefore, our main results come from column (3) of Panel B of Tables 2.2 and 2.3. The impact of a state deregulating is about a 1.7% decline in state-level fertility (a decline of 1.25 births per 1000 women aged 20-44 from a base of 74

[12]We wanted to select the most recent relevant year in our sample to divide the data. The year when a median number of banks had deregulated would be 1985. Referencing the figures B.1 and B.2, we chose our split year to be 1989. Results are not very different for the early batch of states if we choose 1985 as our splitting year.

births per 1000) and a 2.7% decline in individual-level fertility using CPS data (a decline of 0.19 percentage points in the likelihood of having given birth in the past year).

2.5.3 Dynamic Impacts

Since fertility is a variable where we expect to see lagged impacts, we re-estimate our difference-in-differences equation in 2.2 as a dynamic equation (Bertrand and Mullainathan, 2003), by replacing our binary dummy with 5 separate binary dummies for bank reform: the first equal to one for up to two years prior to deregulation, the next equal to one only for the year of deregulation, the third equal to one only for the year after deregulation, the fourth equal to one in the second year after bank deregulation, and the last equal to one for all years two onward after deregulation. The results are reported in Table 2.4. As expected, in column (3) after accounting for state and year fixed effects as well as state controls, we see that the entire decline in state fertility rates is driven by impacts in the years after bank deregulation. There are statistically insignificant declines in fertility in the year after deregulation and two years after deregulation but we see our average impacts of 1.7% in Table 2.2 come almost entirely from an average decrease of 3.78% in the post two-year period after bank reforms are enacted, averaged with insignificant and smaller magnitude of decreases in the two immediate years after bank reform is enacted. The impacts in year of and years prior to bank reform are positive and statistically insignificant.

Once again, we re-estimate our difference-in-differences equation in 2.3 as a dynamic equation, by replacing our binary dummy with the same 5 binary dummies for bank reform as above, with our individual data using CPS. The dynamic results are reported in Table 2.5. Once again, the largest declines in fertility come in the years after bank deregulation. We see that

women are 7.88% less likely to give birth within the last year two years after bank deregulation, and 7.02% less likely to give birth within the past year after more than two years post bank deregulation. The impacts in the years before, the year of, and the year after bank deregulation are statistically insignificant and not statistically distinguishable from each other.

2.5.4 Possible Mechanisms

2.5.4.1 Outside Options for Women

As discussed in section 2.2, bank deregulation significantly increased economic growth within a state. In particular, it contributed to job growth, an uptick in entrepreneurship, and the decrease in the gender gap in labor force participation. This could have offered women more opportunities in the labor force, which, even if not taken up, could increase their bargaining position within the household. As a consequence of increased demand for labor, women could want to now decrease their fertility since the opportunity cost of having children (and subsequent care if seen as time away from the labor force) rose significantly.

We test this in two different ways. Popov and Zaharia (2019) show how the gender gap in labor force participation falls in deregulated states due to a combination of multiple channels: an increase in net job creation (with a particular expansion in service sector jobs) and jobs requiring female-specific skills. We test whether the decrease in fertility is larger in magnitude for those women who are in the labor force, in service industries or in female-dominated industries.[13] The results are shown in Figure 2.5. We can see that while these sub-populations do see a decrease in

[13]We define female-dominated industries in a similar way to Popov and Zaharia (2019), assigning industries a value of 1 ("female dominated") if the proportion of women in that industry are more than the average share of women across all industries in 1977.

their likelihood of giving birth in the last year that is significantly different from zero, we cannot say that they are distinct from the full sample impact on women. Therefore, women who took up new opportunities in the labor market as a consequence of state banking reform may not have been the sole drivers of a decrease in fertility rates in the state. In the same Figure 2.5, we also plot the impact on individual fertility for women who have some college education and those who are self-employed since these are also groups that saw an increase in their numbers as a result of bank deregulation. Again, we find that the impacts on these groups are significantly negative, but not distinguishable from the full sample of women. However, it could be that even women who do not enter the labor force, self-employ or enroll in college see their bargaining power within their households increase due to the mere existence of these outside options, allowing them to delay fertility.

One way we can test if the decline in fertility comes from a change in the relative bargaining power of women is to check if access to fertility allowed them to delay marriages or initiate divorces (thereby decreasing total fertility for women) since they are no longer tied to husbands for credit. We show the results of estimating a difference-in-differences regression as in 2.3, for dependent variables related to marriage and fertility in Table 2.6. For most of the variables, we do see results in the expected directions but the magnitudes are economically small and statistically insignificant. Women decrease their likelihood of being married by 0.93% in column (1) and see an analogous increase in the likelihood of being divorced in column (2) by 3.46% albeit both magnitudes are not statistically significant. In column (3), we see the impact on the age at which women have their first child increased by a statistically insignificant 0.14%, and the total number of children a woman has decreased by 2.11% in column (4) after bank deregulation. The decline in total fertility is statistically significant. The decrease in the likelihood of giving birth in the last

year is lower for married women in column (5) at 1% than for our full sample in column (5) of Table 2.3 at 2.7%, but once again this estimate is not statistically significant. From these results, we find at best suggestive evidence of fertility declines being caused by changes in the labor market opportunities or bargaining power of women in households; we do not have the statistical power to conclude that this is an important driver of fertility decline.

2.5.4.2 Access to Credit for Previously Excluded Groups

Bank deregulation also increased economic growth by specifically boosting incomes and job opportunities for groups marginalized by race or income, as well as extending credit to these groups who may have been previously excluded by the banking industry (Demyanyk, 2008; Beck et al., 2010; Levine et al., 2012).

As a result, these households may have increased their labor force participation and therefore decreased their fertility. We show the impacts on individual fertility by race and income in Table 2.7. The declines in fertility for white women in Panel A are almost entirely driven by point estimates of the decline in fertility for women in the lowest quartile of the income distribution in column (1), with fertility declining by 6.5% after a state deregulates for low-income white women. Non-white women in Panel B on the other hand, see fertility declines across almost all quartiles of income, and non-white women in the lowest income quartile experience fertility declines five times as large as those experienced by white women in the lowest income quartile, of 30.88% post state deregulation. However, once again none of our impacts are statistically significant, and we cannot conclude definitively that access to credit or banking for previously excluded groups was impactful in decreasing fertility.

2.5.4.3 Housing Prices

Bank deregulation also had the effect of increasing housing prices in deregulated states (Favara and Imbs, 2015). This has the possibility of going in two directions. While homeowners would have seen an increase in their housing equity, which in previous literature has predicted to correlate positively with fertility Dettling and Kearney (2014), renters would have seen their barriers to home ownership rise, thereby perhaps decreasing their fertility.

When we compute the impact of bank reform on the likelihood of giving birth within the last year separately for homeowners and non-homeowners (renters) in Table 2.8, we see that the likelihoods decrease for both the groups in columns (1) and (2), but the negative impact on fertility is much larger in magnitude (and comparable to our full sample results) for renters with a decrease of 4.6% in their fertility after state deregulation. As before, the lack of statistical significance in our impacts implies that we cannot conclude that housing prices are an important driver of the fertility decline.

The statistical insignificance of all the estimates further lends doubt to their credibility when exploring our potential mechanisms and we want to be cautious about inferences we make from this entire Section 2.5.4.

2.6 Conclusion

Thus, we find that the deregulation of state banking sectors led to a decline in the fertility of women in deregulated states of 1.7-2.7% across different state and individual-level datasets in the years between 1970 and 1988. The size of the decline we see in fertility represents a meaningful magnitude, although in context, it is smaller than the 4% decline measured in fertility

as a consequence of state abortion laws (Levine et al., 1999), and the 9-11% decline in fertility measured as a result of the Great Recession (Cherlin et al., 2013).

The decline in fertility is driven by the states we see deregulate at the beginning of our sample. Results for declines in fertility after 1989 are not robust to alternate difference-in-differences methodology cited in new econometric literature that corrects for staggered treatment timing.

In testing for possible mechanisms, we find at best suggestive evidence of the fertility declines coming through increases in job market opportunities for women, and non-white and poorer families, which is consistent with previous literature on bank deregulation improving credit access and labor market outcomes for marginalized groups (Demyanyk, 2008; Beck et al., 2010; Levine et al., 2012). We also find limited suggestive evidence that the change in fertility could be attributed to increasing housing prices as a consequence of bank deregulation (Favara and Imbs, 2015; Dettling and Kearney, 2014). For all of our mechanisms, we lack the statistical power to conclusively estimate the driving channels behind the decreased fertility.

Our results highlight that outside economic forces like state banking regulations, which seem disparate to household formation and fertility decisions, could still have economically meaningful and statistically significant impacts on such outcomes. In this chapter, we find a small but statistically significant decline in the fertility rates of women as a result of banking deregulation before 1989, which highlights state-wide differences in fertility at a time when the national fertility rate was fairly stagnant. Thus, in examining the declining U.S. fertility rate, it is important to understand the mechanisms through which fertility behavior within households can evolve in response to outside labor market opportunities and monetary policy effects on local economies.

2.7 Figures

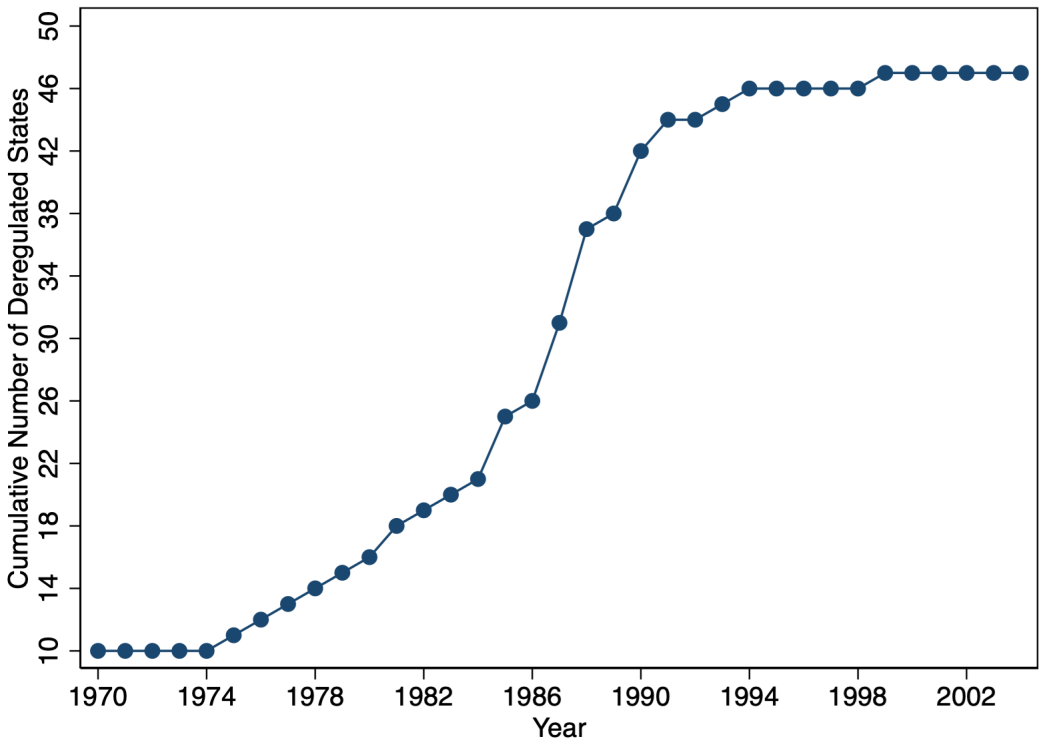


Figure 2.1: **Number of States Deregulating their Banking Sector, by Year**

Notes: Our state-level data starts in 1970, by which time 10 states had already deregulated their banking sectors. 9 deregulated in the 1960s, and VT in 1970 (see Table 2.1 for all state years). By the end of our original sample in 2000, all states had deregulated. We do not include in our analysis SD and DE due to their unique banking structure, as well as AK and HI. We do include DC, for a total of 47 states in our original sample. In our smaller final sample cut off in 1988; 75% of our states (37 in total) had their banks deregulated by the end of the sample period.

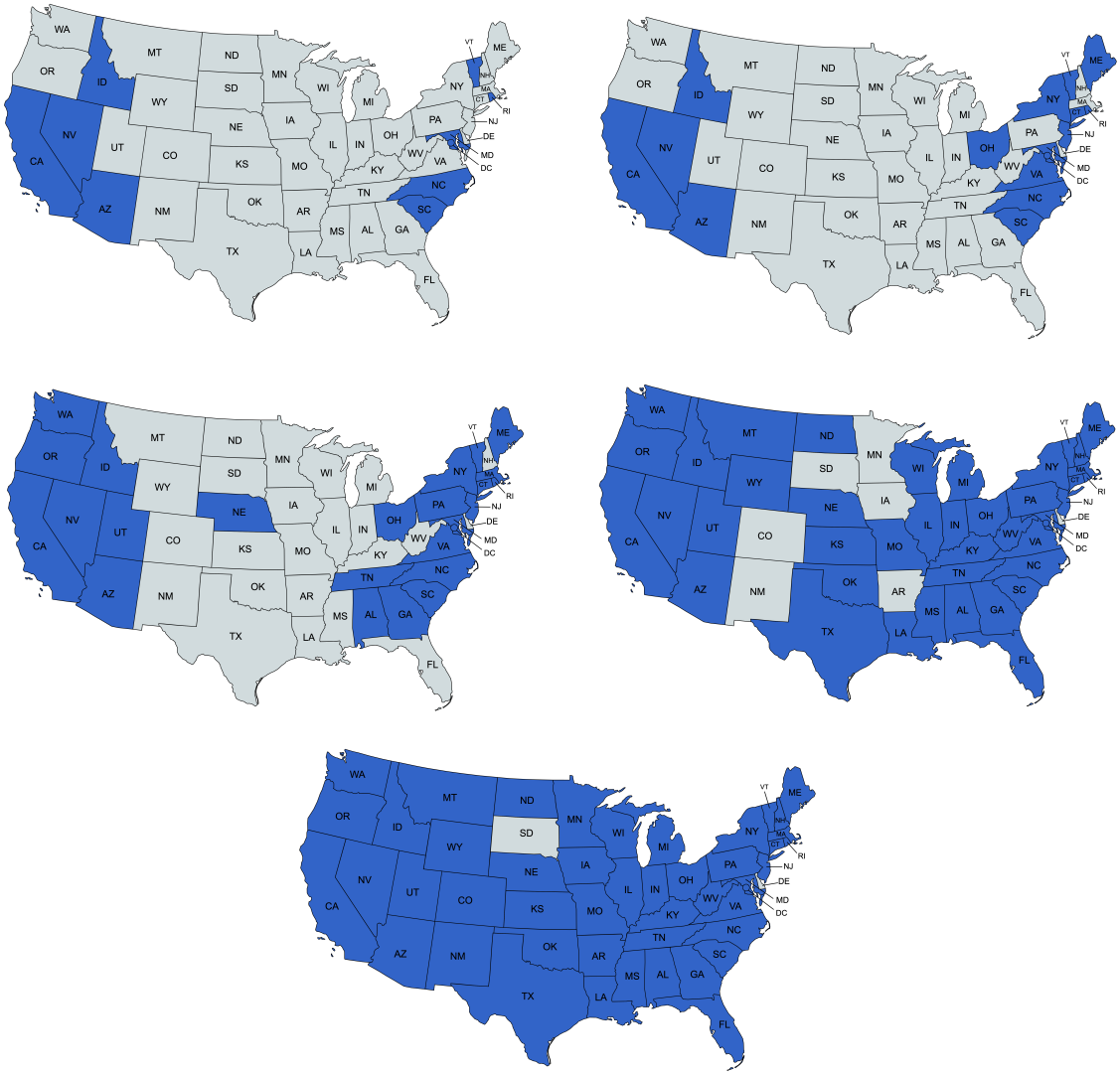


Figure 2.2: **U.S. States Deregulated as of 1970, 1980, 1985, 1990 and 2000.**
Notes: We show a geographic display of states across the US that were deregulated turning blue from grey (from left to right, top to bottom) by 1970, 1980, 1985, 1990, and 2000 (see Table 2.1 for all state-years of deregulation). We do not include in our analysis SD and DE due to their unique banking structure, as well as AK and HI. We do include DC, which gives us a total of 47 states in our original sample, but only 37 states had deregulated in our truncated sample that ends in 1988.

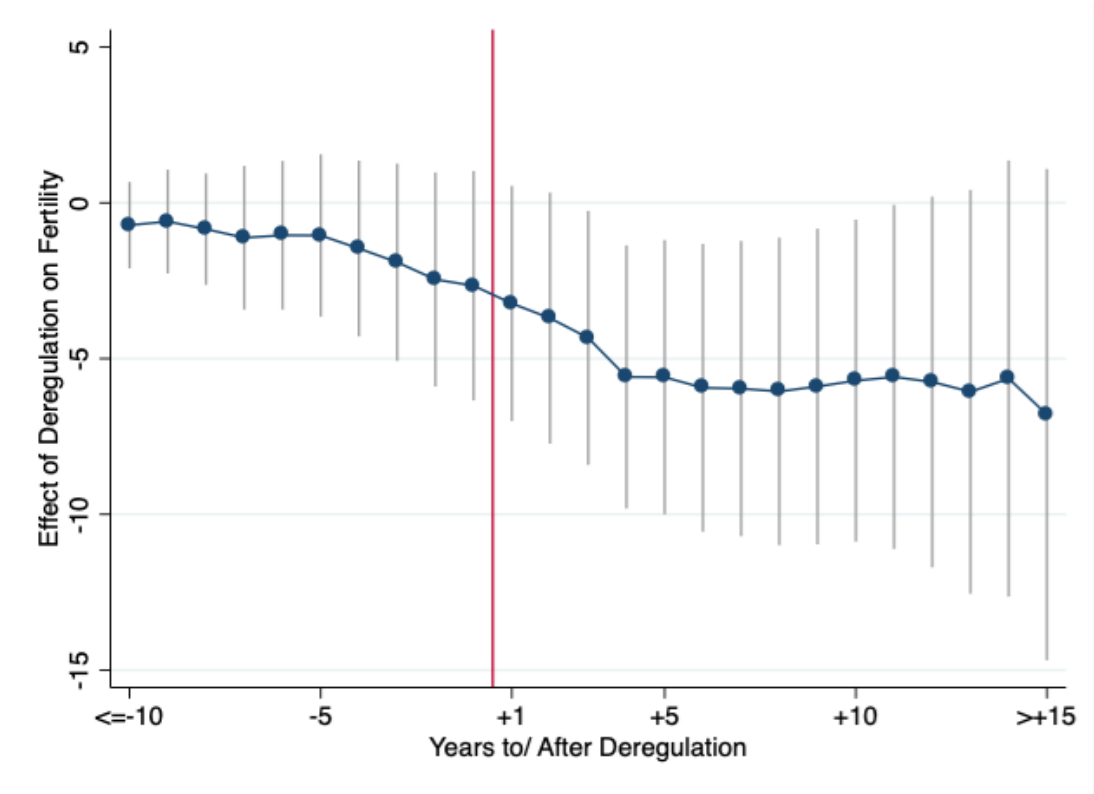


Figure 2.3: **Dynamic Impacts of Banking Deregulation on State Fertility Rate**
Notes: We show the coefficients of a difference-in-differences equation as in equation 2.2, but where we allow our binary dummy for deregulation to vary by year and include state controls: log population, fraction of population employed and fraction of population white. We limit years prior to deregulation to 15 and years after to 10, so the confidence intervals for the farthest lags (or leads) seem tighter since they include more lags (or leads). Specifically, we show the βs from the equation
$Y_{s,t} = \beta_0 + \sum_{j=-10}^{15} \beta_j D^j_{s,t} + \delta_s + \mathbf{X}_{s,t}\Gamma + \phi_t + \epsilon_{s,t}$

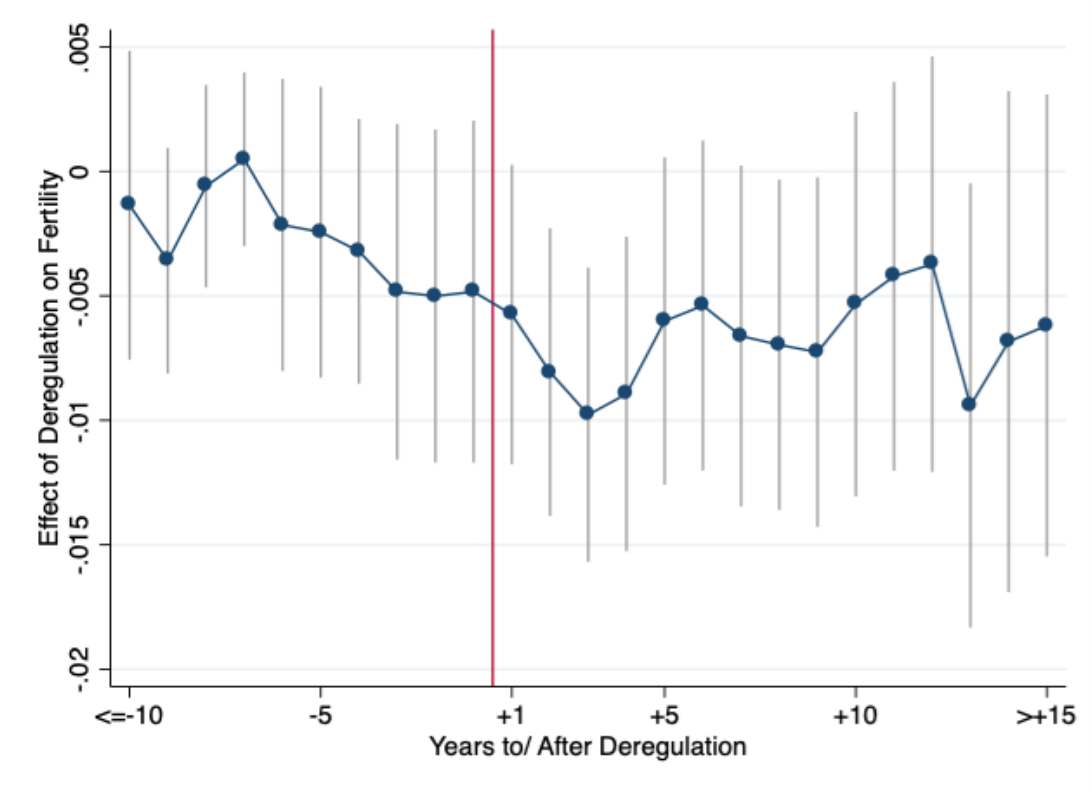


Figure 2.4: **Dynamic Impacts of Banking Deregulation on Individual Fertility**
Notes: We show the coefficients of a difference-in-differences equation as in equation 2.3, but we let our binary dummy for deregulation to vary by year and include state and individual controls. State controls include log population, the fraction of population employed and the fraction of population white. Individual controls include race, HS or more education, and a fourth-degree polynomial in age. We limit years prior to deregulation to 15 and years after to 10, so the confidence intervals for the farthest lags (or leads) seem tighter since they include more lags (or leads) and beyond. Specifically, we show the βs from the equation
$Y_{i,s,t} = \beta_0 + \sum_{j=-10}^{15} \beta_j D^j_{i,s,t} + \mathbf{X}_{s,t}\Gamma + \mathbf{Z}_{i,s,t}\Psi + \delta_s + \phi_t + \epsilon_{i,s,t}$

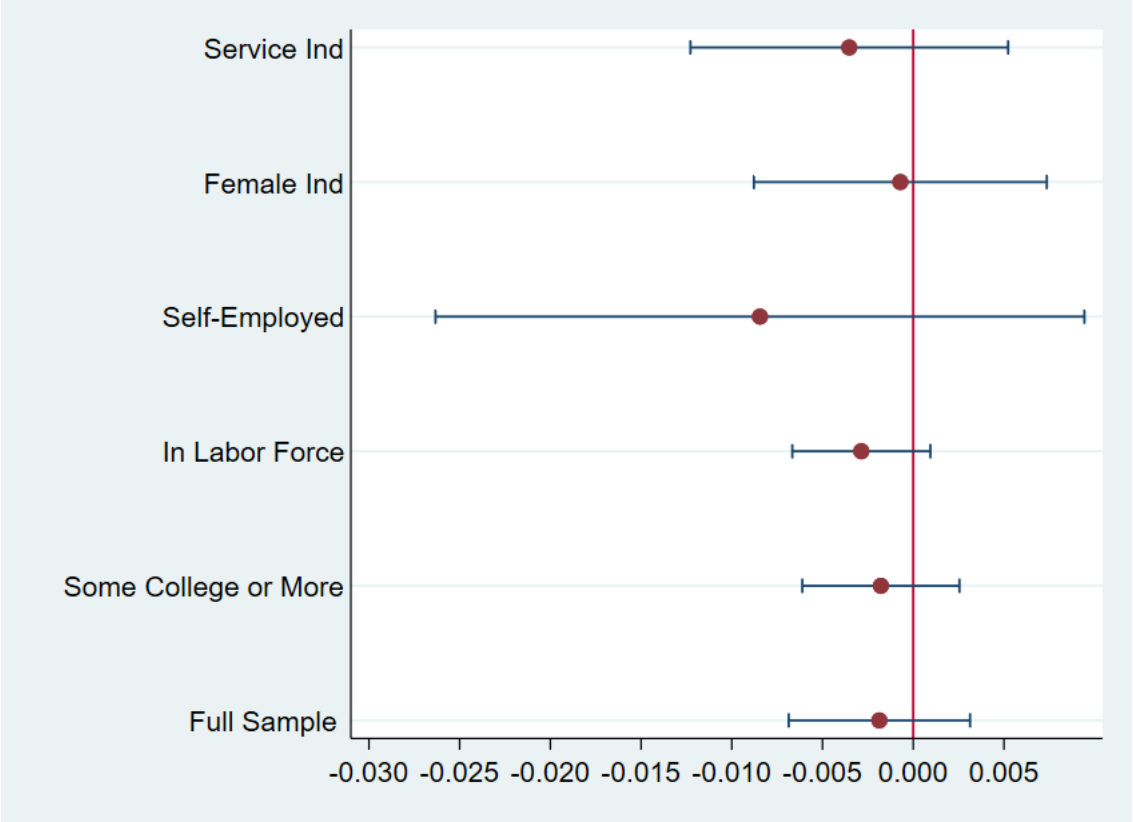


Figure 2.5: **Impacts of Deregulation on Fertility, by Sub-Groups of Women**

Notes: We show the coefficients of the βs from our individual difference-in-differences equation as in equation 2.3, but separately for different sub-groups of women. All specifications include state fixed effects, year fixed effects, state controls for log population, the fraction population white, and the fraction of the population employed. Individual controls include a 4th-degree polynomial in age, race, and high school or more education. Standard errors are clustered at the state-year level. Sample: Women aged 20-44 in the CPS March Supplement 1977-2000. Women are defined to be in service industries based on `ind1950` equal between 806 and 899. Women are defined to be in 'Female Ind' if the proportion of women working in that industry has more than the mean share of women across all industries in 1977.

2.8 Tables

Table 2.1: **Year of Intra-State Bank Deregulation For States**

State	Year Deregulated	State	Year Deregulated
Alabama	1981	Nebraska	1985
Arizona	1960	Nevada	1960
Arkansas	1994	New Hampshire	1987
California	1960	New Jersey	1977
Colorado	1991	New Mexico	1991
Connecticut	1980	New York	1976
District of Columbia	1960	North Carolina	1960
Florida	1988	North Dakota	1987
Georgia	1983	Ohio	1979
Idaho	1960	Oklahoma	1988
Illinois	1988	Oregon	1985
Indiana	1989	Pennsylvania	1982
Iowa	1999	Rhode Island	1960
Kansas	1987	South Carolina	1960
Kentucky	1990	Tennessee	1985
Louisiana	1988	Texas	1988
Maine	1975	Utah	1981
Maryland	1960	Vermont	1970
Massachusetts	1984	Virginia	1978
Michigan	1987	Washington	1985
Minnesota	1993	West Virginia	1987
Mississippi	1986	Wisconsin	1990
Missouri	1990	Wyoming	1988
Montana	1990		

Notes: States in the full sample in alphabetical order and the years in which they deregulated their banking sector. We do not include in our analysis SD and DE due to their unique banking structure, as well as AK and HI. We do include DC, which gives us a total of 47 states in our full sample.

Table 2.2: **Impact of Bank Reform on State-Level Fertility**

	(1)	(2)	(3)
Panel A: Full Sample			
Bank Reform	-6.3577***	-3.6512**	-3.0786***
	(2.3231)	(1.3823)	(1.0746)
Mean Dependent Var.	71.34	71.34	71.34
SD Dependent Var.	(14.63)	(14.63)	(14.63)
N	1,457	1,457	1,457
R^2	0.05	0.89	0.92
Panel B: 1970-1988			
Bank Reform	-3.8248	-3.5307**	-1.2502
	(3.5405)	(1.5625)	(0.8188)
Mean Dependent Variable	74.09	74.09	74.09
SD Dependent Variable	(15.53)	(15.53)	(15.53)
N	893	893	893
R^2	0.01	0.91	0.96
State and Year FE		✓	✓
State-Level Controls			✓

Notes: Coefficients from estimating a difference-in-differences equation 2.2 of the fertility rate on banking reform. Robust standard errors are clustered at state-level in parentheses. The fertility rate is defined as the number of births per 1000 women aged 20-44 in state s in year t from Bailey et al. (2016). State-Level controls in column (3) include the log of the population, the fraction of the population employed, and the fraction of the population white from Bureau of Economic Analysis (2020). The 'Full Sample' in Panel A covers 1970-2000 and the truncated final sample in Panel B covers 1970-1988. We do not include the states of SD, DE, AK, and HI.

Table 2.3: **Impact of Bank Reform on Individual-Level Fertility**

	(1)	(2)	(3)	(4)
Panel A: Full Sample				
Bank Reform	-0.0071***	-0.0056***	-0.0049***	-0.0044***
	(0.0015)	(0.0014)	(0.0013)	(0.0012)
Mean Dependent Variable	0.067	0.067	0.067	0.067
SD Dependent Variable	(0.25)	(0.25)	(0.25)	(0.25)
N	681,017	681,017	681,017	681,017
R^2	0.00	0.00	0.00	0.02
Panel B: 1977-1988				
Bank Reform	-0.0057**	-0.0033	-0.0026	-0.0019
	(0.0020)	(0.0023)	(0.0023)	(0.0023)
Mean Dependent Variable	0.0698	0.0698	0.0698	0.0698
SD Dependent Variable	(0.2548)	(0.2548)	(0.2548)	(0.2548)
State and Year FE		✓	✓	✓
State-Level Controls			✓	✓
Ind.-Level Controls				✓

Notes: Coefficients from estimating a difference-in-differences equation 2.3 of the dependent variable of giving birth in the past year on banking reform. Robust standard errors are clustered at the state-year level in parentheses. State-Level Controls in columns (4) and (5) include the log of the population, the fraction of the population white, and the fraction of the population employed in a state from Bureau of Economic Analysis (2020). Individual Controls in column (5) include race, high school or college education, and a 4th-degree polynomial in age. The 'Full Sample' in Panel A covers 1977-2000 and the truncated final sample in Panel B covers 1977-1988. Sample: Women aged 20-44 in the CPS March Supplement, not including the states of SD, DE, AK and HI.

Table 2.4: **Dynamic Impacts of Bank Reform on Fertility (State-Level)**

	(1)	(2)	(3)
Upto 2 Years Before Bank Deregulation	-5.5410**	-1.0237	0.5335
	(2.3759)	(1.0223)	(0.8298)
Year of Bank Deregulation	-4.8567*	-1.4491	0.4205
	(2.6487)	(1.2761)	(1.1027)
Year After Bank Deregulation	-5.7971*	-1.7938	-0.3857
	(3.3331)	(1.3229)	(1.1115)
2 Years after Bank Deregulation	-6.3875	-2.8758	-0.6692
	(4.1181)	(2.0432)	(1.2961)
$>$ 2 Years after Bank Deregulation	-4.1204	-7.4331**	-2.7999**
	(4.3374)	(3.3138)	(1.2484)
N	893	893	893
R^2	0.02	0.91	0.96
State and Year FE		✓	✓
State-Level Controls			✓

Notes: Coefficients from estimating a difference-in-differences equation of the fertility rate on banking reform, with lags and leads. Robust standard errors are clustered at the state-level in parentheses. The fertility rate is defined as the number of births per 1000 women aged 20-44 in state s in year t from Bailey et al. (2016). State-Level controls in column (3) include the log of the population, the fraction of the population employed, and the fraction of the population white from Bureau of Economic Analysis (2020). The final sample covers 1970-1988 and excludes the states of SD, DE, AK, and HI.

Table 2.5: **Dynamic Impacts of Bank Reform on Fertility (Individual-Level)**

	(1)	(2)	(3)	(4)
Upto 2 Years Before Bank Deregulation	-0.0044	-0.0031*	-0.0024	-0.0024
	(0.0026)	(0.0015)	(0.0015)	(0.0014)
Year of Bank Deregulation	-0.0053	-0.0043	-0.0035	-0.0027
	(0.0039)	(0.0034)	(0.0035)	(0.0039)
Year After Bank Deregulation	-0.0054	-0.0036	-0.0030	-0.0021
	(0.0033)	(0.0026)	(0.0025)	(0.0024)
2 Years after Bank Deregulation	-0.0082**	-0.0072**	-0.0063*	-0.0055
	(0.0034)	(0.0030)	(0.0033)	(0.0033)
$>$ 2 Years after Bank Deregulation	-0.0065*	-0.0072**	-0.0060*	-0.0049
	(0.0031)	(0.0032)	(0.0031)	(0.0031)
N	358,299	358,299	358,299	358,299
R^2	0.00	0.00	0.00	0.03
State and Year FE		✓	✓	✓
State-Level Controls			✓	✓
Individual-Level Controls				✓

Notes: Coefficients from a difference-in-differences regression of the dependent variable of giving birth within the last year on banking reform, with lags and leads. Robust standard errors are clustered at the state-year level in parentheses. State-Level Controls in columns (3) and (4) include the log of the population, the fraction of the population employed, and the fraction of the population white in a state. Individual Controls in column (4) include race, high school or college education, and a 4th-degree polynomial in age. Sample: Women aged 20-44 in the CPS March Supplement 1977-1988, not including the states of SD, DE, AK, and HI.

Table 2.6: **Impact of Bank Reform on Other Outcomes Related to Fertility and Marriage**

	Married	Divorced	Age at 1st Child	No. Children	Married Fertility
	(1)	(2)	(3)	(4)	(5)
Bank Deregulation	-0.0060	0.0045	0.0318	-0.0293**	-0.0009
	(0.0049)	(0.0048)	(0.0428)	(0.0110)	(0.0025)
Mean Dependent Var.	0.64	0.13	22.7	1.39	0.09
SD Dependent Var.	(0.49)	(0.33)	(4.07)	(1.37)	(0.29)
N	358,270	358,270	231,086	358,270	230,619
R^2	0.14	0.03	0.22	0.23	0.06

Notes: Coefficients from a difference-in-differences regression of the dependent variable on banking reform, including state and year fixed effects Robust standard errors are clustered at the state-year level in parentheses. All estimations include state and individual-level controls. State-level controls include the log of the population, the fraction of the population white, and the fraction of the population employed in a state from Bureau of Economic Analysis (2020). Individual controls include race, high school or more education, and a 4th-degree polynomial in age. Sample: Women aged 20-44 in the CPS March Supplement 1977-1988, not including states of SD, DE, AK, and HI.

Table 2.7: **Impact of Bank Reform on Individual Fertility by Household Income and Race**

	Lowest Quartile	2nd Quartile	3rd Quartile	Highest Quartile
	(1)	(2)	(3)	(4)
Panel A: White Women				
Bank Deregulation	-0.0052	-0.0009	0.0032	0.0028
	(0.0057)	(0.0052)	(0.0055)	(0.0030)
Mean Dependent Variable	0.08	0.08	0.07	0.05
SD Dependent Variable	(0.28)	(0.27)	(0.25)	(0.21)
N	68,512	77,880	80,656	83,535
R^2	0.04	0.03	0.03	0.03
Panel B: Non-White Women				
Bank Deregulation	-0.0247	-0.0101	0.0042	-0.0108
	(0.0162)	(0.0154)	(0.0087)	(0.0175)
Mean Dependent Variable	0.08	0.07	0.06	0.05
SD Dependent Variable	(0.28)	(0.26)	(0.23)	(0.22)
N	20,488	11,151	8,503	7,545
R^2	0.04	0.03	0.03	0.03

Notes: Coefficients from a difference-in-differences regression of giving birth in the last year on deregulation. All estimations include state and year fixed effects, as well as state and individual level controls. State controls include the log of the population, the fraction of the population white, and the fraction of the population employed in a state from Bureau of Economic Analysis (2020). Individual controls include high school or college education, and a 4th-degree polynomial in age. Robust standard errors are clustered at state-year level in parentheses. Sample: Women aged 20-44 in the CPS 1977-1988, not including the states of SD, DE, AK and HI.

Table 2.8: **Impact of Bank Reform on Individual Fertility by Homeownership Status**

	Home-Owners	Non Home-Owners
	(1)	(2)
Bank Deregulation	-0.0010	-0.0046
	(0.0022)	(0.0059)
Mean Dependent Variable	0.06	0.08
SD Dependent Variable	(0.24)	(0.26)
N	222,816	135,454
R^2	0.04	0.02

Notes: Coefficients from a difference-in-differences regression of giving birth in the last year on deregulation, by home-ownership status. All estimations include state and year fixed effects, as well as state and individual level controls. State controls include the log of the population, the fraction of the population white, and the fraction of the population employed in a state from Bureau of Economic Analysis (2020). Individual controls include high school or college education, and a 4th-degree polynomial in age. Robust standard errors are clustered at the state-year level in parentheses. Sample: Women aged 20-44 in the CPS 1977-1988, not including SD, DE, AK, and HI.

Chapter 3: Trying to Beat the Heat: Air-Conditioning and Learning

3.1 Introduction

Many environmental factors such as temperature, noise, light, and pollution impact human performance (Echeverria et al., 1994). The negative effects of heat has received particular attention due to its ubiquitous and widespread nature (Jokl, 1982; Ramsey, 1995; Barreca et al., 2016; Kjellstrom et al., 2016). High temperatures have been found to decrease productivity not just in physically demanding jobs such as agriculture, sports, or construction (Hancher and Abd-Elkhalek, 1998; Wendt et al., 2007; Yi and Chan, 2017) but also in sedentary work environments (Seppanen et al., 2006; Kjellstrom et al., 2009; Heal and Park, 2016).

Similar to work environments, high heat has also been shown to cause losses in productivity in learning environments (Cho, 2017; Park et al., 2020). Hot classrooms may prevent children from learning effectively, and teachers from teaching effectively, due to discomfort, exhaustion, or slowed cognition. High temperatures may also lead to increased absenteeism in schools (Randell and Gray, 2016, 2019). A potential policy solution to alleviate these learning losses is for schools to invest in air-conditioning (AC). However, estimating the causal impact of AC on student performance is difficult since AC installation is typically done in conjunction with other infrastructure spending in schools (Cellini et al., 2010; Neilson and Zimmerman, 2014). As such, the current literature is limited to providing correlations between AC coverage and student

performance (Park et al., 2020) and there is little causal evidence as to whether installing AC in schools is an effective tool in improving student outcomes.

In 2013, then Mayor of Chicago, Rahm Emanuel announced that $135 million would be spent to install AC in all classrooms in Chicago Public Schools (Corley, 2013). The campaign was motivated by the sweltering summer temperatures in Chicago (which can reach 90F) as well as reports of inhumane classroom conditions cited by teachers during the district's lengthy teachers' union strike in 2012 (Strauss, 2012; Chambers, 2013). This announcement led to one of the largest ever investments in AC made by a public school district and installed AC in more than 200 schools over the next four years.

We exploit the roll-out of this campaign to study the impacts of AC installation on student performance. During this roll-out, AC was installed in schools over four different waves, starting in the school year 2013-14 and ending in 2016-17 (Chicago Public Schools Press Releases, 2016). We leverage the staggered timing of AC installation across schools using a difference-in-differences strategy to compare students exposed to AC to those that are not, before and after AC installation.

Despite costing $135 million, we find no evidence that the AC installation campaign in Chicago improved student achievement. Using our difference-in-differences design, we find that students whose schools received AC saw no significant improvements in their test scores compared to those whose schools did not. In addition to test scores, we find no significant impact of AC installation on the probability of being held back a grade. Since there may be disruption effects to concurrent test scores in the year of installation and students may only see potential gains in later years, we also look at the impact of AC installation on test scores for each year post treatment. We find there are no significant positive impacts of AC installation for students in

the treated schools even several years after treatment. We also find some evidence that average student attendance at the school level increased by approximately half a day per school year after AC installation. However, for attendance, differentials pre-trends exist.

One potential concern is that some treated schools already had existing AC infrastructure and the Chicago Public Schools campaign merely replaced or updated already functioning AC units. In this case, including schools with preexisting AC could attenuate the estimated effects of the program. To help account for this, we use data on preexisting AC infrastructure in each school that provides information on what fraction of the school was air-conditioned prior to the campaign. After accounting for prior AC infrastructure, we still find no evidence of significant positive impacts of AC installation on student test scores or the probability of being held back a grade.

Classroom AC is one of the many aspects of educational inequality. On top of being less likely to have AC in their schools, due to residential sorting, low-performing and low-income students also have fewer environmental amenities in their neighborhoods such as poorer air quality and hotter temperatures (Banzhaf and Walsh, 2008). Thus, they may face larger test score declines due to heat exposure in schools (Park et al., 2020). We analyze the impact of AC on low-performing students and find null results similar to those found in the full population of students.

Our results show no evidence that the installation of AC in Chicago Public schools had a positive impacts on student achievement and limited evidence of a positive impact on school-level attendance. This analysis covers a temperate region of the United States where temperatures can range from average lows of 17F in January to average highs of 85F in July (see Figure 3.6) and where the typical year has 137 days over 70F and 76 days over 80F, of which approximately 77

and 35 days fall within the school year, respectively (see Figure 3.7).[1] Estimating the causal effect of AC installation on student learning for this type of climate has both benefits and drawbacks. The major drawback is that we cannot determine whether AC is an ineffective tool in combating the detrimental effects of heat in schools or whether there are no detrimental effects of heat on learning in temperate climates like Chicago.[2] On the other hand, this region is similar to other school districts that are on the margin of investing in air-conditioning. As such, our findings have important policy implications. Many large school districts in the US are not fully air-conditioned such as New York City, Philadelphia, Baltimore City, Denver, and Detroit and many are considering large-scale AC installation projects (Barnum, 2017). For example, Mayor DeBlasio announced in 2017 that the Department of Education would spend $29 million to air-condition every classroom in New York City by 2022 (NYC City Hall, 2017). Our results speak directly to the potential benefits (or lack thereof) of these expensive AC infrastructure projects on student learning.

Given the strict budget constraints faced by many public school districts (e.g. Chicago Public Schools cited a deficit of $1 billion in 2013 (Corley, 2013)), our results suggest that the $135 million investment in AC might have been better spent on other educational resources. Even when using estimates at the top of the 95 percent confidence interval, the AC installation program returned test-score gains of less than 0.02 standard deviations, although we acknowledge that test-score gains and other academic benefits may not be the only goal of AC installation.

We outline the context of the AC installation program in Chicago and describe our primary

[1] Monthly averages from 2000-2020 (Lawrimore et al., 2016), daily normals from 2000-2020 (Arguez et al., 2020).

[2] In our analysis, we use end-of-year test scores to measure student learning. As such, we cannot differentiate between the impact of AC on test scores due to changes in student learning accumulated throughout the school year or simply through changes in student performance on test day.

sources of data in section 3.2. Section 3.3 details our methodology followed by a discussion of the results in section 3.4. Section 3.5 outlines the policy implications of our results and concludes.

3.2 Background and Data

To estimate the impact of AC installation on students' academic outcomes we leverage the roll-out of AC in Chicago Public Schools (CPS) from 2013 to 2017. CPS is the third-largest school district in the US (after New York City and Los Angeles) with 355,156 students enrolled across 642 schools in the 2018-19 school year (Chicago Public Schools, 2020). The public school system in Chicago serves an ethnically diverse student body, of which the largest proportion of students are Hispanic (46.6%) and the next largest are Black (35.9%). The district categorizes more than three quarters of the student population as coming from 'Economically Disadvantaged' households. In addition, the district has a history of poor academic performance. Since being called the "worst public school system in the nation" in 1988 by the U.S. Secretary of Education William Bennet, CPS has made vast improvements in high school graduation rates and test scores, but still fares poorly on college readiness nationally (Luppescu et al., 2011; Max, 2023).

In 2012, the Chicago Teachers Union went on a nine-day strike to protest teacher evaluations, pay, and classroom conditions (Pearson and Yan, 2012). The issue of sub-optimal classroom conditions rose again during teacher strikes and protests in 2013 (Chambers, 2013; Ahmed-Ullah, 2013; Peralta, 2013). Partly in response to these concerns, then mayor of Chicago, Rahm Emanuel, announced that $135 million would be spent to install AC in all previously non air-conditioned schools (Chicago Public Schools Press Releases, 2016) – thus providing air conditioning to all students while learning in the classroom. This large expenditure on AC installation occurred

despite CPS facing a looming $1 billion budget deficit which forced CPS to close 47 under-performing schools and provoked city-wide protests in 2013. In defence of these school closures, Mayor Emanuel highlighted that the funds recouped could be better spent on other programs such as "access to libraries, iPads, and air-conditioned classrooms" (Corley, 2013).

The campaign to install AC was implemented in four waves across 212 schools. Using CPS press releases we identify which schools received AC in each of the four waves. Of the 212 schools that received AC, 67 schools received AC during Wave 1 in which installation occurred while school was in session during the 2013-14 school year. In Wave 2, 56 schools received AC during the summer of 2014. In Wave 3, 29 schools received AC in October of 2014. In the fourth and final wave, 60 schools received AC during the 2016-17 school year but prior to spring of 2017. A full list of treated schools by wave of AC installation can be found in Table 3.3.

To measure the academic performance of students, we obtain student-level test scores for math and English from school years ending in 2008 to 2017 for students in grades 3-8 for 603 Chicago schools from the Illinois State Board of Education. The test scores come from standardized tests administered at the end of the year for all students in Illinois.[3] This test is known as a "high stakes" test in the state and is used both to help determine whether a student advances to the next grade and by administrators to evaluate school performance. We normalize student test scores by year and grade using the full Illinois state distribution of test scores. In addition to test scores, we obtain a measure of grade retention (i.e. 'Held Back' a grade) which is a binary variable equal to 1 if a student repeats the same grade. Finally, we also obtain a

[3]Prior to 2015, the Illinois State Board of Education used the Illinois Standards Achievement Test for students in grades 3-8 in math and reading (which we refer to as English for the rest of the paper). Starting in 2015, the State Board mandated all schools to implement the Partnership for Assessment of Readiness for College and Careers test which was created to better reflect the new and updated Common Core standards and replace previous state-wide assessments for all students in grades 3-8 (Citizens For Public Schools, 2017). These are the test scores we use for years 2015-2017 for math and English Language Arts (which we also refer to as English).

school-level measure of the average fraction of days students attend each year. .

Since test scores are only available for 3-8 grades, our analysis does not look at the impact of AC on high school students. Most students in Chicago attend an elementary school from kindergarten to 8th grade, followed by four years of high school. Thus, of the 212 schools that received AC, only the 183 elementary schools and 2 middle schools appear in our sample.[4] In addition, there are 417 'control' schools in our dataset that do not receive AC during this campaign since they were presumed to already have AC installed.[5]

In addition to student outcomes, we gather data on existing AC infrastructure in each school prior to the campaign roll-out. Between 2009 and 2011, the Energy Star Portfolio Manager system collected data on the percentage of school facilities that were air-conditioned and on other physical attributes of the schools (as required by the U.S. Environmental Protection Agency). Of the schools for which we have test score and AC installation data, the Energy Star System has information on approximately 60 percent of those schools. Table 3.1 shows the differences in physical attributes of these schools by treatment status.[6] We consider schools to be 'treated' if they receive AC as part of any of the four waves of AC installation between 2013 and 2017, while the remaining schools that do not receive AC as part of the CPS AC installation during these 4 waves are designated as 'controls'. As can be seen in Table 3.1, treated schools are substantially less air-conditioned than control schools. On average, treated schools had only one-third of their facilities air-conditioned by 2011 while control schools had more than two-thirds of

[4]Of these 185 schools, 66 schools received AC during Wave 1, 50 schools in Wave 2, 29 schools in Wave 3, and 40 in Wave 4.

[5]We only have data on prior AC percentage for 217 of our control schools, which on average have 67% AC as seen in Table 3.1.

[6]Table C.1 shows the difference in physical attributes of these schools both by treatment status and separately for each wave. In addition, both Tables 3.1 and C.1 match the analysis sample and remove 46 schools as discussed in detail in the methodology section.

their facilities air-conditioned. While most of the treated schools had little to no AC prior to the AC installation program, not all control schools were fully air-conditioned. To better illustrate the difference in preexisting AC infrastructure between treated and control schools, Figure 3.1 provides a histogram of the fraction of the schools air-conditioned by treatment status. In addition to preexisting AC infrastructure, control schools are significantly newer and have a lower share of black and low-income students. Also, students in control schools are less likely to be held back, have higher attendance rates, and have higher math and English test scores. In contrast to the discrepancies in AC, all schools are heated.

3.3 Methodology

The AC installation campaign in Chicago provides a natural experiment to measure the potential benefits of having AC in schools on student performance. In particular, the staggered roll-out of AC to schools allows for a straightforward difference-in-differences approach to identify the causal impact of AC on student performance. It also allows us to solely estimate the effect of AC, separate from any other concurrent infrastructure expenditure.

To estimate the effect of having AC in a school on student performance, we estimate a standard difference-in-differences model as follows:

$$y_{ist} = \alpha + \beta \textit{Have AC}_{st} + \gamma_1 \textit{Math}_{is,t-1} + \gamma_2 \textit{English}_{is,t-1} + \mu_s + \lambda_t + \varepsilon_{ist} \tag{3.1}$$

where y_{ist} is the normalized test score (or held back indicator) of student i in school s in year t. *Have AC*$_{st}$ is an indicator equal to one if school s has AC in year t. This variable is equal to one for control schools in all years and equal to one for treated schools starting in the year they

receive AC (and zero before). In addition, we include controls for lagged math ($Math_{is,t-1}$) and English ($English_{is,t-1}$) test scores, school fixed effects (μ_s), and year fixed effects (λ_t). Including lagged math and English test scores allows us to measure the change in test scores from year to year instead of test score levels.[7] This allows us to measure the yearly value-added of having AC in a school on student achievement. Thus, the main coefficient of interest, β, measures the difference in the change in test scores before and after AC installation between students whose schools received AC versus those who did not. We also estimate the effect of AC on attendance. Since our attendance data is at the school level instead of the student level, we estimate equation 3.1 at the school level without including individual lagged test scores. For all estimates we cluster standard errors at the school level.

The main identifying assumption for this model is that the outcomes for treated and control groups ould have parallel trends in the absence of treatment. In our setting, this assumption requires that had the treated schools not received AC, their scores would have moved in parallel with the control schools (which already had atleast some AC). While the counterfactual parallel trend assumption cannot be observed, we can test for parallel trends prior to the treatment. We plot the average test scores of students for each year by treatment status in Figure 3.2 separately for each wave of treatment. Similar figures can be seen for being held back and attendance in Figure 3.4. These figures show that the test scores and the probability of being held back of the treated and control schools appear to move in parallel prior to AC installation. However, for attendance there is some evidence of a pre-trend, particularly for wave 4 schools. To formally test for parallel pre-trends, Figure 3.3 and 3.5 plots the coefficient on $Have\ AC_{st}$ from Equation 3.1

[7] Appendix Table C.2 reports our main estimates without controls for lagged math and English test scores and find similar results.

interacted with each year. These figures show that there is no statistically significant difference in the trend between treated and control schools prior to treatment in each year for each wave for test scores and the probability of being held back. However, again there is a statistically significant pre-trend for attendance for wave 4 schools.

Additionally for the counterfactual parallel trends assumption to hold, there would need to be no other concurrent policy changes that would differently affect AC-receiving schools versus control schools. While that appears to be true for later waves, the AC installation in wave 1 schools coincided with the closure of 47 'under-performing or under-utilized' schools by CPS in the summer of 2013. Students who previously attended these closed schools were assigned by CPS to 48 'Welcoming Schools' (De la Torre et al., 2015). In our data, we observe 46 of the designated 48 'Welcoming Schools'. Of the 66 schools that received AC in wave 1, 33 were 'Welcoming Schools' and 33 were not (while only 13 of 417 control schools were 'Welcoming Schools'). As such, half of the schools treated in wave 1 were simultaneously impacted by being a 'Welcoming School', while few control schools were. Since most of the closed schools were under-performing, 'Welcoming Schools' saw a large influx of low test-score students to their school in 2013 and consequently saw large declines in their average test scores. Therefore, when estimating the impact of AC on student achievement in these 'Welcoming Schools', there will likely be a negative bias because the timing of AC installation coincides with welcoming new low-performing students from closed schools (see Appendix Figure C.1). To account for this potential bias, our main analysis omits these 46 assigned 'Welcoming Schools' from the sample.[8]

Equation 3.1 estimates the impact of AC on student performance based on schools being

[8]In Appendix Figures C.2, C.3 and Table 3.4 we show the results for wave 1 for both 'Welcoming Schools' and 'Non-Welcoming Schools'. While we find null effects for 'Non-Welcoming Schools', the estimates for 'Welcoming Schools' show moderate negative effects on test scores of AC installation consistent with a negative bias due to simultaneously welcoming low-performing students.

part of treated schools in the CPS AC installation program. However, we might be concerned that this does not mean that treated schools experience a binary change in going from zero AC to being fully air-conditioned. To help address this concern we use the Energy Star data on existing AC infrastructure for 354 of the schools prior to the campaign roll-out. Figure 3.1 shows the distribution of AC for treated and control schools in this sample. This figure shows that the modal treated school (32% of schools) had 10% of their school air-conditioned while the modal control school (43% of schools) had 100% of their school air-conditioned. However, 22% of treated schools had more than 50% of their school air-conditioned. Thus, most treated schools already had some non-zero percentage of AC in their school prior to the CPS installation. Hence using an intent-to-treat variable for AC installation in the difference-in-differences model could attenuate the estimates of actually having AC installed in schools.

Therefore, we use an alternative specification to account for the prior AC infrastructure in treated schools:

$$y_{ist} = \alpha + \beta \textit{Fraction AC}_{st} + \gamma_1 \textit{Math}_{is,t-1} + \gamma_2 \textit{English}_{is,t-1} + \mu_s + \lambda_t + \varepsilon_{ist} \tag{3.2}$$

where all variables are the same as in Equation 3.1 except the *Fraction* AC_{st} variable which takes a value from 0 to 1 and is the fraction of the school that was air-conditioned prior to the treatment. For treated schools the value of *Fraction* AC_{st} changes to 1 for all years after AC installation. As such, β measures the impact of a school moving from no AC to being fully air-conditioned on student outcomes. The results for the specifications in Equations 3.1 and 3.2 are reported in Table 3.2.

3.4 Results

3.4.1 Descriptive Results

Before directly estimating our difference-in-differences model, we first look at the trends in student outcomes between treated and control schools. This allows us to test for an effect of AC on student performance after AC installation in the raw data. If AC installation has a positive impact on students, we expect student outcomes to improve in treated schools after treatment relative to control schools. Therefore, in Figures 3.2 and 3.4, we plot the average standardized test scores for math and English, held back, and attendance separately for each wave of treatment, by treatment status over each year. In each sub-graph, the dashed lines represent the treated schools and the solid lines represent the control schools. The vertical line marks when schools in each wave received AC.

In Figure 3.2, we see little evidence that AC installation improved student test scores. For all waves of treatment, we do not see the treated schools' standardized test scores converge post-treatment towards the control schools. In addition, there appears to be parallel trends prior to the treatment. In Figure 3.4 we see little evidence that AC installation decreased the likelihood of being held back, however, there is some evidence of increased attendance for schools treated in wave 1. Overall, this evidence suggests that AC installation had little to no impact on student achievement while potentially improving attendance.

3.4.2 Difference-in-Differences

Next, we estimate the difference-in-differences model using Equation 3.1 and report the results in Panel A of Table 3.2. If installing AC provides better learning conditions for students and teachers, then we would expect positive impacts of AC installation in treated schools post-treatment.

In Panel A of Table 3.2, we find no evidence that students in treated schools saw their math or English test scores improve as compared to students in control schools after AC was installed. Students in treated schools saw statistically insignificant decreases of 0.007 standard deviations in their average math test scores and 0.001 standard deviations in their English test scores post AC installation as compared to control schools. While also statistically insignificant, we estimate that students in schools that received AC were 0.001 percentage points (or 5.3 percent) more likely to be held back after AC was installed. However, for attendance we find that treated schools saw a 0.003 percentage point (or 0.3 percent) increase in attendance. Appendix Table 3.5 shows that the results are very similar when taking into account potential time-varying heterogeneous treatment effects using Borusyak et al. (2023). These results can rule out relatively modest positive impacts of AC installation on student test scores. When measuring returns at the top of the 95th percent confidence interval, the positive impact of AC installation would only increase math and English test scores by 0.016 and 0.020 standard deviations, respectively.

Park et al. (2020) find the beneficial impacts of AC penetration (as measured by survey data from high school counsellors) are larger for marginalized students. The most vulnerable students may be unable to counter the stress of heat at school by going home to an air-conditioned environment. Thus, any potential positive impacts of AC installation may be concentrated on the

already low performing students. To test this, we estimate Equation 3.1 for students in the bottom quartile of the test score distribution in both math and English. These results are presented in Table C.3. We find nearly identical results for students in the bottom of the test score distribution as we do in our full sample. These results show no evidence that AC provides any positive impacts on academic performance – even for low-performing students who may be the most vulnerable to heat in schools.

One may potentially expect negative impacts of AC in the year that AC was installed due to disruption effects from the installation process or construction. Conversely, positive impacts of AC on student achievement could occur some years after AC was installed in schools. To investigate this heterogeneity by years post-treatment, Figures 3.3 plots the coefficients from Equation 3.1 for test scores while allowing them to vary flexibly by each year. For all waves we see that there are no statistically significant impacts after treatment (including no disruption effects in the year of installation), on either math or English test scores. In addition, there are no large differences in estimates before versus after the treatment occurs. These figures also confirm that for each wave we observe parallel pre-trends between treated and control schools, given by confidence intervals that overlap zero for all pre-period estimates. In addition, the lack of heterogeneity across years implies that yearly variation in temperature (at least for the 4 post years) appears to have a limited interactive effect.[9] Hence, even after breaking down the impacts of the AC program by years after installation, we find no evidence of positive impacts on student test scores.

Analogous to Figure 3.3, Figure 3.5 shows the coefficients by year for held back and

[9]For the period of sample, the number of high temperature days in Chicago remain fairly consistent with approximately 25% of school days with a temperature above 70F and less than 5% of school days with a temperature above 90F (see Figure 3.7).

attendance. For attendance there are substantially different patterns by the prior characteristics of the treated schools. For schools treated in waves 2 and 3 – that have similar characteristics to the control schools (see Table C.1) – there are no statistically significant impacts after treatment. For schools treated in wave 1 and 4 – that are lower performing compared to control schools (see Table C.1) – their attendance begins to converge to control schools starting in 2013 (see Figure 3.4). As such, for wave 4 schools who were treated in 2016, Figure 3.5 shows a clear pre-trend. For wave 1 schools the trend occurs right after treatment with no abrupt discontinuity at the treatment year. If AC improves classroom conditions for students, we would likely see an abrupt increase in attendance post AC installation. However, for all waves there are no large differences in estimates right before versus after the treatment occurs and only wave 1 shows a differential trend post treatment. Overall, these results suggest taking the positive effects of AC installation on attendance found in Table 3.2 with caution. In addition, Figure 3.5 show no statistically significant impact after treatment on the likelihood of being held back.

In addition to heterogeneity across years, there may also be heterogeneity by wave due to when different types of schools were assigned to receive AC. Thus, in Panel A of Table 3.6 we report the results of Equation 3.1 separately for each wave. For all waves, we find no positive impacts of AC installation on standardized math or English test scores, or on the probability of being held back in the post-treatment years when compared to students in control schools that did not receive AC. For attendance we find positive effects of AC installation for schools in waves 1 and 4. However, due to the patterns seen in Figure 3.5, these results on attendance should be taken with caution.

Lastly, we test for heterogeneity across grade. Park et al. (2021) find negative effects of heat that are fourfold larger for students in grades 3-5 than grades 6-8. This would suggest there

may be positive effects of AC on student performance for earlier grades but not later grades. In figure 3.8, we show our main results from Equation 3.1 separately for each grade. For both math and English we find no distinguishable difference between the grades.

3.4.3 Energy Star Difference-in-Differences

As discussed in the Methodology section, the above results estimate the impact of a school being part of the CPS AC installation program rather than a change in student performance after a binary change from having no AC to being fully air-conditioned. As shown in Figure 3.1 a substantial number of treated schools already had some non-zero percentage of AC infrastructure in their schools prior to the AC installation campaign. To account for this, we estimate Equation 3.2 as outlined in the Methodology section, which does not just measure the impact of being assigned to a school that receives AC, but modulates the treatment by using information on prior AC infrastructure within the treated schools. Thus, this specification measures the impact of being at a treated school that goes from having no AC to being fully air-conditioned on student achievement. [10]

These results are reported in Panel B of Table 3.2. We find that the estimates are very similar to those in Panel A – although they are slightly more positive and have larger standard errors (as expected due to the reduced number of schools in the sample). Going from having no AC at all to being fully air-conditioned saw a statistically insignificant increase of 0.004 standard deviations on math test scores in post-treatment years for students in the treated schools as compared to control schools, and a statistically insignificant increase of 0.009 standard deviations on English

[10] Alternatively, we also estimate Equation 3.1 restricting the sample of treated schools to only those that had less than 30% of the school air-conditioned prior to being treated. Similar to the full sample, we find null results when making this restriction (see Table C.5).

test scores. In addition, column (3) shows no evidence that going from no AC to being fully air-conditioned impacted the likelihood of a student being held back. However, for attendance the effect sizes are nearly twice as large with a point estimate of 0.005 percentage points (or 0.5 percent). The impacts for the low-performing students are also similar to the full sample (see Figure C.3).[11] With the larger standard errors, when measuring returns at the top of the 95 percent confidence interval, going from having no AC to being fully air-conditioned would increase math and English test scores by 0.045 and 0.041 standard deviations, respectively.

While our results show little to no evidence that the installation of AC had a positive impact on student achievement, we are unable to distinguish whether AC is an ineffective tool in combating the detrimental effects of heat in schools or whether there are no detrimental effects of heat on learning in temperate climates like Chicago. Ideally, we would like to directly estimate the impact of heat on student learning in Chicago over this time period. However, we only have weather variance in Chicago over the nine years in our data and annual end of year test scores for students. While this technically allows us to estimate the direct impact of heat on test scores, all estimates will be based off of this very small sample size. Additionally, over this nine year period there is minimal variation in the number of hot days during the school year (see Figure 3.7). With this very limited data, we find in Table 3.7 that the number of hot days in a school year does not significantly impact student test scores. While these results should be taken with an abundance of caution, they suggest that there is perhaps little detrimental effect of heat in Chicago and, therefore there is little to no margin for AC to be an effect policy tool in this type of temperate climate.

[11] The estimates by wave of treatment are similar to the full sample, and are reported in Panel B of Appendix Table 3.6.

3.5 Conclusion

Although there are well-documented detrimental impacts of heat, our results demonstrate that the AC installation program in Chicago had little impact on students' academic performance. These results are robust to different specifications, sub-populations, and heterogeneity by years post-treatment.

Chicago Public Schools spent $135 million dollars in fixed costs on their AC installation program. This expense averaged to nearly $730,000 per school or $2,600 per student – not including the operational costs such as electricity and maintenance. While AC installation may have improved outcomes along other dimensions, our estimates demonstrate that the AC program resulted in high costs with no observable academic benefits as measured by end-of-year test scores. Even when measuring returns at the top of the 95 percent confidence interval, the $730,000 spent per school led to relatively small test score gains for students of less than 0.02 standard deviations for each future year that the AC unit remains operable (which may vary from 12-20 years depending on upkeep).

Compared to other policy interventions, the Chicago AC installation program compares poorly in terms of test-score improvements, even when we use the 95 percent confidence interval to estimate returns. A meta-study by Fryer Jr (2017) shows that the average returns to school-based educational interventions are 0.05 standard deviation improvements in math and 0.07 standard deviation improvements in English test scores for students. Chetty et al. (2014) show that an improvement in teacher value added by one standard deviation improves math test scores by 0.14 standard deviations and English scores by 0.1 standard deviations. Per Krueger (1999), decreasing student class sizes by one-quarter in Project STAR increased test scores by 0.2 standard

deviations. Alternatively, if the policy goal is to improve racial or SES disparities in student test performance, interventions like high-dosage tutoring may be more effective (Fryer Jr and Howard-Noveck, 2020). While the AC installation program in Chicago may have improved the comfort of the learning environment for students and teachers, this change in environment did not translate to test-score improvements as in other interventions.

Policymakers in Chicago intended to reduce infrastructural disparities between schools and as such improve student performance by installing AC in schools. However, the program had little to no effect in closing the student performance gap between treatment and control schools. Therefore, given Chicago Public Schools' $1 billion deficit (Corley, 2013), the limited funds may have been better spent on other educational interventions if test-score gains was their main objective.

While most schools in the southern United States already have AC in their classrooms, the question of AC installation is still being considered by many school districts in temperate climates such as New York City, Philadelphia, Baltimore City, Milwaukee, Denver, Hawaii, Detroit, Jefferson County, and Long Beach (Barnum, 2017). The results of Chicago's AC installation program from this paper can help guide other marginal school districts when making the expensive choice of whether or not to install AC in classrooms.

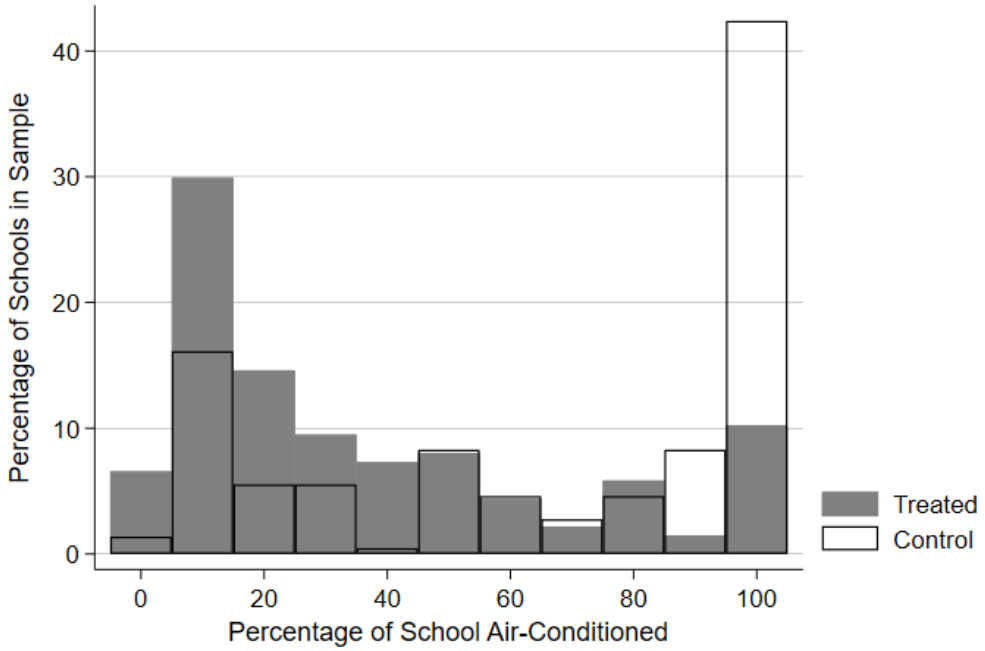


Figure 3.1: **Percentage of School Air-Conditioned by Treatment Status**
Notes: The figure shows the distribution of control and treated schools by decile for the percentage of the school that is air-conditioned prior to the AC installation program. A value of 100 implies the school is fully air-conditioned, while a value of 0 implies the school has no air-conditioning.

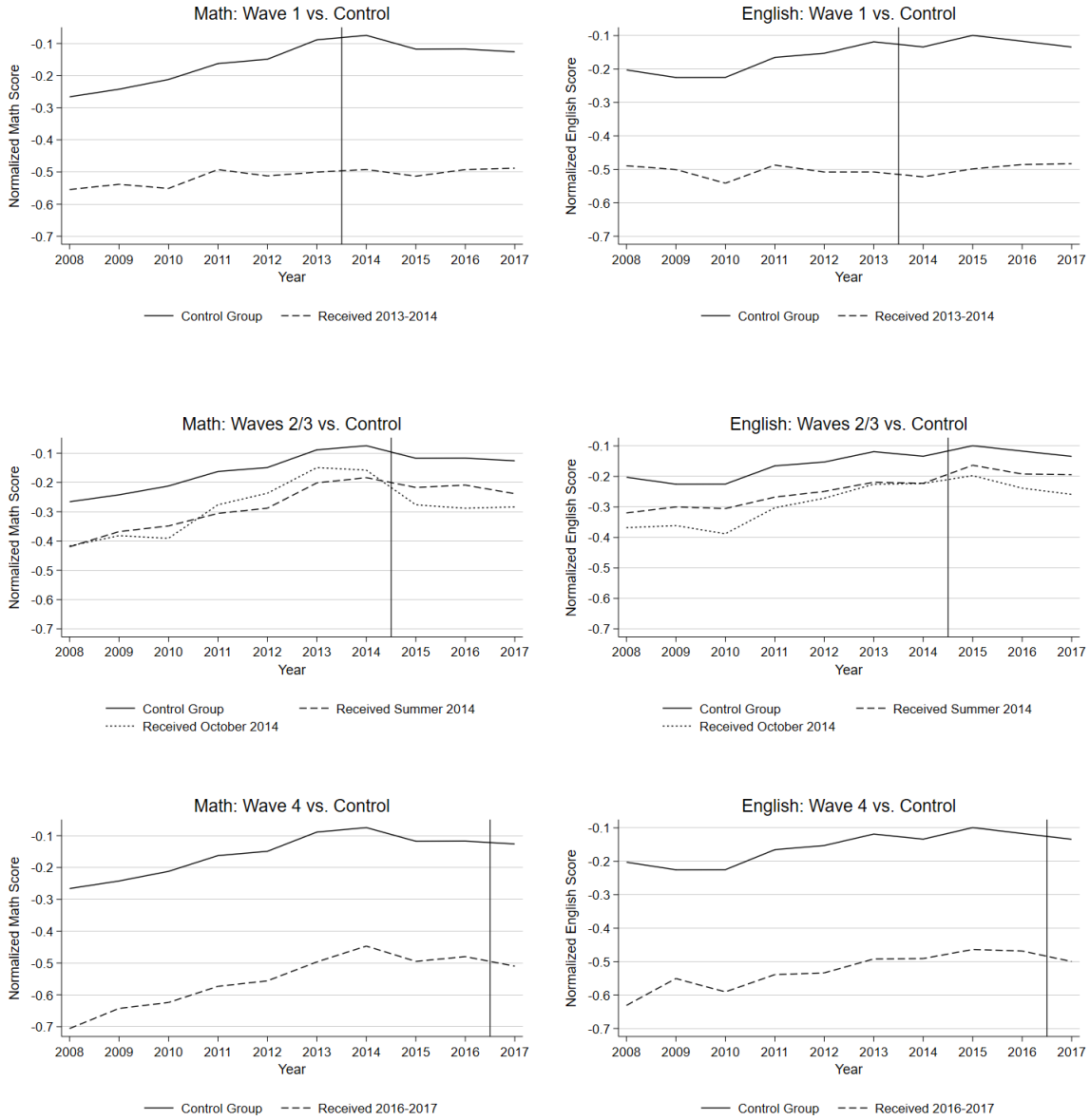


Figure 3.2: **Average Test Scores by Year and Wave**

Notes: The figure reports average annual test scores for math and English for students in treated and control schools for each year from 2008 to 2017, by wave of treatment in the CPS AC installation campaign. Test scores are standardized by year and grade level using the full Illinois distribution of test scores. The vertical line marks treatment year for all sub-figures.

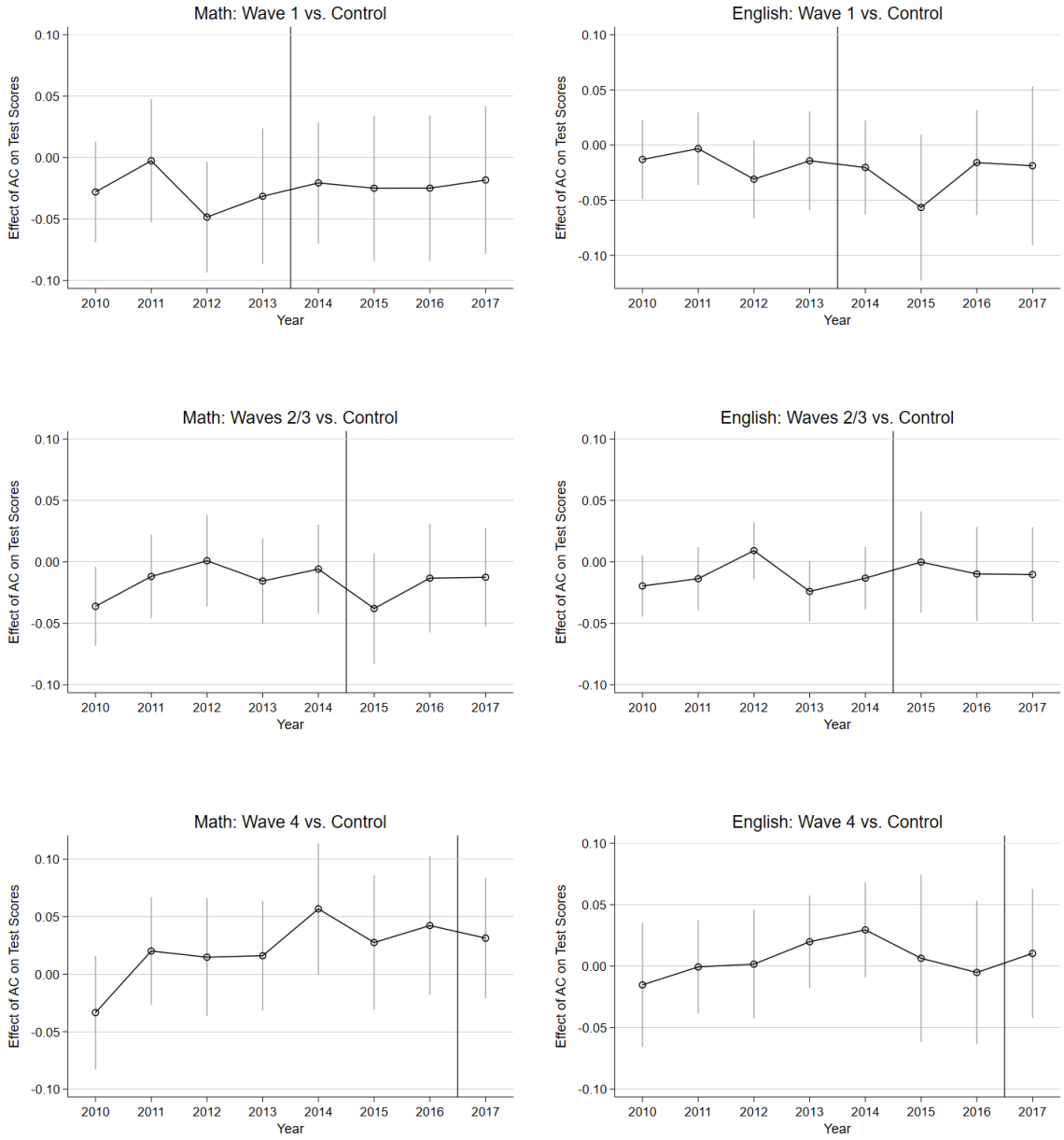


Figure 3.3: **Effects of AC on Test Scores by Wave of Treatment**

Notes: The figure reports difference-in-differences estimates for test score outcomes in math and English for treated and control schools from Equation 3.1 flexibly for each year from 2010 to 2017, by wave of treatment in the CPS AC installation campaign. Bars represent 95 percent confidence intervals. The vertical line marks the treatment year for all sub-figures. Test scores are standardized by year and grade level using the full Illinois distribution of test scores. Equation 3.1 includes year FE, school FE, and controls for the previous year's math and English scores. Robust standard errors are clustered at the school level.

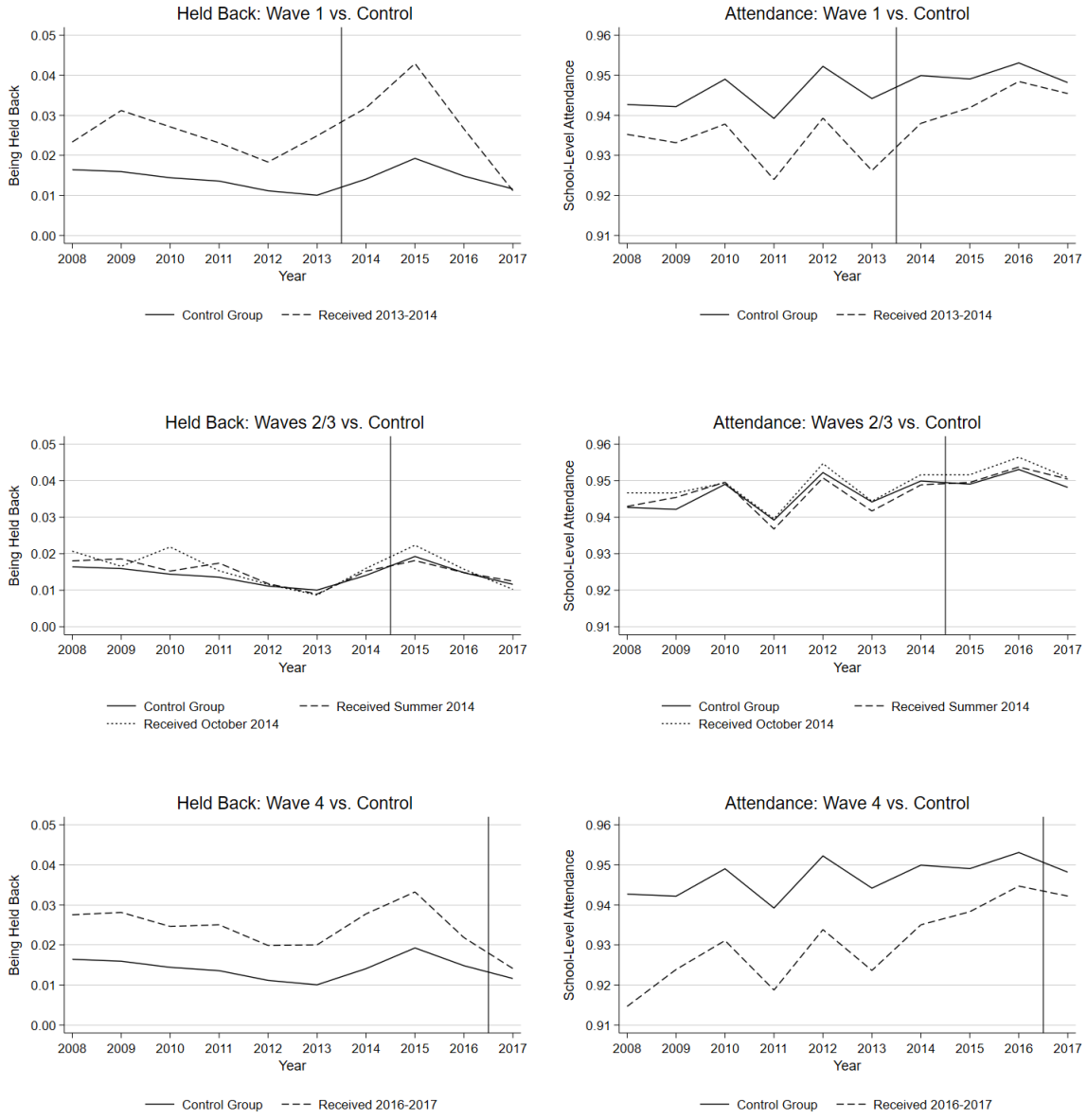


Figure 3.4: **Average Grade Retention and Attendance by Year and Wave**

Notes: The figure reports the average annual likelihood of being held back for students in treated and control schools on the left and average school-level attendance for treated and control schools on the right by wave of treatment in the CPS AC installation campaign. The vertical line marks treatment year for all sub-figures.

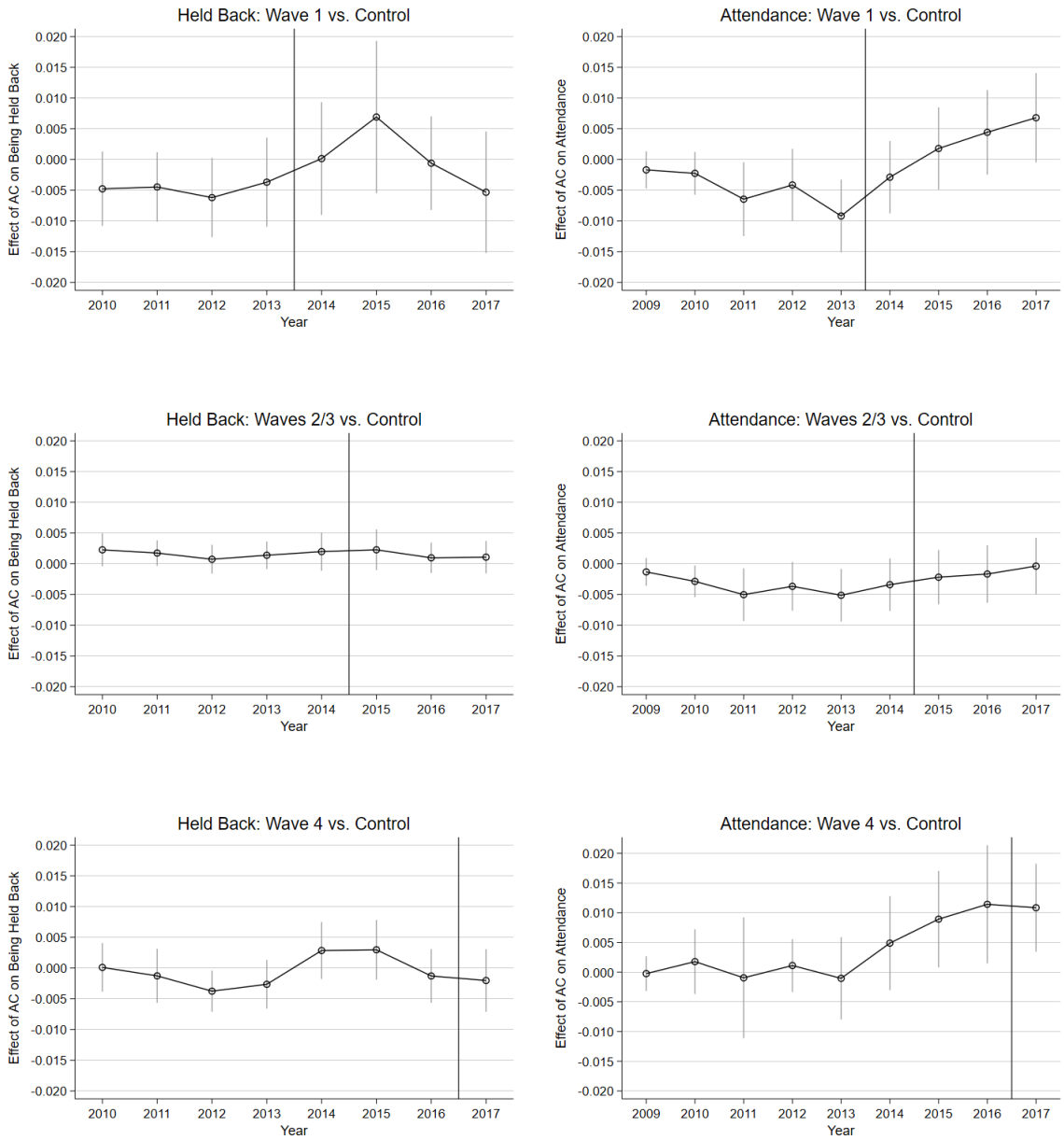


Figure 3.5: **Effects of AC on Grade Retention and Attendance by Wave**

Notes: The figure reports difference-in-differences estimates of the probability of being held back for students in treated and control schools from Equation 3.1 on the left and the difference-in-differences estimates of school-level average student attendance in treated and control schools on the right by the wave of treatment in the CPS AC installation campaign. Bars represent 95 percent confidence intervals. The vertical line marks the treatment year for all sub-figures. Equation 3.1 includes year FE, school FE, and controls for previous year's math and English test scores. However, lagged scores are not included for the attendance specifications since these data are at the school-year level. Robust standard errors are clustered at the school level.

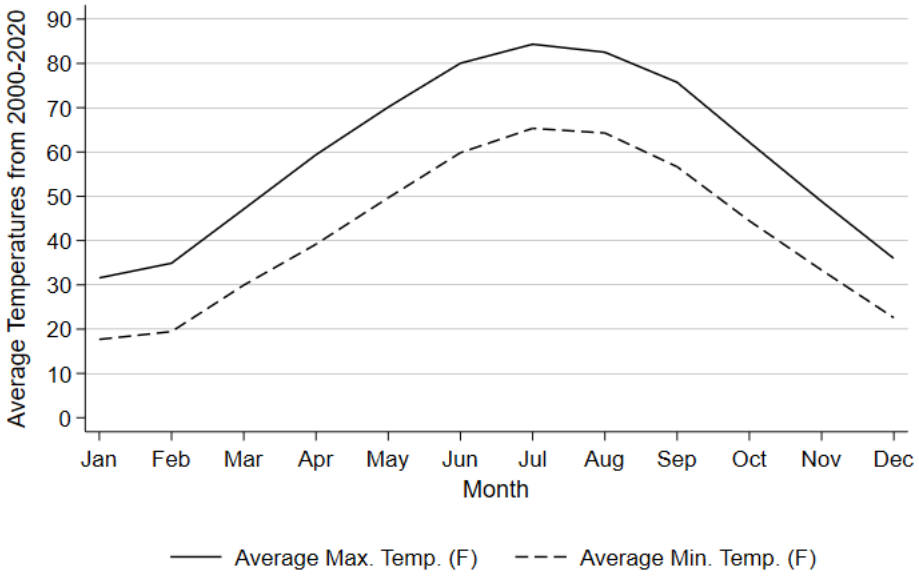


Figure 3.6: **Average Monthly Temperatures in Chicago (2000-2020)**

Notes: Average maximum and minimum temperatures each month in Chicago from 2000 to 2020 (Arguez et al., 2020).

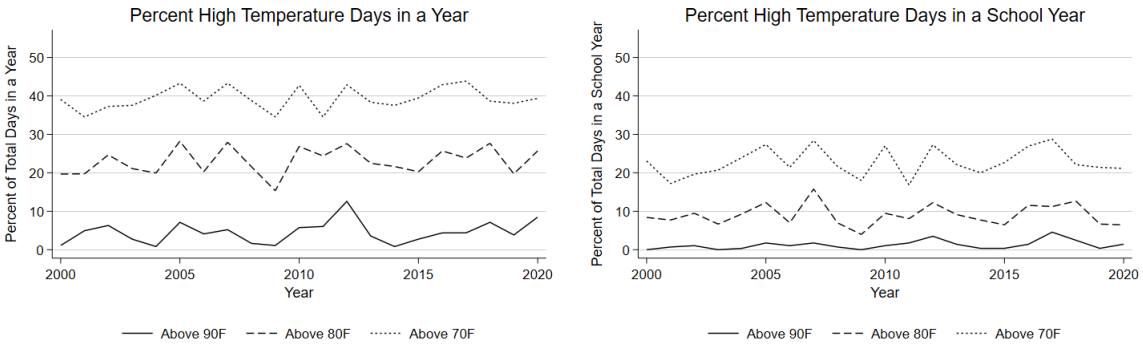


Figure 3.7: **High Temperature Days in Chicago**

Notes: The figure on the left plots the percent of days in each year from 2000 to 2020 that have a maximum temperature above 70F, 80F, and 90F. The figure on the right plots the percent of school days in each school year that have a maximum temperature above 70F, 80F, and 90F. Daily normals are reported from the Chicago O Hare NOAA Station (Arguez et al., 2020).

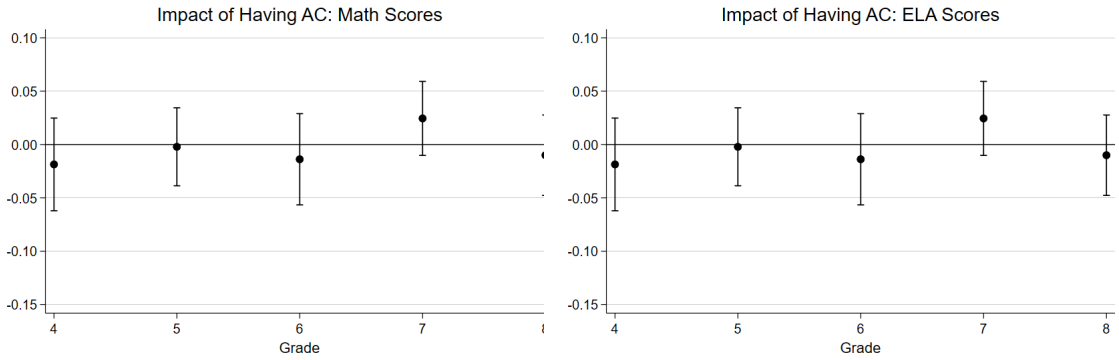


Figure 3.8: **Effect of AC on Test Scores by Grade**

Notes: The figures reports difference-in-differences estimates for math and English test scores for treated and control schools from Equation 3.1 by grade. Bars represent 95 percent confidence intervals. Estimations include year FE, school FE, and controls for previous year's math and English test scores. Robust standard errors are clustered at the school level.

Table 3.1: **School-Level Summary Statistics by Treatment Status**

	Control Mean (Std. Dev.)	Treated Mean (Std. Dev.)	Control-Treated Difference (T-Stat)
Panel A: Full Sample			
Math	-0.295	-0.405	0.110**
	(0.509)	(0.416)	(2.608)
English	-0.281	-0.395	0.114**
	(0.494)	(0.433)	(2.657)
Held Back	0.017	0.019	-0.002
	(0.051)	(0.015)	(-0.779)
Attendance	0.945	0.938	0.006*
	(0.027)	(0.029)	(2.093)
White	11.744	8.096	3.648*
	(18.758)	(15.124)	(2.077)
Black	39.310	54.694	-15.383***
	(41.089)	(42.951)	(-3.435)
Hispanic	43.482	32.496	10.986**
	(37.594)	(36.038)	(2.831)
Low Income	80.876	87.594	-6.718**
	(23.530)	(18.129)	(-3.122)
N	405	152	557
Panel B: Energy Star Sample			
Math	-0.164	-0.404	0.241***
	(0.536)	(0.426)	(4.674)
English	-0.168	-0.391	0.223***
	(0.535)	(0.446)	(4.241)
Held Back	0.013	0.019	-0.006***
	(0.013)	(0.014)	(-3.807)
Attendance	0.944	0.938	0.006
	(0.028)	(0.030)	(1.815)
White	11.811	8.498	3.314
	(18.624)	(15.408)	(1.797)
Black	39.088	53.578	-14.490**
	(40.935)	(42.975)	(-3.116)
Hispanic	43.571	33.003	10.568**
	(37.422)	(35.963)	(2.624)
Low Income	80.848	87.241	-6.392**
	(23.532)	(18.481)	(-2.816)
AC %	66.959	36.058	30.900***
	(36.514)	(31.257)	(8.481)
Year Built	1948	1931	17.255***
	(34.406)	(31.664)	(4.828)
Heated %	100.000	99.270	0.730
	(0.000)	(8.544)	(1.000)
N	217	137	354

Notes: Panel A contains information for the full sample of 557 schools. Panel B contains information on the 354 schools for which we have Energy Star data on AC penetration and other physical school characteristics.

Table 3.2: **Impact of AC From Difference-in-Differences**

	Math (1)	English (2)	Held Back (3)	Attendance (4)
Panel A: Full Sample				
Have AC	-0.0072	-0.0011	0.0010	0.0028**
	(0.0116)	(0.0108)	(0.0009)	(0.0013)
N	1,078,128	1,079,665	1,082,306	3,714
R^2	0.72	0.69	0.01	0.79
Panel B: Energy Star Sample				
Fraction AC	0.0041	0.0086	0.0002	0.0046**
	(0.0204)	(0.0164)	(0.0011)	(0.0020)
N	520,286	521,082	522,688	2,359
R^2	0.71	0.69	0.01	0.83

Notes: This table reports the estimated coefficients from the difference-in-differences model in Equation 3.1. The dependent variables in columns (1) and (2) are standardized math and English test scores, respectively. The dependent variable in col (3) is if a student is held back. The dependent variable in column (4) is average student attendance at the school-level, and therefore does not include lagged student test scores. In Panel A, *Have AC* is the main independent variable and is an indicator equal to one if a school has AC in a given year. In Panel B, *Fraction AC* is the main independent variable and is the fraction of the school that was air-conditioned prior to the AC installation campaign as reported in the 2011 Energy Star report. This variable is equal to 1 after a school receives AC. Robust standard errors clustered at the school level are in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table 3.3: **Schools that Received AC by Wave of Treatment**

Wave 1: 2013-2014	**Wave 2: Summer 2014**	**Wave 3: October 2014**	**Wave 4: Spring 2017**
Alex Haley ES (W)	Ambrose Plamondon ES	Albert R Sabin ES	Amundsen HS
Alice L Barnard ES	Arthur A Libby ES	Carl von Linne ES	Bennett ES
Benjamin E Mays ES (W)	Betty Shabazz – Sizemore	John G Whittier ES	Bogan HS
Bowen HS	Burnside ES	Alcott Humanities HS	Bouchet ES
Bret Harte ES	Charles H Wacker ES	Alexander Hamilton ES	Chicago Tech HS
Burnham Inclusive ES (W)	Charles Kozminski ES	Anna R. Langford ES	Clark G R ES
Carrie Jacobs Bond ES	Charles N Holden ES	Brighton Park ES	Cook ES
Charles Evans Hughes ES (W)	Christian Fenger HS	Cesar E Chavez ES	Cooper ES
Charles G Hammond ES	Christopher Columbus ES	Charles P Caldwell ES	Crown Fine Arts ES
Charles Sumner ES	Daniel Boone ES Daniel	Webster ES	Daniel Hale Williams HS
Clara Barton ES	Ella Flagg Young ES	DeWitt Clinton ES	Darwin ES
Daniel S Wentworth ES (W)	Fairfield ES	Edgebrook ES	Dunbar Vocational HS
Dewey ES of Fine Arts	Fernwood ES	Ernst Prussing ES	Epic Charter HS
Edmond Burke ES	Frank L Gillespie ES	Foster Park ES	Field ES
Ellen Mitchell ES	Frank W Gunsaulus ES	Frank W Reilly ES	Foreman HS
Esmond ES	Friedrich Ludwig Jahn ES	Franklin Art ES	Gage Park HS
Fort Dearborn ES	George B McClellan ES	Henry H Nash ES	Gale ES
Frederic Chopin ES (W)	George M Pullman ES	James Hedges ES	Graham A ES
Genevieve Melody ES (W)	Gurdon S Hubbard HS	Joseph Jungman ES	Harlan Community HS
George Leland ES (W)	Harold Washington ES	Joshua D Kershaw ES	Hirsch Metropolitan HS
George Manierre ES	Harriet Beecher Stowe ES	Mark Sheridan ES	Kelly HS
George W Curtis ES (W)	Helge A Haugan ES	Orville T Bright ES	Kilmer HS
George W Tilton ES (W)	Henry R Clissold ES	Phillip D Armour ES	King ES
George Washington Carver PS	Hiram H Belding ES	Richard J Oglesby ES	Lake View HS
Helen M Hefferan ES (W)	Inter-American Magnet ES	Rowe ES	Lasalle II ES
Ida B Wells Prep ES (W)	James N Thorp ES	Sauganash ES	Lincoln Park HS
Ira F Aldridge ES	James R Doolittle ES	Washington HS	Lovett ES
Irvin C Mollison ES (W)	Johann W Von Goethe ES	William E B Dubois ES	Lowell ES
Isabelle C O'Keeffe ES	John Barry ES	Wolfgang A Mozart ES	Madison ES
James B McPherson ES (W)	John Hay ES		Manley Career HS
James Otis ES (W)	Jonathan Burr ES		Mann ES

Notes: ES: Elementary School. MS: Middle School. HS: High School. (W) : 'Welcoming Schools' that were dropped from our main sample.

Wave 1: 2013-2014	Wave 2: Summer 2014	Wave 3: October 2014	Wave 4: Spring 2017
Jensen ES (W)	Kate S Kellogg ES		Marshall Metropolitan HS
Jesse Sherwood ES (W)	Louis Nettelhorst ES		Mason ES
John B Drake ES (W)	Lyman A Budlong ES		North Lawndale – Christiana
John Fiske ES (W)	Marvin Camras ES		Parkside ES
John Foster Dulles ES (W)	Melville W Fuller ES		Peace & Education HS
John Harvard ES (W)	Newton Bateman ES		Perez ES
John J Pershing ES Magnet (W)	Norman A Bridge ES		Perspectives Leadership HS
John M Smyth ES	North River ES		Perspectives Math Sci HS
John Milton Gregory ES (W)	Park Manor ES		Phillips Academy HS
Jose De Diego ES (W)	Patrick Henry ES		Phoenix Military HS
Laura S Ward ES (W)	Rachel Carson ES		Piccolo Specialty ES
Lawndale ES	Ravenswood ES		Richards Career HS
Leif Ericson ES	Spencer Technology ES		Roosevelt HS
Lorenz Brentano ES	Stephen Decatur ES		Ruggles ES
Ludwig Van Beethoven ES	Stephen K Hayt ES		Shoop Math Sci Tech ES
Mancel Talcott ES	Talman ES		Stagg ES
Maria Saucedo ES	Theodore Herzl ES		Suder Magnet ES
Mary E Courtenay ES (W)	Thomas A Hendricks ES		Sullivan HS
Michael Faraday ES (W)	Thomas J Waters ES		Tanner ES
Mount Vernon ES	Velma F Thomas Center		Tilden Career HS
Nicholson Tech Academy (W)	Washington D Smyser ES		Till Math Sci ES
Northwest MS	William Bishop Owen ES		Univ of Chicago – Donoghue
Owens Community ES (W)	William C Goudy ES		Univ of Chicago – Woodlawn
Paul Revere ES	William J Onahan ES		Urban Prep HS – West
Perkins Bass ES (W)	William Rainey Harper HS		Warren ES
Robert Nathaniel Dett ES (W)			Wells Community HS
Rosario Castellanos ES (W)			Whistler ES
Salmon P Chase ES			Woodson South ES
Scott Joplin ES			Yates ES
South Shore Academy (W)			
Thurgood Marshall MS			
Walter Q Gresham ES			
William C Reavis ES			
William H Ray ES			
William H Ryder ES (W)			
William W Carter ES			

Notes: ES: Elementary School. MS: Middle School. HS: High School. (W) : 'Welcoming Schools' that were dropped from our main sample.

Table 3.4: **Impact of AC From Difference-in-Differences for Wave 1 Schools, if Welcoming School or Not**

	Welcoming Schools				Not Welcoming Schools			
	Math (1)	English (2)	Held Back (3)	Attendance (4)	Math (5)	English (6)	Held Back (7)	Attendance (8)
Panel A: Full Sample								
Have AC	-0.0980***	-0.0692***	0.0046***	-0.0015	0.0004	-0.0143	0.0039*	0.0058**
	(0.0229)	(0.0207)	(0.0015)	(0.0030)	(0.0178)	(0.0182)	(0.0023)	(0.0027)
N	871,517	872,943	875,005	2,786	876,721	878,098	880,334	2,786
R^2	0.71	0.69	0.01	0.75	0.72	0.69	0.01	0.76
Panel B: Energy Star Sample								
Fraction AC	-0.1343***	-0.0802**	0.0068***	-0.0057*	0.0150	0.0060	0.0079**	0.0112**
	(0.0401)	(0.0340)	(0.0026)	(0.0031)	(0.0429)	(0.0371)	(0.0031)	(0.0049)
N	390,621	391,349	392,473	1,674	397,614	398,286	399,594	1,723
R^2	0.71	0.69	0.01	0.82	0.71	0.69	0.01	0.82

Notes: Panel A of the table reports the estimated coefficient on Have AC from the difference-in-differences outlined in Equation 3.1. Similarly, Panel B of the table reports the estimated coefficient on Fraction AC from the difference-in-differences model outlined in Equation 3.2. Columns (1)-(4) reports estimates using only the 33 *Welcoming Schools* treated in wave 1, while columns (5)-(8) reports estimates using only the 33 *Non-Welcoming Schools* in wave 1. Robust standard errors clustered at the school level are in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table 3.5: **Impact of AC From Difference-in-Differences Accounting for Time-Varying Heterogeneity in Impacts**

	Math (1)	English (2)	Held Back (3)	Attendance (4)
Have AC	-0.0037	-0.0004	0.0009	0.0030**
	(0.0118)	(0.0107)	(0.0009)	(0.0014)
N	1,077,554	1,079,096	1,081,732	3,720

Notes: This table reports the estimated coefficients from the difference-in-differences model in Equation 3.1 including lagged student test scores using the Borusyak et al. (2023) method to account for time-varying impacts due to difference in treatment timing across units. The dependent variables in columns (1) and (2) are standardized math and English test scores, respectively. The dependent variable in column (3) is if a student is held back. The dependent variable in column (4) is average student attendance at the school-level, and therefore does not include lagged student test scores. *Have AC* is the main independent variable and is an indicator equal to one if a school has AC in a given year. Wave 1 schools received AC in 2013-14, wave 2 in Summer 2014, wave 3 in October 2014, and wave 4 in 2016-17. Robust standard errors clustered at the school level are in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table 3.6: **Impact of AC From Difference-in-Differences by Wave of Treatment**

	Wave 1				Waves 2 & 3				Wave 4			
	Math (1)	English (2)	Held Back (3)	Attendance (4)	Math (5)	English (6)	Held Back (7)	Attendance (8)	Math (9)	English (10)	Held Back (11)	Attendance (12)
Panel A: Full Sample												
Have AC	-0.0007	-0.0156	0.0041*	0.0065**	-0.0101	0.0036	0.0001	0.0016	0.0149	0.0060	-0.0016	0.0080***
	(0.0178)	(0.0184)	(0.0023)	(0.0027)	(0.0154)	(0.0135)	(0.0008)	(0.0015)	(0.0241)	(0.0236)	(0.0019)	(0.0027)
N	848,371	849,685	851,787	2,656	958,072	959,540	961,618	3,096	843,445	844,854	846,741	2,634
R^2	0.72	0.69	0.01	0.77	0.72	0.69	0.01	0.76	0.72	0.69	0.01	0.81
Panel B: Energy Star Sample												
Fraction AC	0.0122	0.0033	0.0080***	0.0119**	0.0038	0.0063	-0.0013	0.0029	-0.0093	0.0250	-0.0039	0.0093**
	(0.0427)	(0.0371)	(0.0030)	(0.0049)	(0.0254)	(0.0195)	(0.0010)	(0.0021)	(0.0404)	(0.0348)	(0.0025)	(0.0046)
N	385,639	386,285	387,510	1,646	456,723	457,544	458,715	1,954	384,612	385,287	386,327	1,645
R^2	0.72	0.69	0.01	0.83	0.71	0.69	0.01	0.83	0.71	0.69	0.01	0.85

Notes: This table reports the estimated coefficients from the difference-in-differences model in Equation 3.1 separately by each wave of AC installation. In Panel A, *Have AC* is the main independent variable and is an indicator equal to one if a school has AC in a given year. In Panel B, *Fraction AC* is the main independent variable and is the fraction of the school that was air-conditioned prior to the AC installation campaign as reported in the 2011 Energy Star report. This variable is equal to 1 after a school receives AC. Wave 1 schools received AC in 2013-14, wave 2 in Summer 2014, wave 3 in October 2014, and wave 4 in 2016-17. Robust standard errors clustered at the school level are in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table 3.7: **Impact of Temperature on Test Scores**

	Days Above 70F		Days Above 80F		Days Above 90F	
	Math (1)	English (2)	Math (3)	English (4)	Math (5)	English (6)
Panel A: Using Lagged Test Scores						
School Days Above 70F	-0.0010	-0.0015**				
	(0.0006)	(0.0006)				
School Days Above 80F			-0.0000	-0.0013		
			(0.0009)	(0.0009)		
School Days Above 90F					-0.0011	-0.0017
					(0.0021)	(0.0015)
N	1,078,128	1,079,665	1,078,128	1,079,665	1,078,128	1,079,665
R^2	0.72	0.69	0.71	0.69	0.71	0.69
Panel B: Using Student Fixed Effects						
School Days Above 70F	0.0003	0.0001				
	(0.0010)	(0.0011)				
School Days Above 80F			0.0028	0.0019		
			(0.0017)	(0.0021)		
School Days Above 90F					0.0024	0.0017
					(0.0032)	(0.0032)
N	1,471,988	1,468,802	1,471,988	1,468,802	1,471,988	1,468,802
R^2	0.85	0.84	0.85	0.84	0.85	0.84

Notes: This table reports the estimated coefficients from a regression of achievement outcomes on the total number of school days above a certain temperature. The dependent variables in columns (1), (3) and (5) are the standardized math test scores, and in columns (2), (4) and (6) are the standardized English test scores. The main independent variable is the total number of days above a certain temperature in the school year. Controls in Panel A include lagged student test scores in the prior year and school fixed effects. Controls in Panel B include student fixed effects and school fixed effects. Errors are clustered by year. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Appendix A: Appendix to Chapter 1

A.1 Appendix Figures

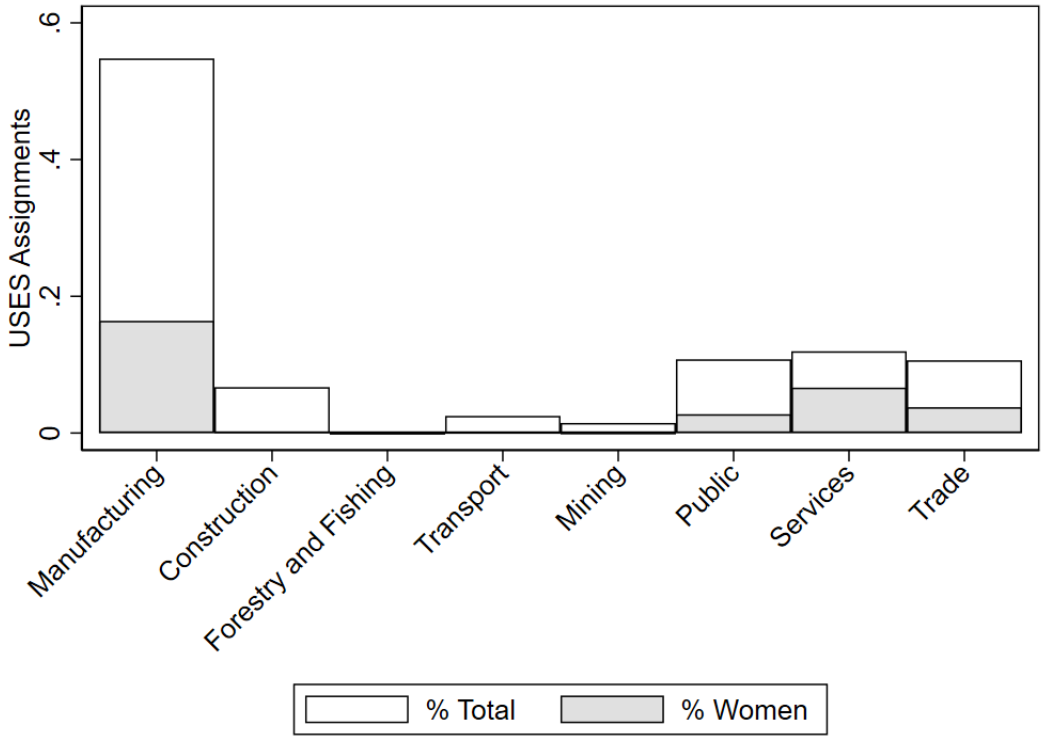


Figure A.1: **Distribution of USES Assignments across All Industries**

Notes: This figure represents the distribution of workers assigned to industries across all industry groups by the United States Employment Service between the third quarter of 1944 and the last quarter of 1946, as in Rose (2018), compiled from ”The Labor Market” reports of the War Manpower Commission’s Bureau of Program Planning and Review prior to 1945, and reports of the same name from the Labor Department since 1945. “% Total” bars represent the share of all USES assignments in the data that are in each industry group. ”% Women” bars in grey represent the share of all female USES assignments in the data that belong to each industry group.

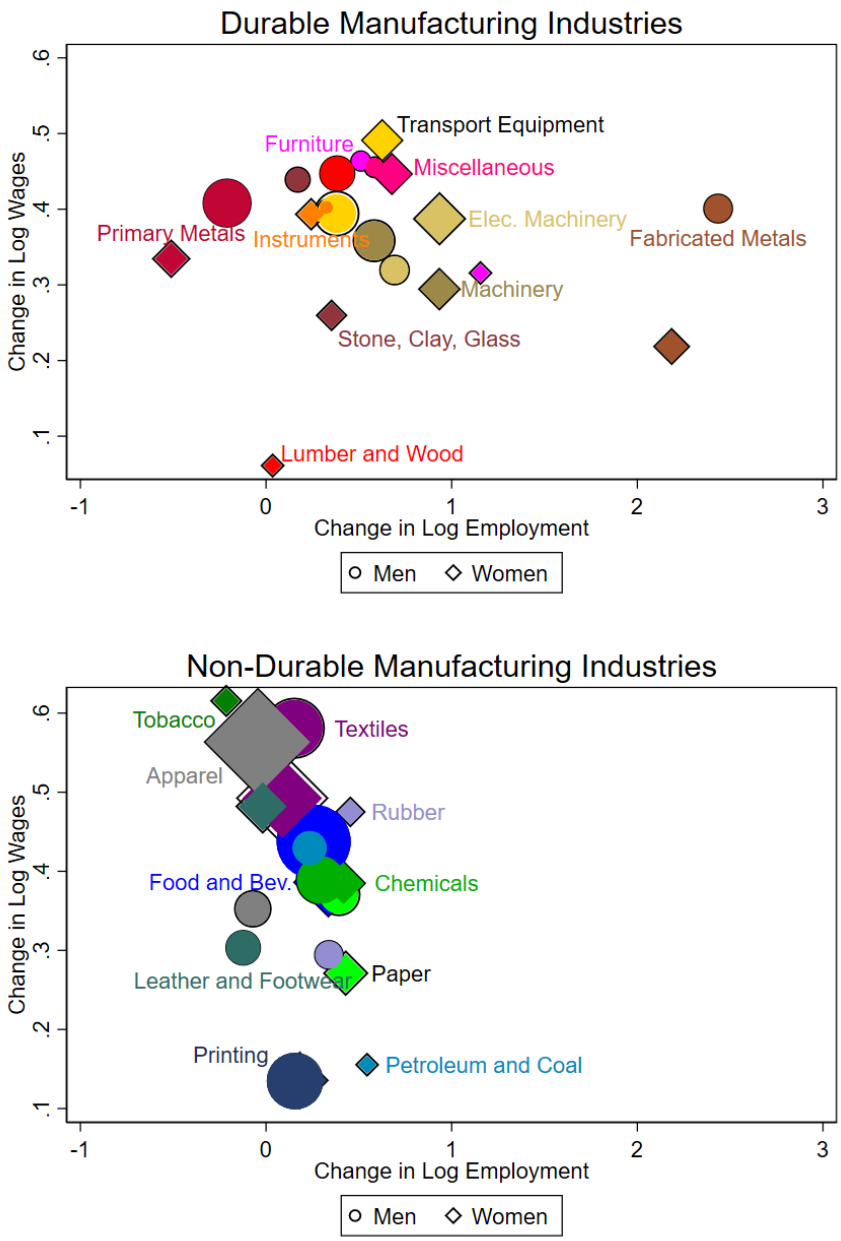


Figure A.2: **Changes in Relative Female-Male Wages and Employment**

Notes: This figure represents the differences in the log average annual wages in a manufacturing industry between 1940 and 1950 on the y-axis, plotted against the differences in the log total annual employment of women (by the diamonds) and men (by the circles) on the y-axis between 1940 and 1950, using the 1% Sample of the Decennial Censuses, excluding AK, HI, NV, DC, NM, MT and WY. Sizes of the diamonds and dots represent the number of workers of that gender in that industry across 1940 and 1950. Employment is the total number of full-time women (or men) employed in an industry, defined as those who work more than 40 weeks in a year and 35 hours per week in 1950 and those who work more than 40 weeks in 1940 (since all reported weeks are considered full-time equivalent in the 1940 Census). Wages in 1950 are inflation-adjusted to be comparable to 1940 $. The top figure with warm colors (red, orange, pink and brown) plots Durable Manufacturing Industries and the bottom figure with cool colors (blues, greens and grays) plots Non-Durable Manufacturing Industries.

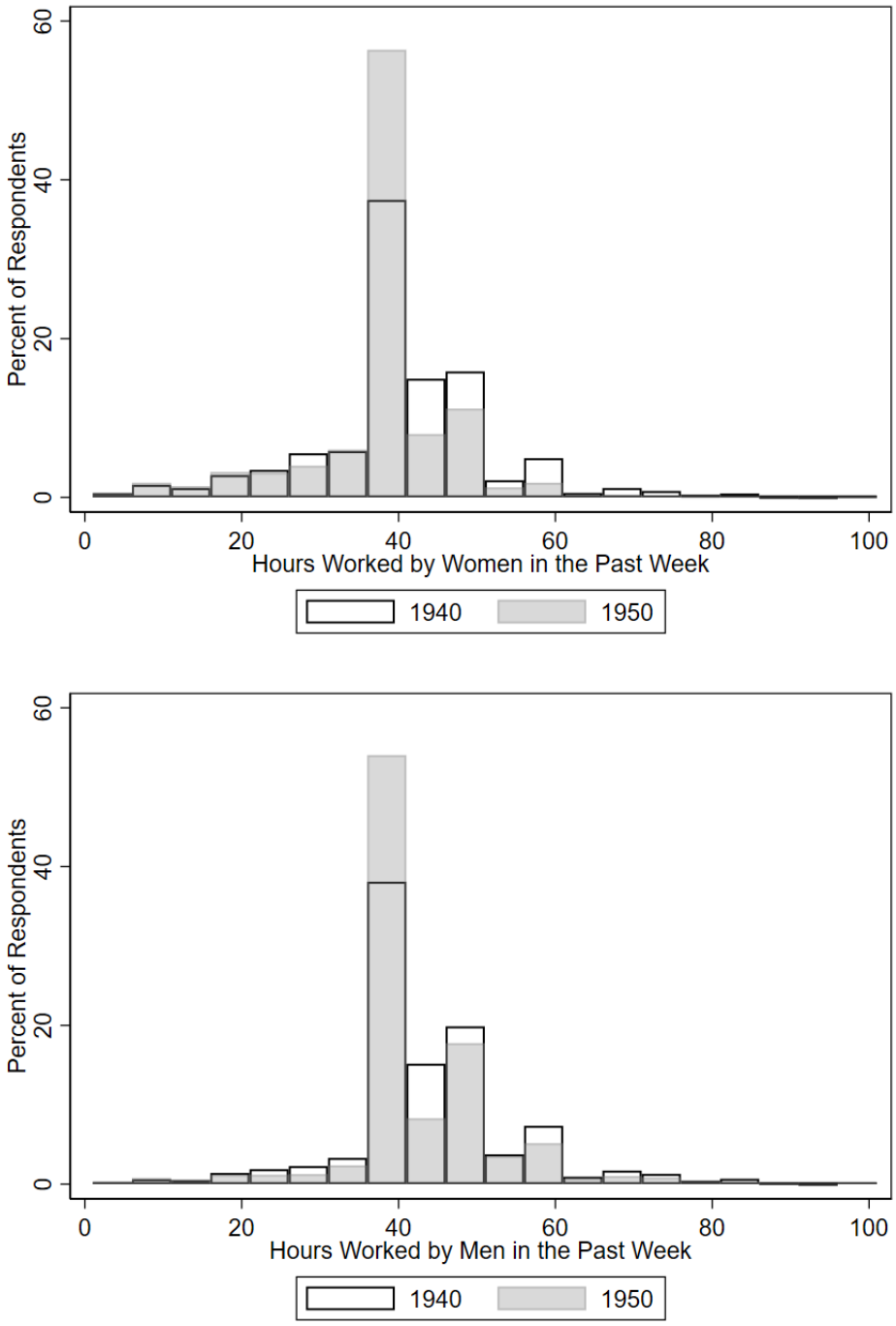


Figure A.3: **Distribution of 'Hours Worked Last Week' reported in the 1940 and 1950 Censuses**

Notes: This figure represents the distribution of "hours worked last week" reported by all workers in the 1940 and 1950 Decennial Censuses, separately for men and women. The sample includes responses from all individuals aged 15-64 who are US-born and excludes the states of AK, HI, NV, and DC.

A.2 Appendix Tables

Table A.1: **Impact of Industry-Level Relative Female Employment Shock on Wages, Including Industry Trends**

	Durables					Non-Durables				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Using Relative Demand Measure 1										
%Δ No. Women by Industry $\times d_{1950} \times f_i$	0.277***	0.222**	0.218**	0.276***	0.234**	-0.030	0.276*	-0.008	0.104	0.511
	(0.093)	(0.102)	(0.091)	(0.092)	(0.100)	(0.136)	(0.167)	(0.142)	(0.139)	(0.331)
Allowing for different time trends for women by proportion of industry:										
With High School Education in 1940		0.693			-0.471		-0.835***			-1.359
		(0.787)			(1.122)		(0.274)			(1.440)
Non-White in 1940			-1.454*		-1.940			0.250		1.110*
			(0.838)		(1.187)			(0.612)		(0.575)
Women in Industry Married in 1940				3.486	7.864				28.430***	-12.432
				(25.833)	(26.154)				(10.095)	(52.518)
N	26,676	26,676	26,676	26,676	26,676	25,990	25,990	25,990	25,990	25,990
R^2	0.36	0.36	0.36	0.36	0.36	0.43	0.43	0.43	0.43	0.43
Panel B: Using Relative Demand Measure 2										
%Δ Female Prop. of Industry $\times d_{1950} \times f_i$	4.178***	3.423**	3.577**	4.992***	4.159***	2.586	4.270	2.212	2.880	1.242
	(1.439)	(1.437)	(1.419)	(1.578)	(1.564)	(5.092)	(4.900)	(5.209)	(4.918)	(5.352)
Allowing for different time trends for women by proportion of industry:										
With High School Education in 1940		1.176*			-0.072		-0.578***			0.247
		(0.713)			(1.058)		(0.208)			(0.713)
Non-White in 1940			-1.900**		-1.837			0.221		0.622
			(0.830)		(1.181)			(0.607)		(0.628)
Women in Industry Married in 1940				-36.509	-23.134				26.083***	37.671
				(28.111)	(27.997)				(9.337)	(32.199)
N	26,676	26,676	26,676	26,676	26,676	25,990	25,990	25,990	25,990	25,990
R^2	0.36	0.36	0.36	0.36	0.36	0.43	0.43	0.43	0.43	0.43

Notes: Results reported from estimating Equation 1.19 in Section 1.5. Measures 1 and 2 are detailed in Section 1.5. The main dependent variable is the log individual annual wages from a pooled 1% sample of U.S.-born full-time workers aged 15-64 from the 1940 and 1950 US Decennial Censuses. All specifications include gender-by-state-by-year, gender-by-industry, and industry-by-year fixed effects, as well as the log levels of relative female-male employment in their industry-state-year and its interaction with the female dummy, instrumented with log levels of relative female-male employment in an industry-state-year in 1930 and its interaction with the female dummy. Individual covariates include being married, number of children under age 5 in the household, 12 or more years of education, race, and a 4th-degree polynomial in potential experience. Robust standard errors are clustered by state and industry, and sample weights are used. Final sample excludes AK, HI, DC, MT, NV, NM, and WY.

Table A.2: **Impact of WWII Employment Shocks on Wages in Durable Manufacturing 1940-1950, With and Without Employment**

	No Emp	W/ Emp	No Emp	W/ Emp	No Emp	W/ Emp
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Using Relative Demand Measure 1						
%Δ No. Women by Industry$\times d_{1950} \times f_i$	0.180***	0.176***	0.184***	0.191***	0.199***	0.277***
	(0.050)	(0.059)	(0.052)	(0.064)	(0.065)	(0.093)
%Δ Total Demand by Industry$\times d_{1950}$	0.295	0.059	0.612*	0.480		
	(0.416)	(0.483)	(0.365)	(0.434)		
%Δ No. Women by State$\times d_{1950} \times f_i$	-0.109	-0.094				
	(0.082)	(0.085)				
%Δ Total Demand by State$\times d_{1950}$	0.000	-0.002				
	(0.002)	(0.003)				
N	26,676	26,676	26,676	26,676	26,676	26,676
R^2	0.37	0.36	0.37	0.36	0.37	0.36
Panel B: Using Relative Demand Measure 2						
%Δ Female Prop. of Industry$\times d_{1950} \times f_i$	3.115***	3.174***	3.204***	3.395***	3.216***	4.178***
	(0.800)	(0.989)	(0.847)	(1.040)	(1.021)	(1.439)
%Δ Total Demand by Industry$\times d_{1950}$	0.151	-0.078	0.460	0.323		
	(0.398)	(0.469)	(0.345)	(0.412)		
%Δ Female Prop. of State $\times d_{1950} \times f_i$	-0.130	-0.113				
	(0.086)	(0.090)				
%Δ Total Demand by State$\times d_{1950}$	0.000	-0.002				
	(0.002)	(0.003)				
N	26,676	26,676	26,676	26,676	26,676	26,676
R^2	0.37	0.36	0.37	0.36	0.37	0.36
Gender-by-State-by-Year FE			✓	✓	✓	✓
Industry-by-Year FE					✓	✓
Individual-Level Controls	✓	✓	✓	✓	✓	✓

Notes: Results reported from estimating Equation 1.17 in columns (1) and (2), Equation 1.15 in columns (3) and (4), and Equation 1.19 in columns (5) and (6) for Durable Manufacturing, with and without including employment. The main dependent variable is the log individual annual wages from a pooled 1% sample of U.S.-born full-time workers aged 15-64 from the 1940 and 1950 US Decennial Censuses. 'Measure 1' is the percent change in the number of women relative to the percent change in all workers during WWII. 'Measure 2' is the percent change in the relative proportion of women in an industry (or state) during WWII defined in Section 1.5. All columns include gender-by-year, gender-by-state, and gender-by-industry fixed effects. Employment is measured as the log of relative female-male employment in the industry-state-year and its interaction with the female dummy, instrumented with the log levels of relative female-male employment in an industry-state-year in 1930 and its interaction with the female dummy in columns (2), (4), and (6). Columns (2) and (5) include gender-by-state-by-year fixed effects, and columns (3) and (6) include those as well as industry-by-year fixed effects. Individual covariates included are being married, number of children under age 5 in the household, 12 or more years of education, race, and a 4th-degree polynomial in potential experience. Robust standard errors are clustered by state and industry, and sample weights are used. Final sample excludes AK, HI, DC, MT, NV, NM, and WY.

Table A.3: **Impact of WWII Employment Shocks on Wages in Non-Durable Manufacturing 1940-1950, With and Without Employment**

	No Emp	W/ Emp	No Emp	W/ Emp	No Emp	W/ Emp
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Using Relative Demand Measure 1						
$\%\Delta$ No. Women. by Industry$\times d_{1950} \times f_i$	-0.257**	-0.180	-0.312**	-0.262**	-0.053	-0.030
	(0.119)	(0.130)	(0.121)	(0.131)	(0.131)	(0.136)
$\%\Delta$ Total Demand by Industry$\times d_{1950}$	0.033	0.085	0.221	0.218		
	(0.718)	(0.716)	(0.529)	(0.536)		
$\%\Delta$ No. Women by State$\times d_{1950} \times f_i$	0.023	0.045				
	(0.059)	(0.063)				
$\%\Delta$ Total Demand by State$\times d_{1950}$	0.000	0.000				
	(0.000)	(0.001)				
N	25,984	25,984	25,990	25,990	25,990	25,990
R^2	0.43	0.43	0.43	0.43	0.43	0.43
Panel B: Using Relative Demand Measure 2						
$\%\Delta$ Female Prop. of Industry$\times d_{1950} \times f_i$	-3.932	-1.274	-5.318	-3.138	1.743	2.586
	(4.759)	(5.112)	(4.883)	(5.141)	(4.978)	(5.092)
$\%\Delta$ Total Demand by Industry$\times d_{1950}$	0.438	0.307	0.756	0.611		
	(0.761)	(0.759)	(0.576)	(0.576)		
$\%\Delta$ Female Prop. of State $\times d_{1950} \times f_i$	0.020	0.044				
	(0.064)	(0.068)				
$\%\Delta$ Total Demand by State$\times d_{1950}$	0.000	0.000				
	(0.000)	(0.001)				
N	25,984	25,984	25,990	25,990	25,990	25,990
R^2	0.43	0.43	0.43	0.43	0.43	0.43
Gender-by-State-by-Year FE			✓	✓	✓	✓
Industry-by-Year FE					✓	✓
Individual-Level Controls	✓	✓	✓	✓	✓	✓

Notes: Results reported from estimating Equation 1.17 in columns (1) and (2), Equation 1.15 in columns (3) and (4), and Equation 1.19 in columns (5) and (6) for Non-Durable Manufacturing, with and without including employment. The main dependent variable is the log individual annual wages from a pooled 1% sample of U.S.-born full-time workers aged 15-64 from the 1940 and 1950 US Decennial Censuses. 'Measure 1' is the percent change in the number of women relative to the percent change in all workers during WWII. 'Measure 2' is the percent change in the relative proportion of women in an industry (or state) during WWII defined in Section 1.5. All columns include gender-by-year, gender-by-state, and gender-by-industry fixed effects. Employment is measured as the log of relative female-male employment in the industry-state-year and its interaction with the female dummy, instrumented with the log levels of relative female-male employment in an industry-state-year in 1930 and its interaction with the female dummy in columns (2), (4), and (6). Columns (2) and (5) include gender-by-state-by-year fixed effects, and columns (3) and (6) include those as well as industry-by-year fixed effects. Individual covariates included are being married, number of children under age 5 in the household, 12 or more years of education, race, and a 4th-degree polynomial in potential experience. Robust standard errors are clustered by state and industry, and sample weights are used. Final sample excludes AK, HI, DC, MT, NV, NM, and WY.

Table A.4: **Impact of Industry-Level and State-Level WWII Employment Shocks on Hourly Wages in Manufacturing 1940-1950, Using Measure 1**

	Durables			Non-Durables		
	(1)	(2)	(3)	(4)	(5)	(6)
$\%\Delta$ No. Women by Industry$\times d_{1950} \times f_i$	-0.006	-0.023	0.109*	-0.262***	-0.281***	-0.148
	(0.040)	(0.038)	(0.057)	(0.072)	(0.072)	(0.091)
$\%\Delta$ Total Demand by Industry$\times d_{1950}$	-0.797*	-0.483		-1.157**	-1.028**	
	(0.419)	(0.314)		(0.565)	(0.440)	
$\%\Delta$ No. Women by State$\times d_{1950} \times f_i$	-0.006			0.078*		
	(0.047)			(0.043)		
$\%\Delta$ Total Demand by State$\times d_{1950}$	0.001			0.000		
	(0.002)			(0.000)		
$ln(\frac{F_{sjt}}{M_{sjt}})$	0.078	0.080	0.090	-0.026	-0.025	-0.022
	(0.053)	(0.057)	(0.060)	(0.033)	(0.033)	(0.033)
$ln(\frac{F_{sjt}}{M_{sjt}}) \times f_i$	-0.067	-0.066	-0.073	-0.064*	-0.062*	-0.063*
	(0.067)	(0.070)	(0.072)	(0.037)	(0.037)	(0.037)
d_{1950}	0.111**	0.127***	0.062	0.248***	0.234***	0.174***
	(0.048)	(0.046)	(0.059)	(0.054)	(0.056)	(0.056)
f_i	-0.532***	-0.476**	-0.499**	-0.403***	-0.391***	-0.379***
	(0.184)	(0.190)	(0.199)	(0.078)	(0.077)	(0.077)
$f_i \times d_{1950}$	0.227*	0.172*	-0.055	0.258**	0.365***	0.166
	(0.118)	(0.102)	(0.113)	(0.130)	(0.111)	(0.132)
N	27,329	27,329	27,329	26,265	26,281	26,281
R^2	0.32	0.32	0.32	0.36	0.36	0.36

Notes: Results reported from estimating Equation 1.17 in columns (1) and (4), Equation 1.15 in columns (2) and (5), and Equation 1.19 in columns (3) and (6) for Durable and Non-Durable Manufacturing industries respectively. The main dependent variable is the log hourly individual wage from a pooled 1% sample of U.S.-born full-time workers aged 15-64 from the 1940 and 1950 US Decennial Censuses. 'Measure 1' is the percent change in the number of women relative to the percent change in all workers during WWII defined in Section 1.5. All columns include gender-by-year, gender-by-state, and gender-by-industry fixed effects, as well as the log of relative female-male employment (in average weeks) in the industry-state-year and its interaction with the female dummy, instrumented with the log levels of relative female-male employment in an industry-state-year in 1930 (in workers) and its interaction with the female dummy. Columns (2) and (5) include gender-by-state-by-year fixed effects, and columns (3) and (6) include those as well as industry-by-year fixed effects. Individual covariates included are being married, number of children under age 5 in the household, 12 or more years of education, race, and a 4th-degree polynomial in potential experience. Robust standard errors are clustered by state and industry, and sample weights are used. Final sample excludes AK, HI, DC, MT, NV, NM, and WY.

Table A.5: **Impact of Industry-Level and State-Level WWII Employment Shocks on Hourly Wages in Manufacturing 1940-1950, Using Measure 2**

	Durables			Non-Durables		
	(1)	(2)	(3)	(4)	(5)	(6)
%Δ Female Prop. of Industry$\times d_{1950} \times f_i$	0.228	0.034	1.551*	1.039	1.605	8.616***
	(0.622)	(0.624)	(0.814)	(2.935)	(2.954)	(3.312)
%Δ Total Demand by Industry$\times d_{1950}$	-0.789*	-0.457		-0.977	-0.833*	
	(0.417)	(0.308)		(0.608)	(0.487)	
%Δ Female Prop. of State$\times d_{1950} \times f_i$	-0.011			0.077*		
	(0.050)			(0.046)		
%Δ Total Demand by State$\times d_{1950}$	0.001			0.000		
	(0.002)			(0.000)		
$ln(\frac{F_{sjt}}{M_{sjt}})$	0.078	0.080	0.091	-0.026	-0.025	-0.021
	(0.053)	(0.057)	(0.060)	(0.033)	(0.033)	(0.033)
$ln(\frac{F_{sjt}}{M_{sjt}}) \times f_i$	-0.068	-0.067	-0.074	-0.065*	-0.064*	-0.066*
	(0.067)	(0.070)	(0.072)	(0.038)	(0.037)	(0.037)
d_{1950}	0.110**	0.126***	0.065	0.244***	0.229***	0.169***
	(0.048)	(0.046)	(0.058)	(0.055)	(0.057)	(0.056)
f_i	-0.535***	-0.484**	-0.475**	-0.408***	-0.406***	-0.389***
	(0.183)	(0.189)	(0.191)	(0.079)	(0.077)	(0.078)
$f_i \times d_{1950}$	0.207***	0.135	0.085	-0.013	-0.000	-0.081
	(0.076)	(0.086)	(0.083)	(0.074)	(0.071)	(0.075)
N	27,329	27,329	27,329	26,265	26,281	26,281
R^2	0.32	0.32	0.32	0.36	0.36	0.36

Notes: Results reported from estimating Equation 1.17 in columns (1) and (4), Equation 1.15 in columns (2) and (5), and Equation 1.19 in columns (3) and (6) for Durable and Non-Durable Manufacturing industries respectively. The main dependent variable is the log hourly individual wage from a pooled 1% sample of U.S.-born full-time workers aged 15-64 from the 1940 and 1950 US Decennial Censuses. 'Measure 2' is the percent change in the relative proportion of women in an industry (or state) during WWII defined in Section 1.5. All columns include gender-by-year, gender-by-state, and gender-by-industry fixed effects, as well as the log of relative female-male employment (in average weeks) in the industry-state-year and its interaction with the female dummy, instrumented with the log levels of relative female-male employment in an industry-state-year in 1930 (in workers) and its interaction with the female dummy. Columns (2) and (5) include gender-by-state-by-year fixed effects, and columns (3) and (6) include those as well as industry-by-year fixed effects. Individual covariates included are being married, number of children under age 5 in the household, 12 or more years of education, race, and a 4th-degree polynomial in potential experience. Robust standard errors are clustered by state and industry, and sample weights are used. Final sample excludes AK, HI, DC, MT, NV, NM, and WY.

Appendix B: Appendix to Chapter 2

B.1 Appendix Figures

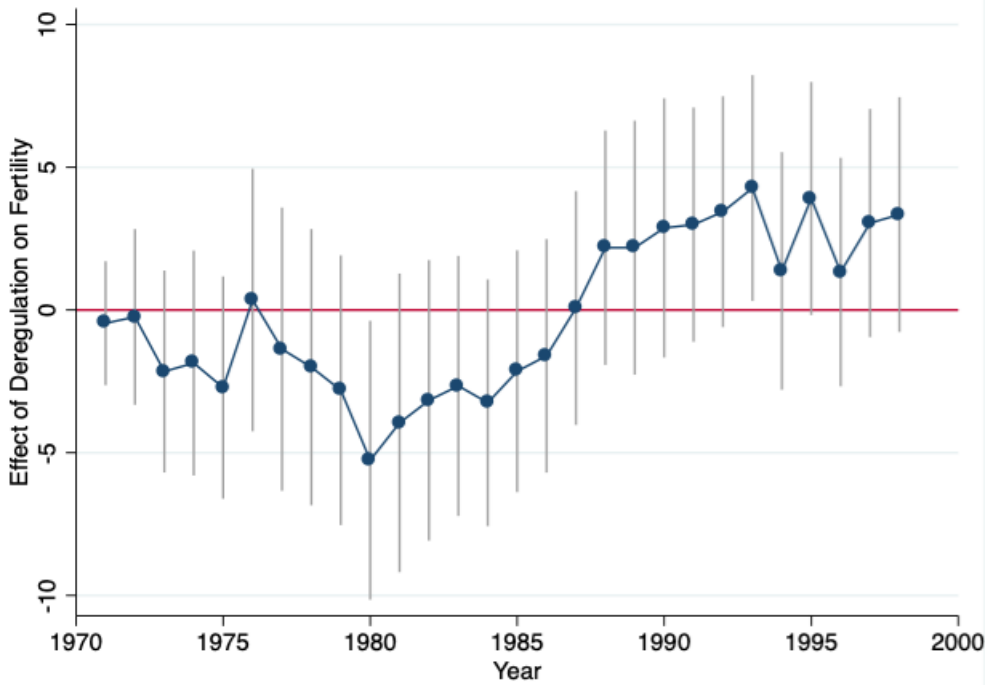


Figure B.1: **Impacts of Bank Reform on State-Level Fertility, by Year**
Notes: We show the coefficients of the βs from our individual difference-in-differences equation as in 2.2, but separately for each year. All specifications include state and year fixed effects, and state controls for the log of the population, the fraction of the population white, and the fraction of the population employed. Standard errors are clustered at the state-year level.

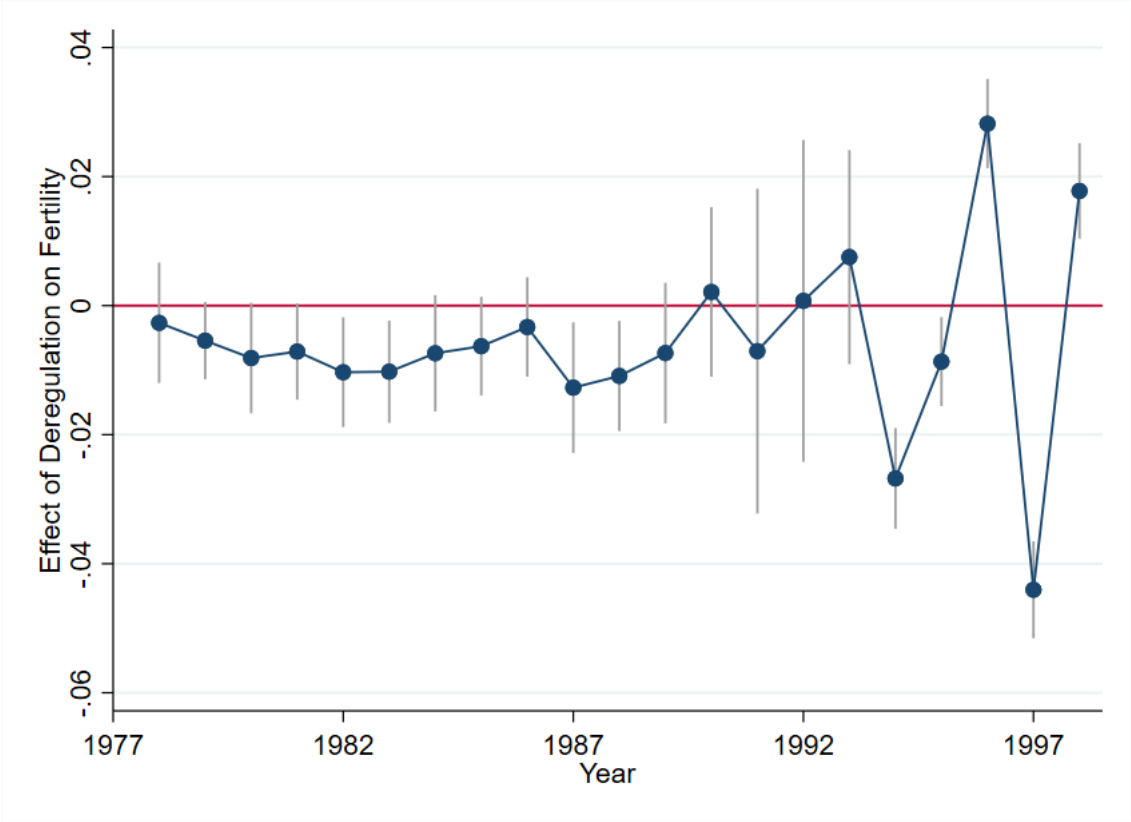


Figure B.2: **Impacts of Bank Reform on Individual-Level Fertility, by Year**

Notes: We show the coefficients of the βs from our individual difference-in-differences equation as in 2.3, but separately for each year. All specifications include state and year fixed effects, state controls for the log of the population, the fraction of the population white, and the fraction of the population employed. Individual controls include a 4th-degree polynomial in age, race, and high school or college education. Robust standard errors are clustered at the state-year level.
Sample: Women aged 20-44 in the CPS March Supplement 1977-2000.

B.2 Appendix Tables

Table B.1: **Weibull Hazard Model to Predict Time to Bank Reform, State-Level**

	(1)	(2)	(3)	(4)
Fertility Rate	-0.0003	0.0025	0.0001	-0.0022
	(0.0075)	(0.0144)	(0.0070)	(0.0049)
Small Bank Share of Assets			4.0658**	7.5221***
			(1.7945)	(2.7403)
Capital Ratio of Small to Large Banks			9.6077*	4.9604
			(5.0632)	(3.8462)
Bank Sells Insurance			-3.7695***	-1.2793
			(1.3893)	(0.9441)
Relative size of Insurance Inc (Sell)			9.8252**	4.1577
			(4.2052)	(3.3043)
Relative size of Insurance Inc (Cannot Sell)			-0.8890*	-0.0760
			(0.5278)	(0.7898)
Small Firm Share			-11.0676**	-12.8578***
			(5.1106)	(3.5232)
Unit Banking			0.3097***	0.1592*
			(0.1012)	(0.0813)
If State Changes Bank Insurance Law			0.0762	0.2444***
			(0.0749)	(0.0810)
N	593	593	593	593
State-Level Contemporaneous Controls		✓	✓	✓
State Banking Sector Characteristics			✓	✓
Region Dummies				✓

Notes: Coefficients from estimating a Weibull Hazard equation of the time to deregulation on the fertility rate. Robust standard errors are clustered at the state-level in parentheses. The fertility rate is defined as the number of births per 1000 women aged 20-44 in state s in year t from Bailey et al. (2016). State-Level controls in columns (3) and (4) include the log of the population, the fraction of the population employed, and the fraction of the population white from Bureau of Economic Analysis (2020). State banking sector characteristics are from Kroszner and Strahan (1999) The truncated final sample covers 1970-1988. We do not include the states of SD, DE, AK, and HI.

Table B.2: **Weibull Hazard Model to Predict Time to Bank Reform, Individual-Level**

	(1)	(2)	(3)	(4)	(5)
Gave Birth in Last Year	0.0030	0.0026	-0.0013	-0.0049	-0.0049
	(0.0081)	(0.0079)	(0.0069)	(0.0061)	(0.0058)
Small Bank Share of Assets			2.1567**	3.8043***	3.8042***
			(0.8460)	(0.9543)	(0.9542)
Capital Ratio of Small to Large Banks			1.4576	1.6371	1.6308
			(2.7223)	(2.4160)	(2.4146)
Bank Sells Insurance			-1.9429**	-1.0148**	-1.0160**
			(0.8638)	(0.5092)	(0.5099)
Relative size of Insurance Inc (Sell)			4.3192*	2.1127	2.1168
			(2.2782)	(1.4218)	(1.4242)
Relative size of Insurance Inc (Cannot Sell)			-0.8878***	-0.6823***	-0.6825***
			(0.2337)	(0.2552)	(0.2550)
Small Firm Share			-2.2292	-6.5893**	-6.5848**
			(7.1953)	(2.6087)	(2.6050)
Unit Banking			0.2050	0.2068**	0.2065**
			(0.1406)	(0.0867)	(0.0866)
If State Changes Bank Insurance Law			0.0198	0.1024	0.1024
			(0.0496)	(0.0871)	(0.0870)
N	146089	146089	146089	146089	146089
State-Level Contemporaneous Controls		✓	✓	✓	✓
State Banking Sector Characteristics			✓	✓	✓
Region Dummies				✓	✓
Individual Controls					✓

Notes: Coefficients from estimating a Weibull Hazard Model to the timing of bank deregulation on the likelihood of giving birth in the past year. Robust standard errors are clustered at the state-year level in parentheses. State-Level Controls in columns (4) and (5) include the log of the population, the fraction of the population white, and the fraction of the population employed in a state from Bureau of Economic Analysis (2020). Individual Controls in column (5) include race, high school or college education, and a 4th-degree polynomial in age. The sample covers 1977-1988. State banking sector characteristics are from Kroszner and Strahan (1999). Sample: Women aged 20-44 in the CPS March Supplement, not including the states of SD, DE, AK and HI.

Table B.3: **Impact of Bank Reform on State-Level Fertility using both Classic TWFE and Borusyak et al. (2023) TWFE**

	Classic TWFE		Robust TWFE	
	(1)	(2)	(3)	(4)
Panel A: Full Sample				
Bank Reform	-6.3577***	-3.6512**	-6.3577***	-1.3632*
	(2.3231)	(1.3823)	(2.4250)	(0.7040)
Mean Dependent Var.	71.34	71.34	71.34	71.34
SD Dependent Var.	(14.63)	(14.63)	(14.63)	(14.63)
N	1,457	1,457	1,457	1,073
R^2	0.05	0.89		
Panel B: 1970-1988				
Bank Reform	-3.8248	-3.5307**	-3.8248	-4.0698***
	(3.5405)	(1.5625)	(3.4257)	(0.6935)
Mean Dependent Var.	74.09	74.09	74.09	74.09
SD Dependent Var.	(15.53)	(15.53)	(15.53)	(15.53)
N	893	893	893	703
R^2	0.01	0.91		
Panel C: 1989-2000				
Bank Reform	-0.0734	0.7683	-0.0734	-0.2542
	(3.2411)	(0.6680)	(3.3154)	(0.4225)
Mean Dependent Var.	66.98	66.98	66.98	66.98
SD Dependent Var.	(11.85)	(11.85)	(11.85)	(11.85)
N	564	564	564	90
R^2	0.00	0.97		
State and Year FE		✓		✓

Notes: Coefficients from estimating a classic difference-in-differences equation 2.2 of the fertility rate on banking reform in columns (1) and (2) and estimating a specification robust to the timing of treatment as per Borusyak et al. (2023) in columns (3) and (4) . Robust standard errors are clustered at state-level in parentheses. The fertility rate is defined as the number of births per 1000 women aged 20-44 in state s in year t from Bailey et al. (2016). The 'Full Sample' in Panel A covers 1970-2000, the sample in Panel B covers 1970-1988, and Panel C covers 1989-2000. We do not include the states of SD, DE, AK, and HI.

Table B.4: **Impact of Bank Reform on Individual-Level Fertility using Classic TWFE and Borusyak et al. (2023) TWFE**

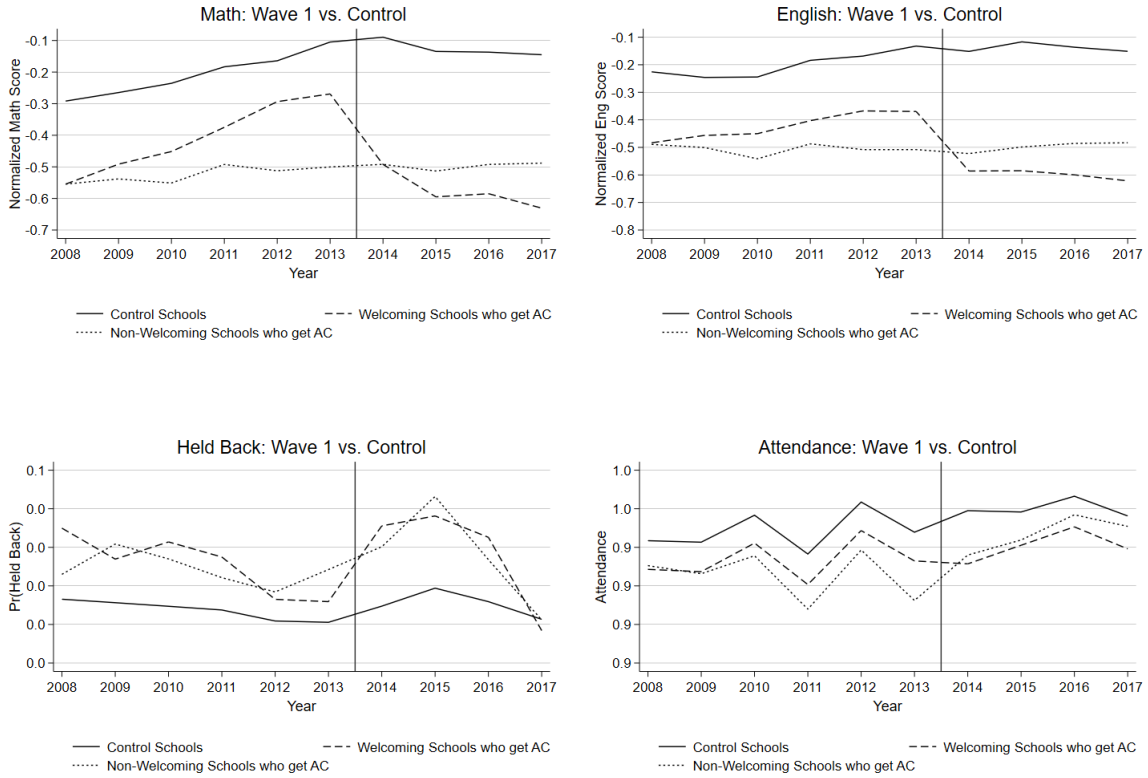
	Classic TWFE		Robust TWFE	
	(1)	(2)	(3)	(4)
Panel A: Full Sample				
Bank Reform	-0.0071***	-0.0056***	-0.0075***	-0.0010
	(0.0015)	(0.0014)	(0.0015)	(0.0021)
Mean Dependent Var.	0.067	0.067	0.067	0.067
SD Dependent Var.	(0.25)	(0.25)	(0.25)	(0.25)
N	681,017	681,017	681,048	416,656
R^2	0.00	0.00		
Panel B: 1977-1988				
Bank Reform	-0.0057**	-0.0033	-0.0056***	-0.0026*
	(0.0020)	(0.0023)	(0.0021)	(0.0015)
Mean Dependent Var.	0.0698	0.0698	0.0698	0.0698
SD Dependent Var.	(0.2548)	(0.2548)	(0.2548)	(0.2548)
N	358,270	358,270	358,299	237,558
R^2	0.00	0.00		
Panel C: 1989-2000				
Bank Reform	-0.0062***	-0.0017	-0.0049**	-0.0006
	(0.0012)	(0.0034)	(0.0025)	(0.0030)
Mean Dependent Var.	0.0638	0.0638	0.0638	0.0638
SD Dependent Var.	(0.2444)	(0.2444)	(0.2444)	(0.2444)
N	322,747	322,747	322,749	32,005
R^2	0.00	0.00		
State and Year FE		✓		✓

Notes: Coefficients from estimating a classic difference-in-differences equation 2.2 of the likelihood of giving birth in the last year on banking reform in columns (1) and (2) and estimating a specification robust to the timing of treatment as per Borusyak et al. (2023) in columns (3) and (4). Robust standard errors are clustered at the state-level in parentheses. The 'Full Sample' in Panel A covers 1970-2000, the sample in Panel B covers 1970-1988, and Panel C covers 1989-2000. Sample: Women aged 20-44 in the CPS March Supplement, not including the states of SD, DE, AK, and HI.

Appendix C: Appendix to Chapter 3

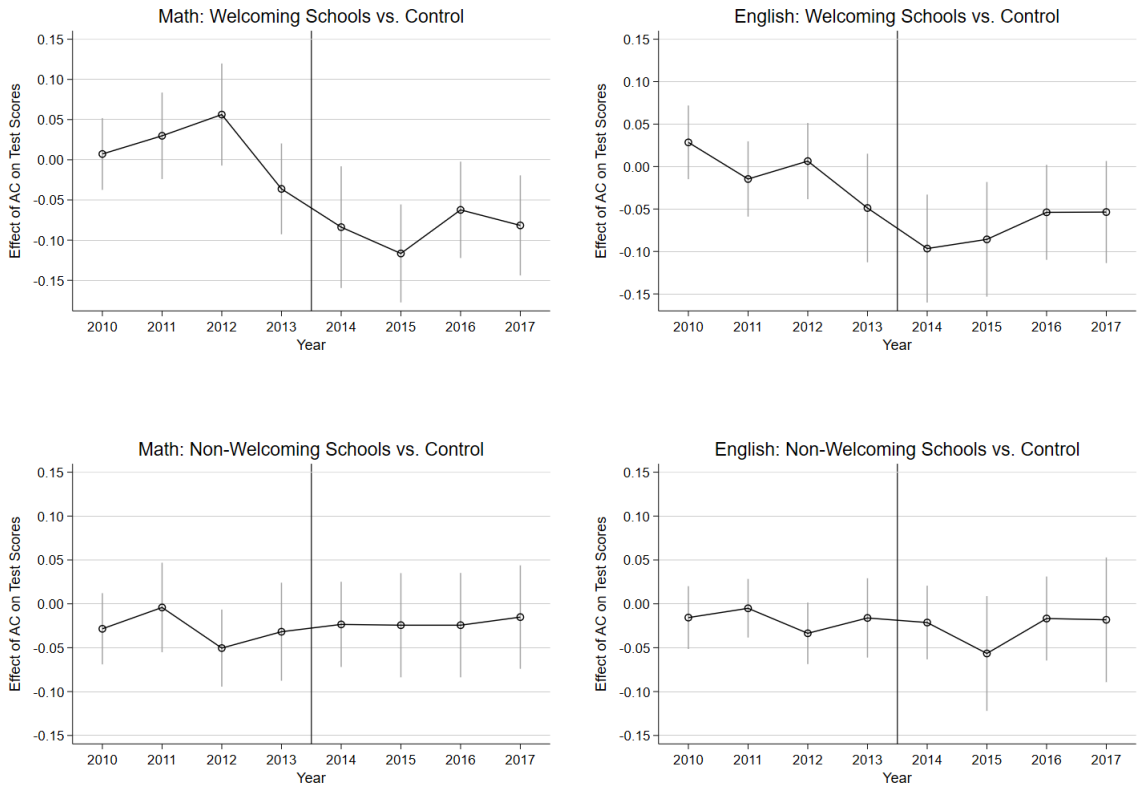
C.1 Appendix Figures

Figure C.1: **Average Test Scores, Grade Retention, and Attendance by Year For Wave 1, Welcoming and Non-Welcoming Schools**



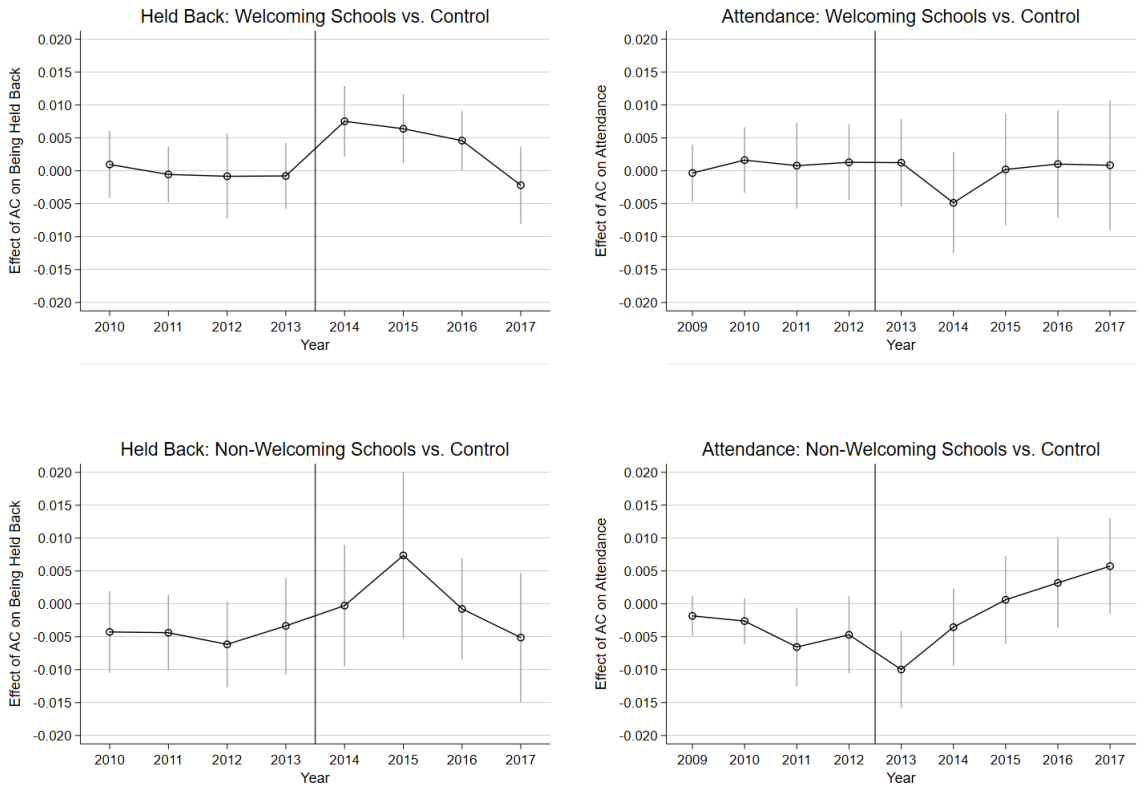
Notes: The figure shows the average annual test scores for math and English, the probability of being held back a grade, and the average school-level attendance in treated and control schools for Wave 1 of the CPS AC installation campaign by whether the school was a welcoming or non-welcoming school. Test scores are standardized by year and grade level using the full Illinois distribution of test scores. The vertical line marks treatment year, 2013-2014. The 'Welcoming Schools' sample includes 33 schools that received AC while also being designated to receive students from the 47 schools that were shut down by CPS that summer, while the 'Non-Welcoming Schools' sample includes 33 schools that received AC but were not designated by CPS to receive students from closed schools.

Figure C.2: **Effects of AC on Test Scores for Wave 1, Welcoming and Non-Welcoming Schools**



Notes: The figure reports difference-in-differences estimates for math and English test scores for treated and control schools from Equation 3.1 for Wave 1 of the CPS AC installation campaign by whether the school was a welcoming or non-welcoming school. Bars represent 95 percent confidence intervals. Vertical line marks treatment year, 2013-2014. Equation 3.1 includes year FE, school FE, and controls for previous year's math and English test scores. Robust standard errors are clustered at the school level. The 'Welcoming Schools' sample includes 33 schools that received AC while also being designated to receive students from the 47 schools that were shut down by CPS that summer, while the 'Non-Welcoming Schools' sample includes 33 schools that received AC but were not designated by CPS to receive students from closed schools.

Figure C.3: **Effects of AC on Grade Retention and Attendance for Wave 1, Welcoming and Non-Welcoming Schools**



Notes: The figure reports difference-in-differences estimates of the probability of being held back for students in treated and control schools from Equation 3.1 on the left and estimates for school-level average student attendance on the right, for wave 1 of the CPS AC installation campaign by whether the school was a welcoming or non-welcoming school. Bars represent 95 percent confidence intervals. The vertical line marks treatment year for all sub-figures. Equation 3.1 includes year FE, school FE, and controls for previous year's math and English test scores for grade retention. Prior math and English controls are not included for attendance because the attendance data is only available at the school level. Robust standard errors are clustered at the school level. The 'Welcoming Schools' sample includes 33 schools that received AC while also being designated to receive students from the 47 schools that were shut down by CPS that summer, while the 'Non-Welcoming Schools' sample includes 33 schools that received AC but were not designated by CPS to receive students from closed schools.

C.2 Appendix Tables

Table C.1: **School-Level Summary Statistics by Wave of Treatment**

	Control Mean (Std. Dev.)	Wave 1 Mean (Std. Dev.)	Waves 2 & 3 Mean (Std. Dev.)	Wave 4 Mean (Std. Dev.)	Control-Wave 1 Difference (T-Stat)	Control-Waves 2 & 3 Difference (T-Stat)	Control-Wave 4 Difference (T-Stat)
Panel A: Full Sample							
Math	-0.295	-0.510	-0.292	-0.544	0.215**	-0.004	0.249***
	(0.509)	(0.335)	(0.446)	(0.353)	(3.377)	(-0.067)	(4.057)
English	-0.281	-0.500	-0.266	-0.563	0.219***	-0.015	0.282***
	(0.494)	(0.278)	(0.442)	(0.448)	(4.020)	(-0.265)	(3.763)
Held Back	0.015	0.026	0.028	0.022	-0.011***	0.002	-0.007*
	(0.015)	(0.013)	(0.111)	(0.017)	(-4.592)	(0.646)	(-2.598)
Attendance	0.945	0.936	0.946	0.922	0.008*	-0.001	0.022**
	(0.027)	(0.016)	(0.022)	(0.044)	(2.493)	(-0.464)	(2.846)
White	11.744	2.263	12.835	2.695	9.481***	-1.091	9.049***
	(18.758)	(4.932)	(18.659)	(6.267)	(6.345)	(-0.446)	(5.546)
Black	39.310	73.790	41.295	67.439	-34.479***	-1.985	-28.129***
	(41.089)	(39.783)	(42.368)	(36.985)	(-4.586)	(-0.360)	(-4.038)
Hispanic	43.482	20.417	39.711	27.376	23.065**	3.772	16.107*
	(37.594)	(35.798)	(35.808)	(33.732)	(3.402)	(0.794)	(2.534)
Low Income	80.876	92.250	82.955	93.906	-11.374***	-2.079	-13.030***
	(23.530)	(10.151)	(22.268)	(7.484)	(-4.834)	(-0.702)	(-6.504)
N	404	33	80	40	437	484	444
Panel B: Energy Star Sample							
Math	-0.164	-0.519	-0.280	-0.586	0.355***	0.116	0.423***
	(0.536)	(0.352)	(0.449)	(0.340)	(4.754)	(1.833)	(6.089)
English	-0.168	-0.509	-0.252	-0.604	0.341***	0.084	0.436***
	(0.535)	(0.292)	(0.445)	(0.456)	(5.236)	(1.334)	(4.993)
Held Back	0.013	0.026	0.015	0.024	-0.013***	-0.001	-0.010**
	(0.013)	(0.013)	(0.012)	(0.017)	(-5.074)	(-0.653)	(-3.230)
Attendance	0.944	0.936	0.946	0.922	0.008*	-0.002	0.022**
	(0.028)	(0.016)	(0.022)	(0.045)	(2.233)	(-0.646)	(2.726)
White	11.811	2.463	13.337	2.774	9.348***	-1.526	9.037***
	(18.624)	(5.146)	(18.866)	(6.351)	(5.846)	(-0.601)	(5.303)
Black	39.088	71.280	41.073	66.453	-32.192***	-1.985	-27.365***
	(40.935)	(41.020)	(42.468)	(37.133)	(-3.964)	(-0.349)	(-3.831)
Hispanic	43.571	22.435	39.216	28.212	21.136**	4.355	15.358*
	(37.422)	(37.065)	(35.565)	(33.922)	(2.876)	(0.894)	(2.353)
Low Income	80.848	92.331	82.414	93.790	-11.483***	-1.566	-12.941***
	(23.532)	(10.371)	(22.553)	(7.573)	(-4.563)	(-0.508)	(-6.158)
AC %	66.959	34.483	34.933	40.000	32.476***	32.025***	26.959***
	(36.514)	(32.248)	(31.295)	(30.923)	(5.011)	(7.308)	(4.549)
Year Built	1948	1934	1925	1941	13.817*	22.856***	7.546
	(34.406)	(32.746)	(28.251)	(35.997)	(2.121)	(5.697)	(1.128)
Heated %	100.000	100.000	100.000	96.970	0.000	0.000	3.030
	(0.000)	(0.000)	(0.000)	(17.408)	(.)	(.)	(1.000)
N	217	29	75	33	246	292	250

Notes: Panel A contains information for the full sample of 557 schools. Panel B contains information on the 354 schools for which we have Energy Star data on AC penetration and other physical school characteristics. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table C.2: **Impact of AC From Difference-in-Differences in Levels (Without Lagged Scores)**

	Full Sample			Low-Performing Students		
	Math (1)	English (2)	Held Back (3)	Math (4)	English (5)	Held Back (6)
Panel A: Full Sample						
Have AC	-0.0126	-0.0054	0.0006	-0.0119	0.0051	0.0011
	(0.0182)	(0.0183)	(0.0015)	(0.0129)	(0.0136)	(0.0027)
N	1,570,976	1,566,628	1,577,349	172,765	173,020	173,817
R^2	0.19	0.17	0.01	0.05	0.04	0.01
Panel B: Energy Star Sample						
Fraction AC	-0.0169	-0.0071	0.0004	-0.0006	0.0316	0.0016
	(0.0269)	(0.0270)	(0.0020)	(0.0194)	(0.0208)	(0.0032)
N	822,858	821,551	826,652	83,086	83,188	83,644
R^2	0.20	0.19	0.01	0.04	0.04	0.01

Notes: This table reports the estimated coefficients from the difference-in-differences model in Equation 3.1 not including lagged student test scores. The dependent variables in columns (1) and (2) are standardized math and English test scores, respectively. The dependent variable in column (3) is if a student is held back. Columns (4), (5), and (6) are analogous to the first three columns but restrict the sample to students performing in the bottom quartile of both the math and English test score distributions. In Panel A, *Have AC* is the main independent variable and is an indicator equal to one if a school has AC in a given year. In Panel B, *Fraction AC* is the main independent variable and is the fraction of the school that was air-conditioned prior to the AC installation campaign as reported in the 2011 Energy Star report. This variable is equal to 1 after a school receives AC. Robust standard errors clustered at the school level are in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table C.3: **Impact of AC From Difference-in-Differences, Low-Performing Students**

	Math (1)	English (2)	Held Back (3)
Panel A: Full Sample			
Have AC	-0.0147	-0.0011	0.0012
	(0.0112)	(0.0117)	(0.0027)
N	172,638	172,898	173,687
R^2	0.16	0.22	0.01
Panel B: Energy Star Sample			
Fraction AC	-0.0064	0.0151	0.0014
	(0.0174)	(0.0186)	(0.0031)
N	82,990	83,094	83,544
R^2	0.13	0.20	0.01

Notes: This table reports the estimated coefficients from the difference-in-differences model in Equation 3.1 for students in the bottom of both the math and English test score distributions. The dependent variables in columns (1) and (2) are standardized math and English test scores, respectively. The dependent variable in column (3) is if a student is held back. In Panel A, *Have AC* is the main independent variable and is an indicator equal to one if a school has AC in a given year. In Panel B, *Fraction AC* is the main independent variable and is the fraction of the school that was air-conditioned prior to the AC installation campaign as reported in the 2011 Energy Star report. This variable is equal to 1 after a school receives AC. Robust standard errors clustered at the school level are in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table C.4: **Impact of AC From Difference-in-Differences by Wave of Treatment in Levels (Without Lagged Scores)**

	Wave 1			Waves 2 & 3			Wave 4		
	Math (1)	English (2)	Held Back (3)	Math (4)	English (5)	Held Back (6)	Math (7)	English (8)	Held Back (9)
Panel A: Full Sample									
Have AC	-0.0525	-0.0611**	0.0032	0.0024	0.0170	0.0008	0.0520	0.0127	-0.0113*
	(0.0326)	(0.0274)	(0.0027)	(0.0223)	(0.0231)	(0.0018)	(0.0365)	(0.0324)	(0.0063)
N	1,229,844	1,226,925	1,234,935	1,392,397	1,388,484	1,397,759	1,224,627	1,221,647	1,229,571
R^2	0.20	0.18	0.01	0.19	0.18	0.01	0.20	0.18	0.01
Panel B: Energy Star Sample									
Fraction AC	-0.0416	-0.0500	0.0047	-0.0102	0.0030	0.0005	0.0247	0.0269	-0.0133
	(0.0547)	(0.0416)	(0.0037)	(0.0331)	(0.0354)	(0.0023)	(0.0375)	(0.0455)	(0.0110)
N	607,891	607,099	610,766	718,691	717,611	721,789	606,090	605,311	608,771
R^2	0.21	0.20	0.01	0.19	0.19	0.01	0.20	0.20	0.01

Notes: This table reports the estimated coefficients from the difference-in-differences model in Equation 3.1 not including lagged student test scores. In Panel A, *Have AC* is the main independent variable and is an indicator equal to one if a school has AC in a given year. In Panel B, *Fraction AC* is the main independent variable and is the fraction of the school that was air-conditioned prior to the AC installation campaign as reported in the 2011 Energy Star report. This variable is equal to 1 after a school receives AC. Wave 1 schools received AC in 2013-14, wave 2 in Summer 2014, wave 3 in October 2014, and wave 4 in 2016-17. Robust standard errors clustered at the school level are in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table C.5: **Impact of AC From Difference-in-Differences for Schools with Lowest Prior AC Coverage**

	Math (1)	English (2)	Held Back (3)	Attendance (4)
Have AC	0.0147	0.0131	0.0009	0.0030**
	(0.0214)	(0.0172)	(0.0011)	(0.0013)
N	436,584	437,335	438,576	2,394
R^2	0.71	0.69	0.01	0.83

Notes: This table reports the estimated coefficients from the difference-in-differences model in Equation 3.1 while restricting the sample of treated schools to only those that had less than 30% of the school air-conditioned prior to being treated. The dependent variables in columns (1) and (2) are standardized math and English test scores, respectively. The dependent variable in column (3) is if a student is held back. The dependent variable in column (4) is average student attendance at the school-level, and therefore does not include lagged student test scores. *Have AC* is the main independent variable and is an indicator equal to one if a school has AC in a given year. Robust standard errors clustered at the school level are in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Bibliography

Evan K Rose. The rise and fall of female labor force participation during world war ii in the united states. *The Journal of Economic History*, 78(3):673–711, 2018.

Daron Acemoglu, David H Autor, and David Lyle. Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. *Journal of political Economy*, 112 (3):497–551, 2004.

Dina Shatnawi and Price Fishback. The impact of world war ii on the demand for female workers in manufacturing. *The Journal of Economic History*, 78(2):539–574, 2018.

Yuliya Demyanyk. Us banking deregulation and self-employment: A differential impact on those in need. *Journal of Economics and Business*, 60(1-2):165–178, 2008.

Thorsten Beck, Ross Levine, and Alexey Levkov. Big bad banks? the winners and losers from bank deregulation in the united states. *The Journal of Finance*, 65(5):1637–1667, 2010.

Ross Levine, Alexey Levkov, Yona Rubinstein, et al. Bank deregulation and racial inequality in america. *RISK*, 2012.

Alexander Popov and Sonia Zaharia. Credit market competition and the gender gap in labor force participation: Evidence from local markets. *European Economic Review*, 115:25–59, 2019.

R Jisung Park, A Patrick Behrer, and Joshua Goodman. Learning is inhibited by heat exposure, both internationally and within the united states. *Nature human behaviour*, 5(1):19–27, 2021.

Kirill Borusyak, Xavier Jaravel, and Jann Spiess. Revisiting event study designs: Robust and efficient estimation, 2023.

Claudia Goldin. The role of world war ii in the rise of women's employment. *The American Economic Review*, pages 741–756, 1991a.

Claudia Goldin and Claudia Olivetti. Shocking labor supply: A reassessment of the role of world war ii on women's labor supply. *American Economic Review*, 103(3):257–62, 2013.

Claudia Goldin. Marriage bars: discrimination against married women workers, 1920's to 1950's. In Patrice LR Higonnet, David S Landes, and Henry Rosovsky, editors, *Favorites of fortune: technology, growth, and economic development since the Industrial Revolution*. Harvard University Press, 1991b.

Raquel Fernández, Alessandra Fogli, and Claudia Olivetti. Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics*, 119(4): 1249–1299, 2004.

Raquel Fernández. Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, 103(1):472–500, 2013.

Matthias Doepke, Moshe Hazan, and Yishay D Maoz. The baby boom and world war ii: A macroeconomic analysis. *The Review of Economic Studies*, 82(3):1031–1073, 2015.

Matthew F Larsen, Theodore J McCarthy, Jeremy G Moulton, Marianne E Page, and Ankur J Patel. War and marriage: assortative mating and the world war ii gi bill. *Demography*, 52(5): 1431–1461, 2015.

Ken Burns. *War production*. Public Broadcasting Service, Sep 2007. URL `https://www.pbs.org/kenburns/the-war/war-production`.

Susan M. Hartmann. *The home front and beyond: American women in the 1940s*. Boston: Twayne Publishers, 1982.

Ruth Milkman. *Gender at work: The dynamics of job segregation by sex during World War II*, volume 314. University of Illinois Press, 1987.

Tracey Warm. Wartime production. *OAH Magazine of History*, 16(3):47–52, 2002. ISSN 0882228X. URL `http://www.jstor.org/stable/25163526`.

Mary Elizabeth Pidgeon. Women workers and recent economic change. *Monthly Labor Review*, pages 666–671, 1947.

USES. Occupations suitable for women. *Occupational Analysis Section*, July 1942.

Mildred A. Joiner and Clarence M. Welner. Bulletin: Employment of women in war production. *Reports and Analysis Division, Bureau of Employment Security*, July 1942.

Encyclopedia.com. Mobilization of women — encyclopedia.com, 2021. URL `https://www.encyclopedia.com/history/news-and-education-magazines/mobilization-women`.

Gary S Becker. Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2):9–49, 1962.

USES. A short history of the war manpower commission. *Technical Service Division, Dept. of Labor*, June 1948.

Ruth Milkman. Redefining 'women's work': The sexual division of labor in the auto industry during world war ii. *Feminist Studies*, 8(2):337–372, 1982.

Peter J Brennan. The us employment service: Forty years of progress. *Manpower Administration, Dept. of Labor*, 1973.

Barbara R Bergmann and Irma Adelman. The 1973 report of the president's council of economic advisers: The economic role of women. *The American Economic Review*, 63(4):509–514, 1973.

Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. Ipums-usa, 2020. URL `https://doi.org/10.18128/D010.V10.0`.

David H Autor, Lawrence F Katz, and Alan B Krueger. Computing inequality: have computers changed the labor market? *The Quarterly journal of economics*, 113(4):1169–1213, 1998.

David Card, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):S13–S70, 2018.

Lawrence F Katz and Kevin M Murphy. Changes in relative wages, 1963–1987: supply and demand factors. *The quarterly journal of economics*, 107(1):35–78, 1992.

Reports War Manpower Commission. The labor market 1943-1945. *Reports and Analysis Service*, 1944.

Karen Tucker Anderson. Last hired, first fired: Black women workers during world war ii. *The Journal of American History*, 69(1):82–97, 1982.

USES. War and post -war trends in employment of negroes. *Monthly Labor Review*, January 1945.

D'Ann Campbell. *Women at war with America: Private lives in a patriotic era*, volume 2. Living Water Institute AG, 1984.

James H Stock and Motohiro Yogo. Testing for weak instruments in linear iv regression, 2002.

Associated Press AP. Us birth rate sees biggest fall for nearly 50 years. *The Guardian*, 2021. URL `https://www.theguardian.com/us-news/2021/may/05/us-birth-rate-sees-biggest-fall-for-nearly-50-years`.

Sabrina Tavernise. The u.s. birthrate has dropped again. the pandemic may be accelerating the decline. *The New York Times*, 2021. URL `https://www.nytimes.com/2021/05/05/us/us-birthrate-falls-covid.html`.

Courtney Vinopal. The u.s. birth rate began dropping years before the pandemic. here's why. *pbs.org*, 2021. URL `https://shorturl.at/ABLNZ`.

Brady E. Hamilton, Joyce A. Martin, and Michelle J.K. Osterman. *Births: Provisional Data for 2020*. 2021. URL `https://www.cdc.gov/nchs/data/vsrr/vsrr012-508.pdf`.

Melissa S. Kearney and Phillip B. Levine. *The coming COVID-19 baby bust: Update*. 2020. URL https://www.brookings.edu/blog/up-front/2020/12/17/the-coming-covid-19-baby-bust-update/.

Sarah Adelman, Mia Charifson, Eunsil Seok, Shilpi S Mehta-Lee, Sara G Brubaker, Mengling Liu, and Linda G Kahn. State-specific fertility rate changes across the usa following the first two waves of covid-19. *Human Reproduction*, page dead055, 2023.

Martha J Bailey, Janet Currie, and Hannes Schwandt. The covid-19 baby bump: The unexpected increase in us fertility rates in response to the pandemic. Technical report, National Bureau of Economic Research, 2022.

Deirdre McPhillips. Covid-19 'baby bump' brought an increased us fertility rate in 2021 – but also record high preterm births. *CNN*, January 2023. URL https://www.cnn.com/2023/01/31/health/us-birth-rate-fertility-trends-2021/index.html.

Raquel Fernandez and Alessandra Fogli. Culture: An empirical investigation of beliefs, work, and fertility. *American economic journal: Macroeconomics*, 1(1):146–77, 2009.

Robert J Willis. A new approach to the economic theory of fertility behavior. *Journal of political Economy*, 81(2, Part 2):S14–S64, 1973.

William P Butz and Michael P Ward. The emergence of countercyclical us fertility. *The American economic review*, 69(3):318–328, 1979.

Claudia Goldin and Lawrence F Katz. The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of political Economy*, 110(4):730–770, 2002.

Phillip B Levine, Douglas Staiger, Thomas J Kane, and David J Zimmerman. Roe v wade and american fertility. *American Journal of Public Health*, 89(2):199–203, 1999.

Dean F Amel and J Nellie Liang. The relationship between entry into banking markets and changes in legal restrictions on entry. *The Antitrust Bulletin*, 37(3):631–649, 1992.

Jith Jayaratne and Philip E Strahan. Entry restrictions, industry evolution, and dynamic efficiency: Evidence from commercial banking. *The Journal of Law and Economics*, 41(1): 239–274, 1998.

Martha Bailey, Price Fishback, Michael Haines, Shawn Kantor, Edson Severnini, and Anna Wentz. Us county-level natality and mortality data, 1915-2007. 2016.

Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, and J Robert Warren. Ipums-cps. *Integrated Public Use Microdata Series, Current Population Survey: Version 8.0 [dataset]*, 2020.

Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021.

Brantly Callaway and Pedro HC Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021.

Jith Jayaratne and Philip E Strahan. The finance-growth nexus: Evidence from bank branch deregulation. *The Quarterly Journal of Economics*, 111(3):639–670, 1996.

Astrid A Dick and Andreas Lehnert. Personal bankruptcy and credit market competition. *The Journal of Finance*, 65(2):655–686, 2010.

Sandra E Black and Philip E Strahan. Entrepreneurship and bank credit availability. *The Journal of Finance*, 57(6):2807–2833, 2002.

William R Kerr and Ramana Nanda. Democratizing entry: Banking deregulations, financing constraints, and entrepreneurship. *Journal of Financial Economics*, 94(1):124–149, 2009.

Tara Rice and Philip E Strahan. Does credit competition affect small-firm finance? *The Journal of Finance*, 65(3):861–889, 2010.

Yuliya Demyanyk, Charlotte Ostergaard, and Bent E Sørensen. Us banking deregulation, small businesses, and interstate insurance of personal income. *The Journal of Finance*, 62(6):2763–2801, 2007.

Donald P Morgan, Bertrand Rime, and Philip E Strahan. Bank integration and state business cycles. *The Quarterly Journal of Economics*, 119(4):1555–1584, 2004.

Stephen Teng Sun and Constantine Yannelis. Credit constraints and demand for higher education: Evidence from financial deregulation. *Review of Economics and Statistics*, 98(1):12–24, 2016.

Giovanni Favara and Jean Imbs. Credit supply and the price of housing. *American Economic Review*, 105(3):958–92, 2015.

Lisa J Dettling and Melissa S Kearney. House prices and birth rates: The impact of the real estate market on the decision to have a baby. *Journal of Public Economics*, 110:82–100, 2014.

Michael F Lovenheim and Kevin J Mumford. Do family wealth shocks affect fertility choices? evidence from the housing market. *Review of Economics and Statistics*, 95(2):464–475, 2013.

BLS Bureau of Labor Statistics. National longitudinal survey of youth 1979 cohort, 1979-2016 (rounds 1-27). *U.S. Department of Labor*, 2019. URL `https://www.cnn.com/2023/01/31/health/us-birth-rate-fertility-trends-2021/index.html`.

Randall S Kroszner and Philip E Strahan. What drives deregulation? economics and politics of the relaxation of bank branching restrictions. *The Quarterly Journal of Economics*, 114(4):1437–1467, 1999.

US Bureau of Economic Analysis. Regional economic accounts. 2020.

Marianne Bertrand and Sendhil Mullainathan. Enjoying the quiet life? corporate governance and managerial preferences. *Journal of political Economy*, 111(5):1043–1075, 2003.

Andrew Cherlin, Erin Cumberworth, S Philip Morgan, and Christopher Wimer. The effects of the great recession on family structure and fertility. *The Annals of the American Academy of Political and Social Science*, 650(1):214–231, 2013.

D Echeverria, V Barnes, and A Bittner. The impact of environmental conditions on human performance: A handbook of environmental exposures. volume 1. Technical report, Nuclear Regulatory Commission, Washington, DC (United States)., 1994.

MV Jokl. The effect of the environment of human performance. *Applied Ergonomics*, 13(4): 269–280, 1982.

Jerry D Ramsey. Task performance in heat: a review. *Ergonomics*, 38(1):154–165, 1995.

Alan Barreca, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S Shapiro. Adapting to climate change: The remarkable decline in the us temperature-mortality relationship over the twentieth century. *Journal of Political Economy*, 124(1):105–159, 2016.

Tord Kjellstrom, David Briggs, Chris Freyberg, Bruno Lemke, Matthias Otto, and Olivia Hyatt. Heat, human performance, and occupational health: a key issue for the assessment of global climate change impacts. *Annual Review of Public Health*, 37:97–112, 2016.

Donn E Hancher and Hersham A Abd-Elkhalek. The effect of hot weather on construction labor productivity and costs. *Cost Engineering*, 40(4):32, 1998.

Daniël Wendt, Luc JC van Loon, and Wouter D Marken Lichtenbelt. Thermoregulation during exercise in the heat. *Sports Medicine*, 37(8):669–682, 2007.

Wen Yi and Albert PC Chan. Effects of heat stress on construction labor productivity in hong kong: a case study of rebar workers. *International Journal of Environmental Research and Public Health*, 14(9):1055, 2017.

Olli Seppanen, William J Fisk, and QH Lei. Effect of temperature on task performance in office environment. 2006.

Tord Kjellstrom, Ingvar Holmer, and Bruno Lemke. Workplace heat stress, health and productivity–an increasing challenge for low and middle-income countries during climate change. *Global Health Action*, 2(1):2047, 2009.

Geoffrey Heal and Jisung Park. Reflections—temperature stress and the direct impact of climate change: a review of an emerging literature. *Review of Environmental Economics and Policy*, 10(2):347–362, 2016.

Hyunkuk Cho. The effects of summer heat on academic achievement: a cohort analysis. *Journal of Environmental Economics and Management*, 83:185–196, 2017.

R Jisung Park, Joshua Goodman, Michael Hurwitz, and Jonathan Smith. Heat and learning. *American Economic Journal: Economic Policy*, 12(2):306–39, 2020.

Heather Randell and Clark Gray. Climate variability and educational attainment: Evidence from rural ethiopia. *Global Environmental Change*, 41:111–123, 2016.

Heather Randell and Clark Gray. Climate change and educational attainment in the global tropics. *Proceedings of the National Academy of Sciences*, 116(18):8840–8845, 2019.

Stephanie Riegg Cellini, Fernando Ferreira, and Jesse Rothstein. The value of school facility investments: Evidence from a dynamic regression discontinuity design. *The Quarterly Journal of Economics*, 125(1):215–261, 2010.

Christopher A Neilson and Seth D Zimmerman. The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics*, 120:18–31, 2014.

Cheryl Corley. Discord grows over chicago's planned school closings. *NPR.org*, Mar 2013. URL https://www.npr.org/2013/03/28/175550937/chicago-school-closings-widens-gap-between-teachers-city.

Valerie Strauss. Why shouldn't chicago teachers ask for air conditioned schools? *Washington Post*, Sep 2012. URL https://www.washingtonpost.com/blogs/answer-sheet/post/why-shouldnt-chicago-teachers-ask-for-air--conditioned-schools/2012/09/12/453c61fe-fc34-11e1-a31e-804fccb658f9_blog.html.

Sarah Chambers. Chicago teacher: Inhumane working conditions are inhumane learning conditions. *Education Week*, Sep 2013. URL http://blogs.edweek.org/teachers/living-in-dialogue/2013/09/chicago_teacher_inhumane_worki.html.

cps.edu Chicago Public Schools Press Releases. Mayor emanuel and cps announce every classroom will be air conditioned by spring 2017. *Chicago Public Schools: Press Releases*, Aug 2016. URL https://web.archive.org/web/20161027141458/http://cps.edu/News/Press_releases/Pages/PR1_08_22_2016.aspx.

H Spencer Banzhaf and Randall P Walsh. Do people vote with their feet? an empirical test of tiebout. *American economic review*, 98(3):843–63, 2008.

Jay Lawrimore, Ron Ray, Scott Applequist, Bryant Korzeniewski, and Matthew Menne. Global summary of the month and year, version 1.0. *NOAA National Centers for Environmental Information*, 2016.

A Arguez, I Durre, S Applequist, M Squires, R Vose, X Yin, and R Bilotta. Noaa's us climate normals (1981–2020) daily summaries: Chicago o hare, noaa national centers for environmental information, 2020.

Matt Barnum. Exclusive: Too hot to learn: Records show nearly a dozen of the biggest school districts lack air conditioning. *the74million.org*, Jun 2017. URL https://shorturl.at/pvKX5.

Press Office NYC City Hall. Mayor de blasio, chancellor fariña and city council announce every classroom will have air conditioning by 2022. *The Official Website of the City of New York*, Apr 2017. URL https://shorturl.at/bns68.

cps.edu Chicago Public Schools. Stats and facts, 2020. URL https://www.cps.edu/about/stats-facts/.

Stuart Luppescu, Elaine Allensworth, Paul Moore, Marisa de la Torre, James Murphy, and Sanja Jagesic. Trends in three eras of cps. *Consortium on Chicago School Research*, 2011. URL https://consortium.uchicago.edu/sites/default/files/2018-10/Trends_in_Three_Eras_of_CPS.pdf.

Hannah Max. Jun 2023. URL https://www.illinoispolicy.org/chicago-students-score-lower-fewer-graduate-fewer-go-to-college/.

Michael Pearson and Holly Yan. Official: No deal yet between chicago teachers and school system. *CNN.com*, Sep 2012. URL http://www.cnn.com/2012/09/10/us/illinois-chicago-teachers-strike/index.html?hpt=hp_t1.

Noreen S. Ahmed-Ullah. Protests held at chicago school headquarters. *Chicago Tribune*, Apr 2013. URL https://www.chicagotribune.com/news/ct-xpm-2013-04-24-chi-cps-protests-20130424-story.html.

Eydar Peralta. In chicago, dozens arrested as they protest school closures. *NPR.org*, Mar 2013. URL https://www.npr.org/sections/thetwo-way/2013/03/27/175524955/in-chicago-dozens-arrested-as-they-protest-school-closures.

High-Stakes Testing Citizens For Public Schools. Parcc fact sheet. *citizensforpublicschools.org*, 2017. URL https://www.citizensforpublicschools.org/mcas-and-high-stakes-testing/parcc-fact-sheet/.

Marisa De la Torre, Molly F Gordon, Paul Moore, and Jennifer Cowhy. School closings in chicago. Technical report, Research Report. Consortium, 2015.

Roland G Fryer Jr. The production of human capital in developed countries: Evidence from 196 randomized field experiments. In *Handbook of Economic Field Experiments*, volume 2, pages 95–322. Elsevier, 2017.

Raj Chetty, John N Friedman, and Jonah E Rockoff. Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9):2593–2632, 2014.

Alan B Krueger. Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2):497–532, 1999.

Roland G Fryer Jr and Meghan Howard-Noveck. High-dosage tutoring and reading achievement: Evidence from new york city. *Journal of Labor Economics*, 38(2):421–452, 2020.