Title

Essays on Labor and Employment in the Macroeconomy

Author

Eisner, Emily Murphy

Publication Date

2021

Peer reviewed|Thesis/dissertation

Essays on Labor and Employment in the Macroeconomy

by

Emily Murphy Eisner

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Christina Romer, Chair
Professor Emi Nakamura
Professor Abhay Aneja

Summer 2021

Essays on Labor and Employment in the Macroeconomy

Abstract

Essays on Labor and Employment in the Macroeconomy

by

Emily Murphy Eisner

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Christina Romer, Chair

This dissertation reflects research on the US labor markets over the course of large macroeconomic events. Each chapter studies a particular facet of how to understand movements in labor markets and individual lives as the macroeconomy shifts. In my first chapter, I introduce and motivate each of the subsequent chapters as essential for furthering our collective understanding of human well-being in an ever evolving and fluctuating macroeconomy.

In the second chapter of my dissertation, I study the impact of technological change on workers by examining the adoption of the tractor in American agriculture. The paper uses complete count census data at both the county and individual level to study the impact of farm mechanization in the first half of the twentieth century. I examine the individual-level outcomes and adjustment mechanisms for impacted individuals. I find that the generation of farm workers that was older than 18 when the tractor was adopted saw large disemployment effects and an exodus from the labor market, though no strong migratory patterns. In contrast, youth who grew up on farms were likely to stay in farm work and do not show strong propensity to migrate out of their county of origin. I conclude that changing technology requires new tasks and skills, making it difficult for incumbent workers to adjust. New entrants, however, have the capacity to adapt to the changing work requirements.

In the third chapter, co-authors and I show that the standard implementation of X13-ARIMA yields predictable revisions to initial releases of employment growth. At the onset of severe recessions, initial growth rates tend to be revised downward. This causes the severity of recessions to be underestimated in real-time– a major concern for policy-makers. The problem arises from the combination of concurrent seasonal adjustment, and the standard forecasting procedure used as part of the seasonal adjustment algorithm, when applied to abrupt movements such as those that occur at the onset of a recession. This problem was avoided in the COVID crisis only due to the triggering of outlier detection clauses, which counteracted this effect.

In the fourth chapter, I study the empirical accuracy of home production in macroeconomic models. I begin by building a simple RBC-style model with home production and two income earners. I solve and simulate the model to demonstrate how shocks to productivity may impact time allocation between two spouses in the home. I then use the American Time Use Survey to study time allocation in response to the Great Recession. I use state-level variation in the unemployment rate over the Great Recession to measure how men and women's time spent on market work, home production, childcare and leisure varied with labor market conditions. I find that men's market work does decline with state-level unemployment and that men's home production and childcare do increase. Women's time use does not appear to change. Finally, I test the model by estimating how men's and women's time in home production and market work relate. I find that contrary to the prediction of the model, men's and women's time in home production do not co-move. This finding calls for further work to understand how to better model time allocation and substitution within a household.

To my parents, who have always believed in the best parts of me.
And to Lily, my sister and best friend, who has always challenged me to be better.

Contents

List of Figures

List of Tables

Acknowledgments

There is no question that I could not have completed this dissertation without the love and support of so many people. First and foremost, I'd like to thank the many faculty at the University of California, Berkeley who gave their time to teach and advise me. This group includes Abhay Aneja, Brad DeLong, Ellora Derenoncourt, Hilary Hoynes, Emi Nakamura, Martha Olney, David Romer, and Jón Steinsson. Each of you has made an indelible impact on my life and character. Most importantly, Christina Romer has been a role model and steadfast mentor through this challenging phase of my life. Thank you for your warmth as well as your determination in making me a strong and inquisitive economist.

The administrative staff in the Economics department have been amazing to work with. Thanks, in particular, to Patrick Allen and Heather Iwata for being such kind and caring people in our community.

Many people and institutions are to thank for technical and financial support. Thanks to Carl Boe, Carl Mason, and Josh Goldstein at the UC Berkeley Demography Department for access to and help with the Full Count Census Data. Thanks to Leah Boustan, Ran Abramitzky, Kathryn Eriksson and Helen Kissel for their help with linking the census records. Thanks the National Science Foundation for financial support during my time in graduate school. And thanks to the Economic History group for providing many helpful conversations and some much-needed structure.

The writing of this dissertation coincided with some of the most challenging times in my life to date. I can't overstate the gratitude I have for my friends and family for supporting me as I navigated my way through the storm - always showing me light in the darkness, and creating calm in the turmoil. Thanks to Ingrid Haegele, Ben Scuderi, Nina Roussille, Eva Lyubich, Peter McCrory, Ceren Baysan for your friendship in graduate school. Thanks to Jacob Rosenberg for joining with me to make a truly difficult year into something full of laughter and stimulating conversation. Thanks to Eliza Huntington, Kathryn Peake, and Eva Petzinger for riding the tides with me throughout the last decade. Thanks to my parents, sister and the whole Eisner-Murphy-Wiberg family who have all shown up for me every time I've asked (and some times when I haven't). Thanks to Sarah Scheckter and Jackie Gotthold, who have helped me embrace myself fully and unconditionally. Finally, thanks to my friend Alex Kroeger, who supplied me with seemingly bottomless silliness, support, and peanut-butter pretzels as I put this dissertation together.

Chapter 1

Introduction

This dissertation studies labor markets through the lens of macroeconomics and macroeconomic shocks. Each chapter focuses on a unique set of data and a unique research question, all intended to provide new information on how macroeconomic policy-makers should understand and react to events such as a recession or technological change.

Historically, macroeconomics has employed aggregate data and time series variation to study the workings of the economy. More recently, macroeconomists have moved towards exploiting cross-sectional data to better understand the mechanisms through which the economy evolves and adjusts in response to shocks. My research straddles both approaches, taking the view that many of the assumptions in our economic models and data measurement require closer examination. In particular, I focus my investigation on three key issues in macro-labor: First, how technological change impacts individuals and their progeny over the medium- and long-run; Second, how real time aggregate employment data can underestimate the severity of an economic downturn; Third, how intra-household divisions of labor may stray dramatically from standard assumptions made in macroeconomic models and why this may be of increasing importance to the overall functioning of the economy. Overall, these studies contribute to a growing literature reevaluating the role of heterogeneity, frictions to adjustment, and data measurement in macro-labor.

In Chapter 2, I study the medium- and long-run response of individuals to a changing technological landscape. I use tractor adoption in the first half of the twentieth century to observe the individual response of farmers to rapidly improving agricultural technology. This period in history offers an episode in which technological change forced large and lasting changes to labor demand. Rich micro-data available for the period provides an opportunity to study the micro-level adjustments that individuals made as the technology in their local community shifted. My data allows me to not only see regional changes, but to observe individual-level adjustments in terms of both labor market participation and migration. I find that tractor adoption on local farms decreased employment for incumbent workers and increased “out of labor force” status. In contrast, new entrants to the labor market appear to find local agricultural employment. The observed distinction between incumbents and new entrants suggests an important insight into adoption of new skills that meet the requirements of new technology. By exposing this heterogeneity, my research adds to our collective knowledge of economic change and human well-being in the face of evolving technology.

Chapter 3 focuses on measurement of macroeconomic data in real time. In this study, my coauthors and I focus on the real time employment data published by the Bureau of Labor Statistics each month. We show that the seasonal adjustment methodology employed by the Bureau of Labor Statistics can bias real-time data during an economic recession. By examining the specific mechanisms of the seasonal adjustment methodology, we offer a set of alternative procedures that improve the real-time accuracy of employment data. This paper contributes methodologically and theoretically to underlying questions regarding data validity and empirical methods in macroeconomics. That is, to be able to study and understand labor markets, we must first understand the data that we use to study them.

Finally, my fourth chapter examines home production over the business cycle. This project uses time-use data from the American Time Use Survey to show how individual's time use varies over the course of a recession, and how, in particular, the time use of two heterosexual spouses varies in ways not currently modeled in our basic business cycle models. The project contributes to the literature on individual heterogeneity as a key input to the aggregate economy. Further, the project works to establish complementarities and dependencies in individual behavior that may be of key importance to the modern economy. If we have learned anything from the COVID-19 recession, it is that our home lives and work lives are deeply entangled and thus understanding our home lives is key to understanding the health of the economy as a whole.

Overall, my dissertation contributes to the growing literature on observed heterogeneity and complexity in macro-labor markets. Each chapter takes the stance that by deeply investigating the nuances of individual behavior and data collection, we can better improve policy and the economy as a whole.

Chapter 2

Individual Impacts of Technological Change: The Case of Farm Mechanization in 20th Century US

This chapter of my dissertation is taken from a paper I wrote over the past three years. The paper focuses on the disemployment impacts of evolving technology in the first half of the 20th century. The chapter uses micro data on individuals living in the Midwest US as tractors became more commonly used on farms. The work uncovers heterogeneity in how individuals respond to changing technology by age and by industry. The paper contributes to my overall thesis by offering a historical view into issues such as structural change and technological innovation that are increasingly relevant today.

2.1 Introduction

In both the developed and the developing world, technological progress underlies key debates of how to protect and provide for workers and families. In the developing world, where agricultural employment remains high as a proportion of total employment, much of economic research is devoted to understanding and inducing individuals' decisions to adopt agricultural technology (Duflo, Kremer, and Robinson, 2011; Suri, 2011; Kremer, Rao, and Schilbach, 2019). In contrast, in the developed world, where agricultural employment has reached all-time lows, technological change dominates debates over the future of employment, income inequality, and market capitalism. With the decline of the labor share of income and the increase in income inequality, technology competes as one of the most talked-about challenges to the US workforce and to the US economy more broadly (Acemoglu and Restrepo, 2020; Acemoglu and Restrepo, 2018b; Autor, Dorn, and Hanson, 2013; Karabarbounis and Neiman, 2014; Elsby, Hobijn, and Şahin, 2013; Abdih and Danninger, 2017).

In this paper, I study technological change in US agriculture in the first half of the 20th century, a period that occupies the nexus between a 'developing' and 'developed' nation state. In particular, I study the adoption of the tractor and the tractor's subsequent impact on labor markets and individual workers. I build off of foundational work by Olmstead and Rhode, 2001 by combining the agricultural census data for US counties between 1900 and 1940 to the linked historical Decennial US Census micro-data.[1,2] With this expanded dataset, I am able to document the medium- and long-run consequences of technological change on agricultural workers and their progeny.

The main results of my study are as follows. Using a linked sample of individuals in the 1910 and 1930 Decennial Census files, I find that for incumbent agricultural workers in 1910, the introduction of tractors onto farms led to a significant reduction in 1930 employment (and a parallel increase in "out of labor force" status). In contrast, for farm residents under the age of 18 in 1910, local tractor adoption significantly increased 1930 agricultural employment, but only slightly increased the overall probability of 1930 employment. These results align closely with a task-based model of automation which includes both a displacement effect and a reinstatement effect associated with newly adopted technology. With technological adoptions that alter the set of tasks required in a given sector, incumbent workers lose necessary skills, but their children (the new entrants into the labor market) are able to build skills and remain employed in the respective sector. This result is important as it sheds light on the relative propensity for humans to adjust to changing labor markets and work environments. My research makes clear that adjustment to technological change occurs between generations rather than within an individual's life.

Examining outcomes for a sample of individuals linked between the 1910 and 1940 Census records shows a smaller, but still significant disemployment effect for incumbent agricultural workers. The attenuation in estimates for the 1910-1940 relative to the 1910-1930 sample could be explained by the fact that over a longer time horizon, the impact on incumbent workers smooths geographically. Alternatively, the diminished observed impact could be because the 1910-1940 linked sample is different in composition than the 1910-1930 linked sample. Regardless, over a 30 year time horizon, incumbent workers see small but significant disemployment effects in response to tractor adoption. For 1910 youth farm residents, I continue to find large increases in the probability of 1940 agricultural employment along with declines in manufacturing employment, declines in relative wage income, and declines in the likelihood of attaining some high school education. These declines in education and income come alongside evidence of increased non-wage earnings and a decline in the probability of migrating out of one's county of origin. Thus, the welfare impacts of tractor adoption remain ambiguous.

Taken as a whole, these findings align with recent works by Eckert and Peters, 2017 and Alvarez-Cuadrado and Poschke, 2011. In their work, Eckert and Peters, 2017 find

[1]Haines, Michael, Fishback, Price, and Rhode, Paul. United States Agriculture Data, 1840 - 2012. Inter-university Consortium for Political and Social Research [distributor], 2018-08-20. https://doi.org/10.3886/ICPSR35206.v4

[2]Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 10.0 1900-1940 Full Count USA. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D010.V10.0

that initial agricultural employment shares are strongly negatively correlated with average earnings and uncorrelated with subsequent net population outflows. The authors attribute these findings to large spatial frictions, making it hard for agricultural families to adjust to aggregate structural change. Similarly, Alvarez-Cuadrado and Poschke, 2011 find that labor pull factors appear to dominate structural change, rather than labor push factors (which would include farm mechanization).

By employing empirical techniques to isolate the causal impact of the tractor on individual labor market outcomes and local transition dynamics, my work advances both the economic history literature as well as the growing macroeconomic literature studying technological change. There is a long economic history literature surrounding the diffusion, adoption, and subsequent impact of the tractor in the twentieth century United States economy. For example, Steckel and White, 2012 find that as of 1954 the tractor contributed an additional 8 percent to growth in the overall US economy relative to a counterfactual economy dependent on 1910 levels of technology. Olmstead and Rhode, 2001 find that the adoption of the tractor resulted in reduced labor hours of around 20 percent between 1944 and 1959. The authors also cite the USDA's calculations showing that the tractor was responsible for an 8 percent decline in labor requirements in agriculture by 1944. My work builds off of the work of Olmstead and Rhode, 2000 and Steckel and White, 2012 by adding individual-level outcomes data to study the impact of technological change on agricultural workers and their children. By looking at individual-level outcomes, I am able to measure migration and long-run employment and educational attainment for those exposed to tractors. My work largely confirms the results of the literature; I find reductions in employment for incumbent agricultural workers. However, I also find increased probability of agricultural work for farm youth exposed to tractors. Thus, I do not see a pure movement off of the farm and out of agricultural work entirely.

To expand on the tractor adoption literature, I exploit increasingly good methods for linking historical US Census records (Abramitzky, Mill, and Pérez, 2020; Acemoglu and Restrepo, 2020; Bailey et al., 2017). I link historical US census records across decades, a task that is computationally challenging and still hotly debated in the literature (Bailey et al., 2017). With the individual-level linked data, I observe the adjustments of individual farmers and their children as the aggregate US labor market shifted away from agriculture and into manufacturing. I construct multiple datasets of linked individuals between 1900 and 1940. My method of linking preserves a measure of link quality that allows researchers to test for potential bias in my research design associated with link quality. Once constructed, the linked historical census offers the ability to track an individual over multiple decades, as they respond to changing labor markets. In this way, my work is very similar to recent work by Feigenbaum and Gross, 2020 who study the fate of telephone operators who experienced the rise of mechanical switching technology between 1920 and 1940. The authors employ linked historical Census data to show that incumbent telephone operators were likely to experience either disemployment or lower-paying employment a decade after technological adoption. For future cohorts of workers, growing job opportunities meant that the decline in demand for telephone operators did not impact overall employment. These results mirror my own; incumbent workers face a large cost to technological change, but future cohorts find opportunities.

An even longer literature describes the adoption of the tractor in US agriculture. Much of the literature circles around the question of why the tractor took so long to be widely adopted. Seminal work by Clarke, 1991, and subsequent work by Fishback, 2017 find that tractor adoption accelerated rapidly in the 1930s under the New Deal. Many researchers suggest that borrowing constraints and the high cost of farm capital were barriers to tractor adoption (Rajan and Ramcharan, 2015; Olmstead and Rhode, 2000). Other researchers such as Manuelli and Seshadri, 2014 suggest that tractor adoption occurred slowly due to the slow incremental improvements in tractor technology. Along these lines, Gross, 2018 shows that the tractor was first adopted onto farms that primarily grew wheat, because early tractors were not as useful on row crops such as corn due to the specific tasks involved. As will be discussed in more depth below, I use the finding from Gross, 2018 as motivation for my instrumental variable approach, which is designed to purge the data of variation in tractor adoption that is correlated with farm productivity.

My study overcomes a number of key obstacles in the literature studying technological change. First, a key challenge to studying the impact of technological change on individual labor market participants is that we have few natural experiments that isolate a change in the production function of a firm that is orthogonal to other key firm characteristics. Technological adoption by a firm frequently corresponds with the initial productivity of the firm. Therefore, studying technological change may be confounded by the fact that only the best firms adopt a technology in the first place. A second challenge to studying technological change in the United States is that while agricultural employment was declining, employment in other sectors was growing rapidly. Figure 2.1 illustrates the structural shift taking place in the United States over this period. Aggregate changes in the economy are a result of both ‘pull factors’ (i.e. increasing demand in some sectors, such as manufacturing or services) and ‘push factors’ (i.e. declining demand in another sector, such as agriculture). In fact, Alvarez-Cuadrado and Poschke, 2011 argue that labor pull factors dominated push factors in causing internal migration in the US over the period. A third concern is that much technological change in the recent US history has been developed (either through private or public investment) with the direct purpose of lessening costs for employers via displacing labor, enhancing labor productivity, or enabling global outsourcing. If technological change endogenously responds to factor prices and productive capacities, then studying the causal impact of technological change requires an innovation that is useful only for specific tasks, specific places, or specific people for reasons unrelated to their productive capacities and costs.

To address these challenges, I isolate the push factors by carefully controlling for observable labor market conditions that might threaten identification. Because some omitted variables may be unobservable, I exploit an engineering constraint in the development of the tractor to predict where, geographically, tractors were adopted. Early tractors rode too low to the ground to be used on row crops such as corn, cotton, and potatoes. Instead, they were most useful on small grains (Gross, 2018). By harnessing geographic variation in the pre-tractor crop mix, I am able to isolate variation in tractor adoption that is not correlated with labor productivity or labor costs. Finally, I check the robustness of my results using a within-county design that exploits variation in an individual's farm residency status to estimate the causal effect of tractor adoption on workers' lives. By looking at farm-residing individuals relative to peer individuals in the same county, I am able to control for county-level changes that may go unobserved but correlate with tractor adoption. By using a within-county comparison of individuals, the empirical design looks similar to a difference-in-difference model in which the first difference is between individuals within a county by whether or not they live on a farm in the pre-tractor period and the second difference compares counties of the US based on which crops they grew most intensely before the rapid diffusion of tractors. This identification strategy relies on the assumption that crop varieties were not evolving differently in terms of labor demand for reasons other than the shock that I want to study.

With my research design, I isolate the impact of the tractor as a newly adopted tool used to increase the productive capacity of farms in the first half of the 20th century. I develop a model of technological change and spatial sorting in the context of a developing economy to act as a framework through which to understand my empirical results. By doing so, I contribute to the burgeoning literature on spatial structural change as well as the growing literature on technological change and automation (Eckert and Peters, 2017; Moll, Rachel, and Restrepo, 2021; Acemoglu and Restrepo, 2018a; Acemoglu and Restrepo, 2020; Acemoglu and Restrepo, 2018b).

There is a large and growing literature on automation and technological change in the present era. In particular, researchers look to the declining cost of capital and rise in automation technology to help explain recent growth in income and wealth inequality. Karabarbounis and Neiman, 2014 use country-industry variation to argue that decreases in the cost of capital have contributed largely to the declines in the labor share of income. In contrast, Elsby, Hobijn, and Şahin, 2013 find little evidence that changing investment costs have caused decreases in the labor share of income in the US, using time series aggregate and industry-level data from the US economy. Acemoglu and Restrepo, 2020 use industry variation to generate a shift-share empirical design applied to US commuting zones. They find that commuting zones exposed to industries with higher robot adoption also exhibit declining wages and employment. As with my empirical strategy, their strategy relies on geographical variation within the United States, which makes it difficult to estimate the aggregate impact of changing technology for the economy as a whole. Unlike my strategy, the authors are not able to observe migratory behavior, nor individual-level sectoral adjustments in response to changing local labor markets. In contrast, Graetz and Michaels, 2018 find no major disemployment effects of robot adoption using a task-based instrumental variable model and industry-country level data from Europe. Likewise, Aghion et al., 2020 find evidence of employment and wage *increases* using firm-level data from French firms. My paper contributes to this literature by providing new estimates of how individuals adjust to technological change using a shift-share instrument constructed over US counties. Relative to the current literature, my study contributes evidence of how individuals adjust to changes in both the incumbent generation and the generation of new labor market entrants.

Lastly, I provide a brief analysis of oral histories collected in the 1980s in Minnesota regarding the farm economy over the previous decades. These oral histories are provided by the Minnesota Historical Society and largely regard the changing landscape of agriculture over the early twentieth century. While not solely about technological change, the accounts demonstrate the importance of rapidly changing technology to the rapidly adjusting economy. The oral histories also shed light on the changing tasks associated with farming as the technological landscape changed.

In summary, this paper contributes to two major debates in the literature: 1) finding well-identified estimates of labor market responses to technological change, and 2) shedding light on long-run outcomes and margins of adjustment for individuals facing a rapidly changing labor market. The question of how and if individuals across the economy respond - for example via migrating or changing sectors - to large and long-run declines in industry-specific labor demand remains under-researched. Understanding the costs and benefits of technological change and the incidence of those costs and benefits across a variety of demographics will be key to mitigating welfare losses and maximizing productive capacity as technological change continues to force rapid changes in the global economy.

The rest of this paper proceeds as follows. In Section 4.3 I describe the data employed. Section 2.3 describes the historical context, including the rationale for my choice of instrument. In Section 2.4, I discuss my empirical strategy and layout the empirical model I employ. Section 2.5 discusses the results individual-level analysis. In Section 2.6 I provide narrative evidence from Minnesota farmers. In Section 4.5 I conclude.

2.2 Data

I restrict my study to the Midwestern states of the United States of America. These include Michigan, Ohio, Indiana, Illinois, Iowa, Wisconsin, Missouri, Kansas, North Dakota and South Dakota. I make this restriction for two reasons. The first is that this region saw the largest increase in tractors over the period, as shown in Figure 2.3. The second and more important reason is that by restricting to a single region, I isolate the impact of technological change in labor markets that are relatively uniform. In contrast to the Midwestern states, the Southern United States had a much different labor market dynamic, given their history of the enslavement of African and African American people. Likewise, the Western United States was experiencing rapid Western expansion, creating a very different economic environment from the rest of the country (Bazzi, Fiszbein, and Gebresilasse, 2020; Hornbeck and Naidu, 2014).

There are two main sources of data for this project. First, I use the 1900-1940 Agricultural Census for the US.[3] The data is at the county level and available for 1910, 1920, 1925, 1930 and 1940. These datasets include detailed information about crop acreage by county in

[3]Downloaded from ICPSR: Haines, Michael, Fishback, Price, and Rhode, Paul. United States Agriculture Data, 1840 - 2012. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2018-08-20. https://doi.org/10.3886/ICPSR35206.v4

the US. They also include the number of tractors in use on farms by county starting in 1925. Because there were no tractor counts at the county level collected before 1925, I make the assumption that there were zero tractors on farms in 1910. This assumption is fairly reasonable given that Clarke, 1991 estimates that there were a total of 3,000 tractors in the entire US in 1910 - approximately 1 tractor per county.

Second, I use the Full Count US Population Censuses from 1900-1940.[4] These records are at the individual level and available each decade from 1900-1940. Importantly, the records do not include unique individual identifiers across decades. Thus, in order to follow an individual over the whole 1910-1940 period, I match records using reported names, birth-years and place-of-birth. My matching method is discussed further in Appendix A.[5] [6]

In my analysis, I use the US Population Census in two distinct ways: first, I aggregate to the county level and observe changes in local labor markets using these county aggregates. Second, I construct links between Census decades so that I can observe outcomes of individuals over multiple decades even if they migrate. These datasets are based on the same underlying data, but capture distinct sets of outcomes.

My method for linking across historical decades is described in Online Appendix. The method is based on code originally written and used by Ran Abramitzky, Leah Boustan, and Katherine Erikkson.[7] I adapted their code for computing string distances across census decades. I then construct a novel distance metric and preserve this metric to test for bias in the sample. In the Online Appendix, I show that my measure of link quality is slightly negatively correlated with either crop shares or tractor adoption. The link quality metric is such that a higher number indicates a worse match and a lower number indicates a better match. Thus a negative correlation indicates that my links are slightly higher quality in places with more treated crops (likewise with more tractors). This is an important test of the validity of my data and linking method. Because of the correlation between data quality and tractor adoption, my results may be biased for outcome variables such as migration which are correlated with link quality. Using between-group identification methods mitigates this problem, as discussed in Section 2.5.

A key feature of the linked census data is that it only includes individuals who are alive and present in the US in each of the linked decades, 1910-1940. This creates a key difference between the two datasets since the linked dataset does not capture change over time stemming from births, deaths, and immigration from outside the US. Thus the linked dataset is used to shed light on *individual* outcomes and margins of adjustment.

---

[4] Made available through the Minnesota Population Center: Minnesota Population Center and Ancestry.com. IPUMS Restricted Complete Count Data: Version 3.0, 1910-1940 Census Files. Minneapolis: University of Minnesota, 2019.

[5] An online appendix is retrievable at `https://www.dropbox.com/s/ewvvxr3dqfkj6js/Appendix_AgMechanizationUS.pdf?dl=0`.

[6] My code for matching is based on code generously provided by Ran Abramitzky, Leah Boustan and Katherine Erikkson. Similar code is available at `https://ranabr.people.stanford.edu/matching-codes`

[7] The code on which I base my linking can be found at the following URL: `https://ranabr.people.stanford.edu/matching-codes`

Finally, the technological innovations in the tractor and the snapshots of data that I use do not entirely line up. The first key innovation in tractors occurs in 1917 when tractors became widely available to farmers. However, I only begin to observe tractors by county in 1925. Thus, I look at the long changes between 1910 and 1930, or 1940 in outcomes when I run my empirical analysis. The 1900-1910 period is considered the pre-period.

Figure 2.2 demonstrates some basic trends in the linked census data. In Figure 2.2 I present binned scatter plots of 1930 economic outcomes by farm status and age in 1910. The top left panel presents the expectation that an individual will be a farm resident in 1930 conditional on farm residency status in 1910 and age. We see that farm residents under the age of 20 have a reduced probability of still living on a farm in 1930, relative to those older than age 20. The same pattern holds for the bottom right corner, which charts the probability of agricultural employment conditional on employment in 1930. The top right corner plots the probability of migrating out of one's 1910 county of residence by 1930. Error in the linking means that the baseline migration rates reported will be higher than the true migration rate. However, as long as error in linking is not correlated with farm status or age, the shapes of the curves will remain as plotted; only the level will be incorrect. Interestingly, the probability of migrating is highest for those around the age of 10 in 1910 for farm residents, but closer to 20 for non-farm residents. Finally, in the bottom left corner, we see the probability of employment by age and 1910 farm status. We see that farm and non-farm residents show very similar employment patterns by age, with farm residents showing slightly higher probability of employment over the entire life-cycle.

In addition to my two main data sources, I use a dataset of banks per county in 1920 from the Inter-university Consortium for Political and Social Research (ICPSR).[8]

2.3 History

Tractors have been the focus of a large body of research studying innovation and technological change (Gross, 2018; Manuelli and Seshadri, 2014; Olmstead and Rhode, 2000). However, relatively little is understood about how the tractor directly changed labor markets and individuals' lives.

In this section, I first lay out some general history of the tractor and its previously estimated impact on farm labor demand and the US economy generally. I show evidence for how the tractor changed the farm production function and contributed to an agricultural supply shock. I then discuss the specific history that motivates my instrumental variable approach to causal identification.

[8]Federal Deposit Insurance Corporation. Federal Deposit Insurance Corporation Data on Banks in the United States, 1920-1936. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1992-02-16. https://doi.org/10.3886/ICPSR00007.v1

General History of the Tractor

While steam tractors had existed in small numbers since the 1870s, it wasn't until 1917 that the first mass-produced tractor - the Fordson - was invented by Henry Ford, making the tractor much more widely available to farmers. The Fordson was adopted rapidly by farmers and praised highly for its capabilities (Wik, 1964). The Fordson could be used alongside already existing plows and grain combines, however it had a low clearance and a fixed-tread and thus could not be employed for cultivation of row crops such as corn, cotton or tobacco. The first general purpose tractor that could be used for cultivation of row crops - the International Harvester's Farmall Tractor - was invented in 1924 and widely adopted during the 1930s (Gross, 2018).

Clarke, 1991 estimates that there were a total of 3,000 tractors in the entire US in 1910 and by 1930, the Census of Agriculture reports approximately 1 million tractors in use around the US. This growth only really started in 1917 when tractors first became mass produced. The number of tractors in the US increased to approximately 1.5 million by 1940 as the tractor continued to become more versatile and cost effective for farmers (Gross, 2018). Figure 2.3 shows the uptake of tractors per 100 farm arcres all over the US by 1930, specifically highlighting that the most intense region of uptake was the Midwest. Figure 2.4 shows the distribution of tractor uptake over Midwest counties for each year that tractor data is reported in the agricultural censuses.

While farm mechanization truly accelerated in the mid- and late-20th century, studying the impact of the tractor after WWII is difficult given that during this period many other farm technologies proliferated including the combine, the corn-picker, and cotton-pickers (Olmstead and Rhode, 2001). Many of these technologies were actually made possible by the invention of the tractor which allowed machinery to be towed and powered more efficiently. Studying the first half of the twentieth century offers a unique window in which to study the early stages of technological change and its impact on individual lives and labor markets.

The decline in agricultural employment was well underway in the early 1900s, as manufacturing jobs increased and people moved into urban settings. Figure 2.1 shows the decline in agricultural employment as a percent of total employment in the US alongside the rise of manufacturing. However, how much of this decline in agricultural employment is attributable to the rise of tractors and other farm mechanization remains disputed (Alvarez-Cuadrado and Poschke, 2011; Eckert and Peters, 2017).

The impact of the tractor on the US economy and labor markets has been estimated to be relatively large by authors previously. Steckel and White, 2012 estimate that mechanization in agriculture increased GDP by more than 8 percent over the first half of the 20th century. Olmstead and Rhode, 2001 discuss the labor-saving feature of the tractor, citing that "USDA authorities estimated that in 1944 the tractor saved on net roughly 940 million man-hours in field operations and 760 million man-hours in caring for draft animals relative to the 1917-21 period. The combined savings of 1.7 billion man-hours per year, represented about 8 percent of the total agricultural labor requirements in 1944, and translates into about 850 thousand workers."

As early as 1931, researchers considered whether agricultural mechanization was posing a threat to farm labor. Figure 2.5 shows the first page of a Labor Review article that discusses whether labor displacement was resulting from the adoption of the tractor.

Crop Variety as Instrument for Tractor Uptake

In this paper, I estimate the extent of substitution away from labor caused by the tractor using geographical variation at the county level. To attain a causal estimate of the substitution away from labor due to technological advancements in capital, I use an instrumental variable approach based on the early years of tractor diffusion - between 1910 and 1940 - which offers a unique opportunity to study the impact of technological change because of the incremental developments in tractor capabilities (Gross, 2018; Manuelli and Seshadri, 2014; Martini and Silberberg, 2006).

In the late 1910s and early 1920s the tractor was large and had a fixed tread so that it was primarily useful on certain crops and certain terrains. Specifically, tractors were useful on small grain crops such as wheat, oats and hay (Gross, 2018; Martini and Silberberg, 2006). Tractors were less effective in farming of row crops such as corn, potatoes, cotton and tobacco which required more intensive cultivation practices.

The differences between farming small grains and farming row crops stem from a difference in the tasks required. In general, farming requires four main tasks: plowing the fields, planting crops, cultivating (taking care of the growing crops), and harvesting (or 'reaping' as it is often termed for grains). Figure 2.6 lists these four tasks and how the early tractor was or wasn't useful on each task for the two major crop types. Plowing could be done by early tractors for any type of crop. However, while tractors were useful for reaping small grains, tractors could not be used for cultivating and harvesting row crops such as corn, cotton and tobacco. Contemporaries of the early tractor found that fixed-tread tractors, such as the Fordson, could be used for 77 % of farm tasks for producing small grains, but only 38 % for row crops (Gross, 2018; Gunlogson, 1922).

Given these differences in tasks for which early tractors were useful, tractors were more economically viable for small grain farmers and thus adopted most readily by farmers of these small grains (Gross, 2018; Martini and Silberberg, 2006). It wasn't until 1924 that the International Harvester's Farmall tractor was introduced as the first tractor that was truly general purpose in that it was agile enough to cultivate on fields that had actively growing row-crops and it could power other machines in tow (Gross, 2018; Martini and Silberberg, 2006; Olmstead and Rhode, 2001). By 1932, a tractor with rubber tires had been invented which made the tractor much more agile and no longer limited in scope.

Based on the incremental technological development of the tractor, I employ variation in county crop mix as an instrument to study the impact of tractor adoption on individual and regional labor market outcomes.

Figure 2.7 shows the map of Midwest counties where the color intensity corresponds to the proportion of farm acres in the county that were devoted to small grain crops (hay, wheat, and rye). This variation is what I will use to identify the causal impact of the tractor on agricultural labor market participants.

2.4 Empirical Strategy

Capturing the impact of technological change on individual labor market outcomes poses a unique set of challenges. Technology uptake is correlated with a variety of characteristics of farms and counties that may influence other long-run outcomes. Omitted variables due to these factors could dramatically bias the estimated results. To measure the causal impact of tractor technology on labor outcomes for impacted farmers, I employ a variety of controls along with an instrumental variable model that predicts the geographical proliferation of the tractor based on the local crop varieties. In this section I describe the challenges to identification in the study, and my strategy to address these challenges.

To study the impact of tractor adoption on individuals' medium- to long-run outcomes, I use county-level geographical variation in tractor usage across the United States. To test for the causal impact of tractors on agricultural labor demand, labor market composition and individual outcomes, I must isolate variation in tractor diffusion that is orthogonal to other factors that may also impact local labor market changes and individuals' long-run labor market outcomes. My approach is two-fold. First, I control directly for observable differences in counties that may be correlated with both tractor diffusion and labor market outcomes. Second, I instrument for tractor uptake using the county-level crop mix before the rapid diffusion of the tractor.

County-Level Controls

To examine the pre-tractor characteristics of the county-level data, I perform a balance test, reported in Table 2.1. Table 2.1 presents the results of running a simple OLS regression of 1910 county characteristics on tractor adoption. The regression takes the form

$$y_{c,1910} = \beta_1 Tractors_c + \sigma_s + \epsilon_c \tag{2.1}$$

where in columns (2) and (3), the tractor variable is given by the number of tractors per 100 farm acres as of 1930 and in columns (4) and (5), the tractor variable is predicted using the crop-share instrument in a simple 2-stage least squares regression. The regressions do not include any controls but do include a state fixed-effect. We see that in both levels and 1900-1910 changes, the OLS model has a statistically significant correlation between tractor adoption and pre-period county characteristics. The IV model does a bit better, but there is still a statistically significant relationship between 1910 county characteristics and tractor adoption as predicted by the share of total farm acres that is allocated to small-grain production. I discuss potential omitted variable bias in this section and include all of the listed county characteristics in the balance table as controls in my formal analysis.

There are a number of observable factors related to tractor diffusion that one may be concerned would confound my results. First, the adoption of the tractors might be correlated geographically with urbanization rates, manufacturing centers, or proximity to large urban areas. Because tractors were a new technology and produced mainly in assembly lines pioneered by Henry Ford, there is strong reason to believe that tractors were adopted most readily in places nearby manufacturing centers. Because demand for manufacturing labor and urban centers acted as pull factors for laborers, I control directly for these factors. First, I include the percent of the county population reporting as "urban" on the US Population Census in pre-tractor period (1910). Second, I control for manufacturing employment rates in the county from the 1910 US Population Census. Finally, I control for the geographical proximity to urban centers using simple Cartesian distances between counties and defining urban centers as counties with over 90,000 residents as of the 1920 US Population Census.

A further concern might be that tractor uptake would be greater in areas with relatively more access to credit. In fact, Rajan and Ramcharan, 2015 find this to be the case. Because access to credit and local bank activity may be correlated with other economic developments and possibly labor market changes, I control for this factor in two ways. First, I control for the number of banks in each county as of 1920 (the earliest date at which this data is available). Second, because the 1910-1920 period in US farming exhibited a large boom and bust cycle - large land price increases over the WWI period and sharp declines starting in 1920 - I control for the change in agricultural land values in 1910. The boom and bust is depicted in wholesale crop prices in Figure 2.8, where I've highlighted the small grain crops I use for identification in red.

Another large literature explores the fact that tractors may have been most cost effective and thus most readily adopted in places with already declining labor supply. For example, Hornbeck and Naidu, 2014 find increases in tractor adoption after the Great Mississippi Flood caused out-migration of African Americans from the Mississippi Delta. To address this issue, I control for population density in 1910 as well as the change in the county population between 1900 and 1910.

For each of the 1910 county characteristics I control for, I also include the 1900-1910 change. This is because, as is in the balance table, Table 2.1, there are significant trend differences in counties that adopted tractors.

Instrumenting for Tractor Adoption

Even after controlling for each of these important and potentially confounding factors, it remains highly likely that unobservable differences in productivity between farms explains a large proportion of the variation in tractor adoption. In particular, it is likely that highly productive farms were more likely to adopt tractors because of better foresight about technology or better financing. These highly productive farms would also be more likely to hire more workers, thus biasing estimates of changing labor demand positively, i.e. it may look as though technological adoption correlates positively with labor demand more than it truly does. This fact motivates the use of an instrumental variable that captures differences over space in where tractors were adopted onto farms that is uncorrelated with farm productivity differences. The development and adoption of the tractors offers a unique setting in which we can exploit features of the early tractor technological development to isolate the causal impact of farm mechanization on agricultural labor demand.

Perhaps even more important, by exploiting a random technological constraint faced by early tractors, I capture variation in tractor adoption that is not correlated with labor-capital dynamics. Much technological development is to induce more productive workers, or displace workers who demand higher wages. With my empirical technique, I estimate a true causal relationship between technological adoption and subsequent economic outcomes. This may have the effect of showing that technology is better for workers and labor markets than the modern literature suggests, for it may be the case that contemporary technological advancement is designed specifically to reduce labor costs via wages or labor displacement. In my study, the variation I use to predict tractor adoption avoids such dynamics, thereby highlighting the causal impact of technological change.

I use the fact that early tractors were most used on tasks related to small grain crops such as wheat, oats and hay. Small grain crops, in particular saw a change in which tasks could be accomplished using a tractor rather than power from either humans or livestock. The specific features of small grain crops that explain why they were most suitable for tractors are explained in detail in Section 2.3 and summarized in Figure 2.6.

In my empirical model, I use the share of farm acres that were used to cultivate small grain crops such as wheat, oats and hay as an instrument for tractor adoption. In order to avoid potentially endogenous crop choice, I use crop shares as of 1910, 7 years before tractors became mass produced and widely adopted.

Even when using the 1910 crop shares, I get a strong first stage. Figure 2.9 illustrates the strong correlation between the percent of a county's farm acres used to grow wheat, oats, or hay, and the increase in the number of tractors per acre in the 1910-1930 period and the 1910-1940 period. A 10 percentage point increase in the percent of farm acres in a county that are used to grow small grains corresponds to an increase in tractors per 100 acres in the county of 0.026 - 0.040, respectively. This result is highly statistically significant even with a full set of controls, state fixed effects and standard errors clustered at the state level. For the remainder of the paper, I will report the first stage F-statistic as well as a J-statistic for the over-identification test when using each crop share separately as a set of instruments.

2.5 Linked Individual Analysis

In this section, I employ linked historical census data to examine the medium- and long-run outcomes of farm residents and farm workers who were exposed to tractor adoption in the early 20th century. Using the linked data, I am able to observe individual level adjustment patterns, as well as draw conclusions regarding the interaction of local labor markets in response to technological shocks.

Empirical Model

The main empirical specification I use to study the impact of local tractor adoption on individual outcomes and adjustments is a simple IV regression, presented in equation 2.2. This regression uses the percent of a county's farm acres devoted to growing wheat, rye, and hay in 1910 to predict the number of tractors present in the county in year t. Using 1910 as a base year, before tractors were widely available, I then look at time t individual outcomes, $y_{i,c,t}$, for individual i who was living in county c in 1910. I include state fixed effects as well as a vector of controls that are both individual and county specific, but do not vary over time. Controls used include an individual's reported race, their literacy status, a quadratic for the individual's age, their homeownership status, and farm status. County level controls include the number of banks present in the county in 1920 to approximate credit availability, the distance to the closest large metropolitan area, and a series of 1910 county characteristics such as the average farm size, share in agricultural employment, share in manufacturing employment, average farm value, population density, urban share, and 1900-1910 changes for those characteristics. These controls are intended to absorb observable economic variation that may correlate with tractor adoption and individual economic outcomes. I discuss these controls in depth in section 2.4.

$$y_{i,c,t} = \beta_1 \hat{tractors}_{c,t} + \beta_2 X_{i,c} + \mu_s + \epsilon_{i,c,t} \tag{2.2}$$

I employ the 1910 percent of county farm acres in each of hay, wheat, and rye separately as three instruments to predict tractor adoption. The instrument is intended to harness as-good-as-randomly assigned variation in tractor adoption using the fact that tractors were, at first, most useful on small grain crops. The validity of the instrument and its purpose are discussed at length in section 2.4. In particular, I use this instrument to harness variation in tractor adoption that is not due to the unobservable characteristics of adopting farmers such as their latent productivity or business prowess. I compute the IV regressions using GMM estimation and report the F-statistic and J-statistic for each model to examine the strength of the instrument, as well as a test of over-identification. In all results reported, I cluster standard errors at the county level, to account for spatial correlation in the data.

Within-County Empirical Model

My most rigorous empirical strategy harnesses within-county treatment effects using demographic variation in exposure. I consider this my most robust model because I am able to control for any county-level trends that correlate with both individual outcomes tractor adoption using a county fixed effect. This empirical model is based on work done by other economic historians such as Aneja and Avenancio-Leon, 2019 and Alsan and Wanamaker, 2018.

To estimate the causal impact of the tractor on impacted farm residents and farm workers, I measure the outcomes of farm residents relative to the non-farm residents within the same county at the time of tractor adoption. By observing the interaction of tractors and farm status, I measure the differential impact of tractors on farmers (whom we would assume to be most affected by tractor effects). The regression model I use is given by the equation

$$y_{i,c,t} = \beta_0 + \beta_1 \hat{tractors}_{c,t} + \beta_2 \hat{tractors}_{c,t} \times 1\{i \in G\} + \beta_3 1\{i \in G\} + \beta_4 X_{i,c,t} + \mu_c + \epsilon_{i,c,t} \quad (2.3)$$

where I is any group of individuals, such as farm residents, who do not constitute the whole of a county. In most cases I use either farm residents or agricultural workers as the category I and compare to like individuals outside that group. The coefficient β_1 is not identified because tractor adoption is only observed at the county-level and the model already includes county fixed effects, however β_2 is identified and represents the outcomes of impacted farm residents relative to non-farm resident neighbors (when I is farm resident status in 1910). By using this variation in pre-period farm-residency status, I look specifically at the impact of tractor adoption on those whose lives would be most affected by the tractor, compared to those whose lives were most likely not impacted directly by the tractor. The tractor may have had regional-level impacts that affect all residents of the county, regardless of pre-period farm residency status. These region-level effects will not be captured by β_2. Instead, β_2 is designed to measure how farm-residents specifically were impacted by changing technology, net of any county-level changes that took place.

With each of these strategies in place, I estimate the causal impact of tractor adoption on medium- and long-run individual outcomes. In the following sections, I present results.

Main Results

First, I present the results of running the regression specified in equation 2.2 on subsamples of the linked census datasets. I report results separately for the 1910-1930 linked data and the 1910-1940 linked data. As stated above, the data is restricted in all instances to reported males in the Midwest of the United States.

1910-1930 Linked Sample

Figure 2.10 depicts the coefficient in the IV regression of county tractor adoption on employed males. The regression is run separately on males employed in agriculture in 1910 and males who report employment in a sector other than agriculture in 1910. We see that for those employed in agriculture in 1910, the introduction of the tractor between 1910 and 1930 results in a decrease in employment. The decline in employment is accompanied by a similarly sized increase in the "Not in the Labor Force" (NILF) group. We do not see the same movement among those employed in non-agricultural sectors. However, we do see an overall decline in the likelihood that an individual is employed in agriculture in 1930 in response to local tractor adoption regardless of the individual's starting sector. Non-agricultural workers are slightly more likely to work in manufacturing when exposed to local tractor adoption. Figure 2.11 shows that when using farm residents rather than agricultural employment, we see the same pattern among adult males. However, in contrast, when we restrict the sample to males under the age of 18 in 1910, we see a very different pattern. In this case, we see that youth exposed to tractors report greater employment rates regardless of whether or not they lived on a farm in 1910. Those who lived on a farm in 1910 experience a greatly increased probability of working in agriculture in 1930, while those who did not reside on a farm in 1910 see reduced likelihood of working on a farm in 1930. Neither group has increased likelihood of manufacturing employment due to tractor exposure.

Table 2.2 reports the results of the within-county empirical specification. We see that the results remain largely the same. Those employed in agriculture in 1910 experience greatly reduced employment and increased NILF status. Agricultural workers who were exposed to tractors are less likely to go into manufacturing employment than those less exposed.

For those under the age of 18, farm residents are much more likely to be employed in agriculture in 1930 when exposed to tractors and show no other strong outcomes relative to those less exposed. This suggests that for those new entrants to the labor market, farming was still a viable employment option in 1930 and one that young farm residents largely took up. In fact, this evidence suggests that while older farmers may have lost the ability to maintain employment with the introduction of tractors, youth were actually more likely to stay in farming when exposed to tractors, counter to common narratives of technological change and movement off of farms.

It is worth noting that, while the first stage F-statistics are quite high, the J-statistics show that in most cases the model is highly over-identified, suggesting that there may be heterogeneous treatment effects or that the instrument is not valid.

1910-1940 Linked Sample

The same analysis for the 1910-1940 linked census data shows a slightly different story. Figure 2.12 shows that for individuals captured in the 1910-1940 linked dataset, the employment results agricultural employees are muted relative to their strength in the 1910-1930 linked data. There is still a decrease in employment for farmers exposed to tractors, but far less strong than for the 1910-1930 links. This may be because the 1910-1940 linked data does not included those who died between 1930 and 1940. It could also be the case that over the ten additional years, much of the geographical distinctions smoothed out so that there is less of a strong and distinct effect of tractors between locations over a 30 year time horizon than there was over a 20 year horizon.

Additionally, whereas in the 1910-1930 links, those employed in agriculture in 1910 were much less likely to be employed in agriculture in 1930, with the 1910-1940 links, we see no reduced employment in agriculture for those exposed to tractors. On the other hand, for 1910 employed workers in non-agricultural sectors, we see a greatly increased likelihood of employment in manufacturing when exposed to tractors. The left panel of Figure 2.13 shows that the results for male farm residents older than 18 show much the same pattern - small declines in employment that do not seem to load on agricultural employment.

The right panel of Figure 2.13 shows that for those under the age of 18 in 1910, tractors induce increases in employment. For non-farm residents this also means a decrease in the unemployment rate in 1940 and large increases in the probability of manufacturing employment. For farm residents, increases in employment are attributable primarily to increased probability of agricultural employment in 1940.

Table 2.3 displays the within-county results for the 1910-1940 linked data. As before, these results align with the between-county results. In particular, for those employed in agriculture in 1910, we see small declines in employment and increases in NILF status. These workers are more likely to be employed in agriculture and less likely to be employed in manufacturing. For youth growing up on farms in 1910, tractor exposure makes them much more likely to find agricultural employment in 1940. There are no significant employment effects.

In the 1940 data, we can observe education and income, though the data is far from perfect. Taking the data for what it is, we measure education and income outcomes for 1940 and report results in Table 2.4, along with the migration results.

Whereas in the 1910-1930 linked data, we didn't see much in the way of migration differences for farmer and agricultural workers, in the 1910-1940 data, those younger than 18 in 1910 appear less likely to move when exposed to tractors in their home county. This result is somewhat surprising given the fact that the historical narrative seems to state that farm youth would leave home in search for new work during this period. In fact, it may be that in the whole population, young men were likely to move to find new work, but that the emergence of the tractor and subsequent economic growth actually compelled tractor-exposed young men to stay in their locales.

In terms of education, Table 2.4 shows that for adults exposed to tractors, there is little impact of tractor exposure on education rates. This is comforting given that we would not expect adult's educational choices and opportunities to be impacted by a technological change occurring after their schooling years. However, for those under the age of 18 in 1910, we see that education rates decline with exposure to tractors. Most likely this is also due to the fact that these youth had increased probability of staying in agricultural employment which required less formal schooling.

In terms of income, we observe that those employed in agriculture earn nearly \$ 200 less in 1940 wages for each additional tractor per 100 farm acres. In contrast, exposed 1910 agricultural worker are 7 % more likely to earn non-wage income greater than \$ 50. For farm youth, we find even worse wage income in 1940 along with better non-wage income when faced with a greater intensity of tractors adopted locally. These results may be a direct consequence of the fact that farm youth are more likely to stay in agricultural work when exposed to tractors and agricultural workers make lower wages and higher non-wage income on average (as seen by looking at the coefficient on farm resident status or agricultural employment).

Placebo Test Using 1900-1910 Links

To further examine the validity of these results, I present the results of the same regressions discussed previously, but using 1900-1910 linked data.

The results indicate that the data for 1900-1910 is noisier than the other datasets. Unfortunately, we see estimates of relatively high magnitudes for the impact of tractors on pre-period employment and labor force status. This is especially true for those participating in the labor force in 1900 who have significantly impacted unemployment rates, labor force status, and rates of employment in manufacturing.

For those younger than 18 in 1900, we see similar patterns though none statistically significant and most dampened in magnitude. Nonetheless, it is worth acknowledging that these pre-period trends exist in the data despite a full set of controls and state fixed effects.

Perhaps the most startling pre-trend is the highly elevated rates of migration for and departure from the labor force for non-farm residents in the pre-period. It is certainly possible that a changing demography leading up to tractor adoption accounts for some of my results.

2.6 Narrative Evidence

The Minnesota Historical Society holds a collection of oral histories on the farm economy.[9] The Historical Society describes the collection as "depictions of farm life, most of the reminiscences focus on early twentieth-century life," in addition to commentary on the changing economy in the second half of the twentieth century. In order to methodologically analyze the oral histories, I went through each document and extracted passages that contained the words "technology," "tractor," and "machinery" (and other words with the same roots). [10] Here, I discuss a subset of the extracted quotes to demonstrate how these narrative accounts support the empirical results already discussed.

First and foremost, the narrative accounts describe the rapidly evolving agricultural landscape and the economic challenges farmers faced. One woman described farmers under increasing economic stress and the "climate of silence" that fostered a sense of isolation for suffering farmers:

" ...there are a lot of them out there yet who will not talk about it. I have people come here to my house, crying. I have people still calling here, who would like to kill themselves, who have no food in the house... I talk to them about needing some help, about getting a farm advocate. But they are still very insistent that we don't tell anybody that they've been here, or that they've called me. There is still a great deal of–I don't know–maybe pride; that word has been used so much over the years–pride. I don't know if it's pride nearly as much as it is embarrassment. They're embarrassed because they look at their neighbors and it looks as if everybody else is going just wonderful. So part of this quietness and part of this keeping everything inside has made the story not get around." In addition to demonstrating the despair of struggling farmers, this quote also hints at

[9] The oral history collection can be accessed at `http://collections.mnhs.org/cms/display?irn=10469048`.

[10] The extracted paragraphs are available to be read in full at this link: `https://www.dropbox.com/s/glehdjdz2vnbvmd/OralHistoryQuotes.xlsx?dl=0`.

the emotional difficulty of transitioning out of farming, the "pride" that kept people from adjusting, uprooting, or accessing additional help.

Another interviewee describes the decline in rural communities. "Small towns in rural communities are declining. One of our extension sociologists predicted that twenty-five years ago and got kicked around a little bit for it in meetings around the state. How far they'll decline in size, I don't know, but a lot of them are not going to grow much, because so many were established seven to ten miles apart. That was about the distance that a farmer could drive in one day with a team of horses and deliver whatever it was he was going to sell or buy. That's not true anymore. If they're looking for machinery repairs, it may be that there's only one dealer of the two or three main brands in a county. So they go fifty miles for parts if they need to. It's a little more expensive, but maybe that dealer can stock a broader range of parts. I think there are some of those kinds of changes taking place." In this quote, the interviewee suggests that the changing size of the communities has been enabled in part by the increased automobile access and more need for mechanical expertise.

Second, the narratives confirm the increasingly important role of technological change and farm machinery in the agricultural production function. One interviewee discusses how technology changes the need for family members to stay working on the farm. "The technology part is changing in that there used to be a need for more family members involved, and now there's less need for a lot of people involved with it. But, on the other hand, it's maybe easier for other family members to help operate the farm because things are automated. You don't have to evaluate the family members by how much muscle they have ... there's a lot of other ways of valuing people besides that. I think that's the case now, whether it's intelligence or ability to run a computer or skills that you can market off the farm." This quote demonstrates not only the decreasing need for some family members on the farm, but also the changing skills required to operate a farm. These changing skills meant that the younger generation may have had more prowess in running a farm relative to the older generation as the technology evolved. This idea is confirmed by another interview: "The young farmer isn't a better farmer then dad, but he is a lot better at playing the game. When all of a sudden I asked dad for a cash flow and this and that document, I just paralyzed him. Some dads had a problem with trying to figure out why something that had worked, his dad did it that way, I did it this way, why something that had worked for 20 years wasn't working anymore. And the younger farmer who, maybe, went to some ag school, he's probably got a marketing advisor, he probably takes soil samples, knows how to play the pick and roll and all the games that have to be played." This interviewee confirms the hypothesis that the second generation of farmers has learned new skills adept to the new tasks on the farm.

Another farmer interviewed discusses his memory of the time before modern farm technology. "I'm not that old. I mean, this isn't that far back, but I remember as a small child, we were putting up hay on the farm, yet, with a hay loader and some of that. It only happened for a year or two before we switched to baling and stuff, but I still remember that as a small child, even on our own farm. Then we did a lot of farming with relatives and neighbors who were even just a few years behind that. I certainly, at the time, didn't appreciate it as I do now, especially a few years later when it got to be more work, and I was old enough to do work...But looking back, I'm really glad that I had that, because I have not really run into anyone my own age who has the glimpse of that, not just in haying, but in other examples. We did putting up silage, for example, and thrashing, even until I was into high school. So I had a chance of working with all of those older technologies... My dad...just tinkered with stuff and came up with a lot of homemade implements and stuff which I think really were innovative and creative...I think it's changed a lot today because there are fewer farmers that have the time to do that, plus the costs get higher." In this case, the interviewee laments the days in which farming involved more creative and small-scale solutions to improving production. Again, we see how the changing technologies changed the landscape of farm economies and the skills and tasks needed to provide for a family in farming.

The full oral history collection gives a much broader set of commentary on changing farm life. In particular, the narratives describe the changing size of farms in terms of acreage, the shift towards corporate farming in the second half of the twentieth century, and the central role of lending and credit as part of both the expansion of farmers into new technologies as well as the financial strain faced by farmers in bad times.

The oral histories demonstrate that technological change uprooted many of the practices of agricultural life. Increased farm equipment corresponded with increasingly large farms and a departure from family farming. The technology also created many new tasks for those in the agricultural community. These new tasks were frequently taken on by the "younger generation" who were more likely to adapt to the changing environment. Nonetheless, farmers struggled with the changing technology and decreasing incomes as the economy evolved.

2.7 Conclusion

In this paper, I have developed a method to study an episode of technological change. I use a key historical fact regarding the technological improvement of the tractor over time, making the tractor more useful on certain crops. With this identification strategy in hand, I produce a set of results showing a rich picture of how the development of a farm technology that changed production and reduced the need for labor impacted the farm population. My analysis and results are hindered by the fact that I use geographical variation and thus cannot fully see the aggregate impact of the tractor. I find strong evidence for tractor adoption leading to reduced employment for incumbent farm workers, though no increased likelihood of migration. In contrast, children of farm families that were exposed to tractors locally ended up more likely to work in agriculture as adults and less likely to move out of their county of origin. Total incomes in tractor-adopting counties appear to rise, but less so for farm workers than for other county residents, suggesting that tractor adoption most likely hurt farm wages.

The results of this paper suggest that while many farmers stayed on their farms and were not physically displaced, these farmers also appear to have lost significant income and employment opportunities. The younger generation appears to emerge unscathed and still able to find farm employment, however understanding how to support those whose skills and investments incidentally lose value is a key factor in improving welfare.

2.8 Tables and Figures

Table 2.1: Balance Table for Midwest Counties

MidWest Counties	Sample Mean (1)	OLS Levels (2)	Changes (3)	IV Levels (4)	Changes (5)
Population per 100 county acres	8.67	88.865***	11.234***	51.371***	5.990
		(5.306)	(1.262)	(14.706)	(3.391)
Farm residents per 100 county acres	2.52	5.221***	-0.861***	10.583***	-1.652***
		(0.279)	(0.101)	(0.728)	(0.266)
Urban Pop. per 100 county acres	4.31	75.919***	11.340***	33.580*	7.204*
		(4.964)	(1.214)	(13.551)	(3.273)
Ag. employment per 100 county acres	0.84	2.293***	-0.081*	3.863***	-0.426***
		(0.102)	(0.040)	(0.283)	(0.102)
Manu. employment per 100 county acres	0.49	8.114***	2.805***	3.425*	1.858**
		(0.495)	(0.203)	(1.369)	(0.571)
Farm Val/Acre, 1910	72.35	262.754***	0.000	440.580***	0.000
		(6.754)	(0.000)	(22.035)	(0.000)
Farm acres per 100 county acres	81.00	67.099***	-24.593***	328.329***	-85.674***
		(6.386)	(2.809)	(12.890)	(7.060)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table reports county characteristics as of 1910. Column (1) reports county means without any weighting scheme. Column (2) reports the regression coefficient on a regression of the respective county characteristic (in levels) on the number of tractors per 100 farm acres in 1930. The regression includes state fixed effects. Column (3) repeats column (2) but uses 1900-1910 change in county characteristics as the regressor. Columns (4) and (5) replicate columns (2) and (3), but report the estimated coefficient on the predicted number of tractors given the percent of farmland devoted to small grains in 1910.

Table 2.2: Within-County Regressions for 1910-1930 Linked Data

			Ag. employment vs. Other employment			
	(1)	(2)	(3)	(4)	(5)	(6)
	Emp.	NILF	Unemp.	Emp. = Ag.	Emp. = Manu.	Moved
Tractors * Ag.	-0.157***	0.190***	-0.00960	0.00976	-0.0230**	-0.00999
Emp.	(0.0202)	(0.0182)	(0.0126)	(0.0333)	(0.0106)	(0.0309)
Ag. Emp	0.0485***	-0.0360***	-0.0159***	0.196***	-0.00277	-0.0897***
	(0.00453)	(0.00379)	(0.00298)	(0.00738)	(0.00239)	(0.00748)
N	444974	444974	444974	438048	438048	444974
F	99.36	99.36	99.36	100.1	100.1	99.36
J	0.000240	0.000000809	0.429	0.0620	0.0000321	0.0000115
			Farm Residents vs. Non-farm Residents, Age < 18			
	(1)	(2)	(3)	(4)	(5)	(6)
	Emp.	NILF	Unemp.	Emp. = Ag.	Emp. = Manu.	Moved
Tractors * Farm	0.0804+	-0.0249	-0.0457	0.383***	0.00402	-0.0769
Res.	(0.0419)	(0.0198)	(0.0372)	(0.0840)	(0.0271)	(0.0688)
Farm Res.=1	0.000330	0.00175	-0.00448	0.0907***	-0.0163**	-0.0252
	(0.0105)	(0.00538)	(0.00916)	(0.0190)	(0.00682)	(0.0168)
N	45697	45697	45697	45083	45083	45697
F	78.89	78.89	78.89	79.07	79.07	78.89
J	0.145	0.488	0.164	0.0126	0.976	0.0213

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

This table presents the results of running the regression specified in equation 2.3. Standard errors are reported in parentheses with significance stars used to indicated significance at the 90, 95 and 99 percent levels, respectively. All regressions use clustered standard errors at the county level. The F statistic of the first stage regression is reported, as well as the J statistic to test for over-identification of the instrument.

Table 2.3: Within-County Regressions for 1910-1940 Linked Data, Part 1

	Ag. employment vs. Other employment				
	(1)	(2)	(3)	(4)	(5)
	Emp.	NILF	Unemp.	Emp. = Ag.	Emp. = Manu.
Tractors * Ag. Emp.	-0.0389***	0.0443***	-0.00199	0.0631***	-0.0552***
	(0.0131)	(0.0126)	(0.00727)	(0.0215)	(0.0138)
Ag. Emp	0.0197***	-0.0176***	-0.00304	0.175***	-0.0202***
	(0.00617)	(0.00569)	(0.00331)	(0.00870)	(0.00540)
N	233394	233394	233394	233394	233394
F	99.26	99.26	99.26	99.26	99.26
J	0.0102	0.0108	0.401	0.0187	0.0788
	Farm Residents vs. Non-farm Residents, Age < 18				
	(1)	(2)	(3)	(4)	(5)
	Emp.	NILF	Unemp.	Emp. = Ag.	Emp. = Manu.
Tractors * Farm Res.	-0.00659	0.00544	0.0000226	0.331***	-0.0990***
	(0.0104)	(0.00737)	(0.00679)	(0.0317)	(0.0164)
Farm Res.=1	0.0262***	-0.0109***	-0.0150***	0.108***	-0.0220***
	(0.00442)	(0.00305)	(0.00308)	(0.0116)	(0.00651)
N	256743	256743	256743	256743	256743
F	122.6	122.6	122.6	122.6	122.6
J	0.0935	0.174	0.421	0.0000230	0.162

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

This table presents the results of running the regression specified in equation 2.3. Standard errors are reported in parentheses with significance stars used to indicated significance at the 90, 95 and 99 percent levels, respectively. All regressions use clustered standard errors at the county level. The F statistic of the first stage regression is reported, as well as the J statistic to test for over-identification of the instrument.

Table 2.4: Within-County Regressions for 1910-1940 Linked Data, Part 2

	Ag. employment vs. Other employment			
	(1)	(2)	(3)	(4)
	Moved	Some HS	Wage Inc.	Non-Wage Inc. > 50
Tractors * Ag.	0.00414	-0.0272	-185.3***	0.0684***
Emp.	(0.0226)	(0.0191)	(36.25)	(0.0190)
Ag. Emp	-0.0744***	-0.107***	-209.7***	0.0445***
	(0.00966)	(0.00788)	(15.30)	(0.00825)
N	233394	233394	210916	224162
F	99.26	99.26	96.39	98.55
J	0.00122	0.823	0.175	0.0113
	Farm Residents vs. Non-farm Residents, Age < 18			
	(1)	(2)	(3)	(4)
	Moved	Some HS	Wage Inc.	Non-Wage Inc. > 50
Tractors * Farm	-0.0770***	-0.146***	-428.4***	0.161***
Res.	(0.0260)	(0.0266)	(47.40)	(0.0227)
Farm Res.=1	-0.0904***	-0.104***	-191.0***	0.0678***
	(0.0101)	(0.0106)	(17.63)	(0.00891)
N	256743	256743	240881	248651
F	122.6	122.6	120.3	121.5
J	2.54e-08	0.0197	0.210	0.0335

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

This table presents the results of running the regression specified in equation 2.3. Standard errors are reported in parentheses with significance stars used to indicated significance at the 90, 95 and 99 percent levels, respectively. All regressions use clustered standard errors at the county level. The F statistic of the first stage regression is reported, as well as the J statistic to test for over-identification of the instrument.

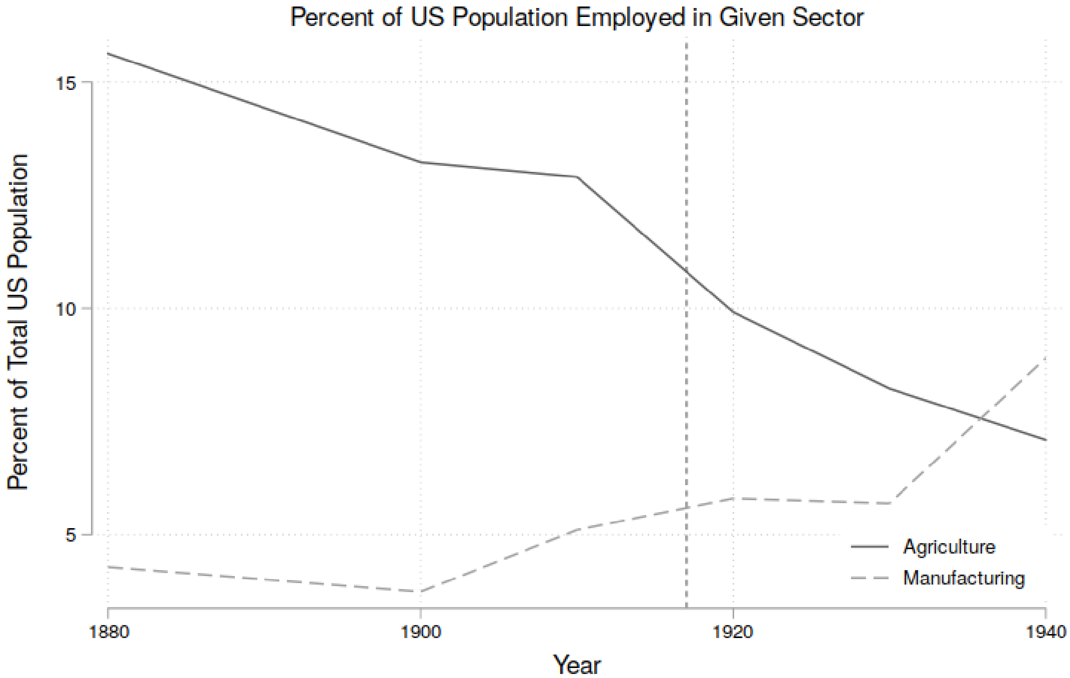


Figure 2.1: Percent of Total Population Employed in Each Sector (not all sectors plotted). Dotted vertical line marks 1917, when the first mass produced tractor became available.

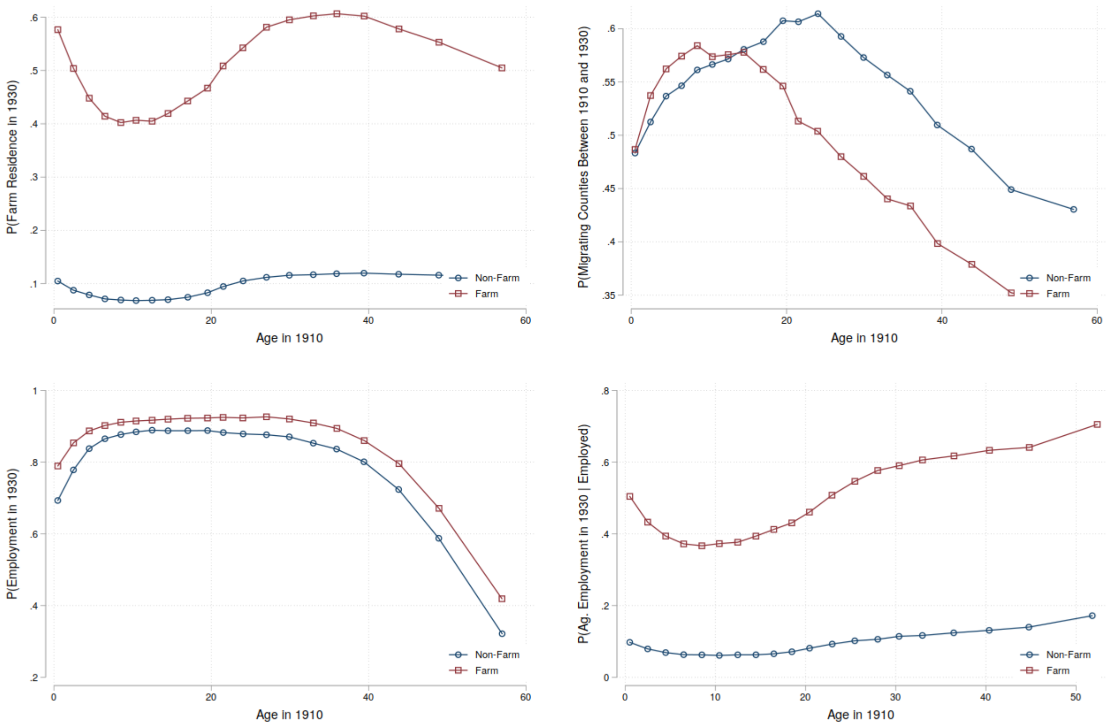


Figure 2.2: Basic trends in linked 1910-1930 census data

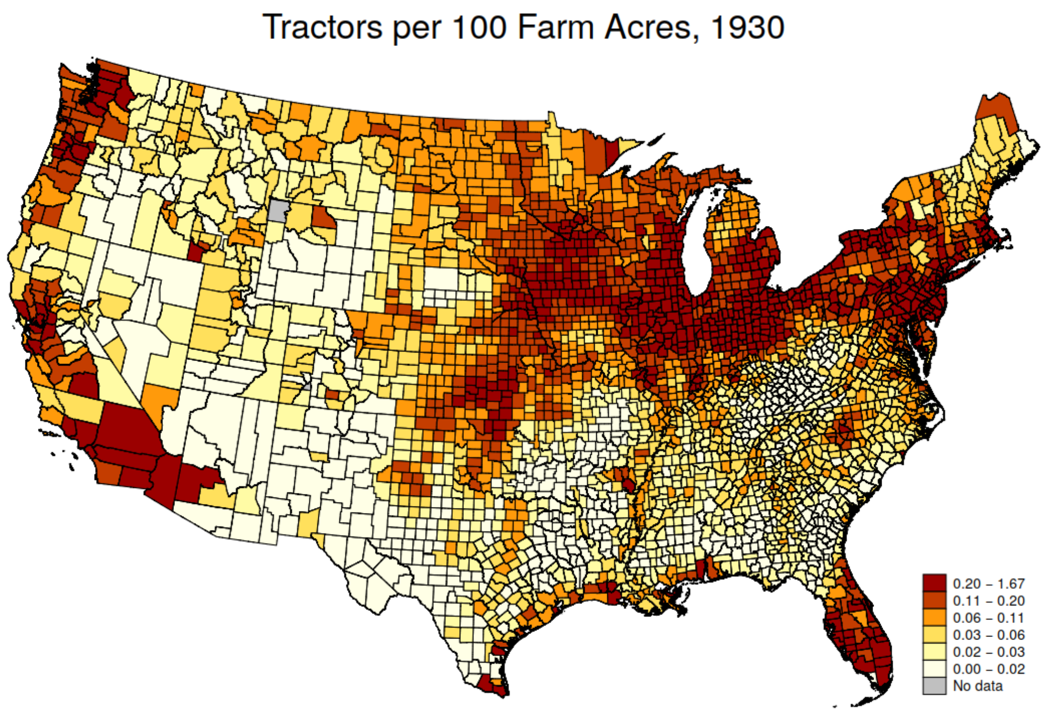


Figure 2.3: Tractors per 100 Farm Acres in 1930, by County

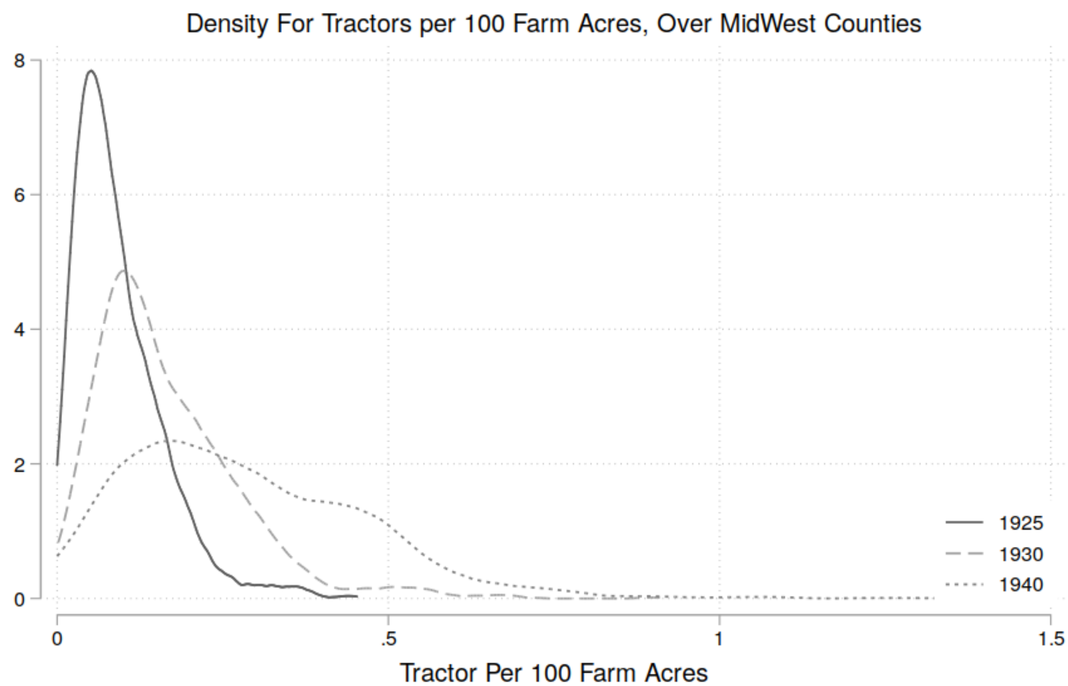
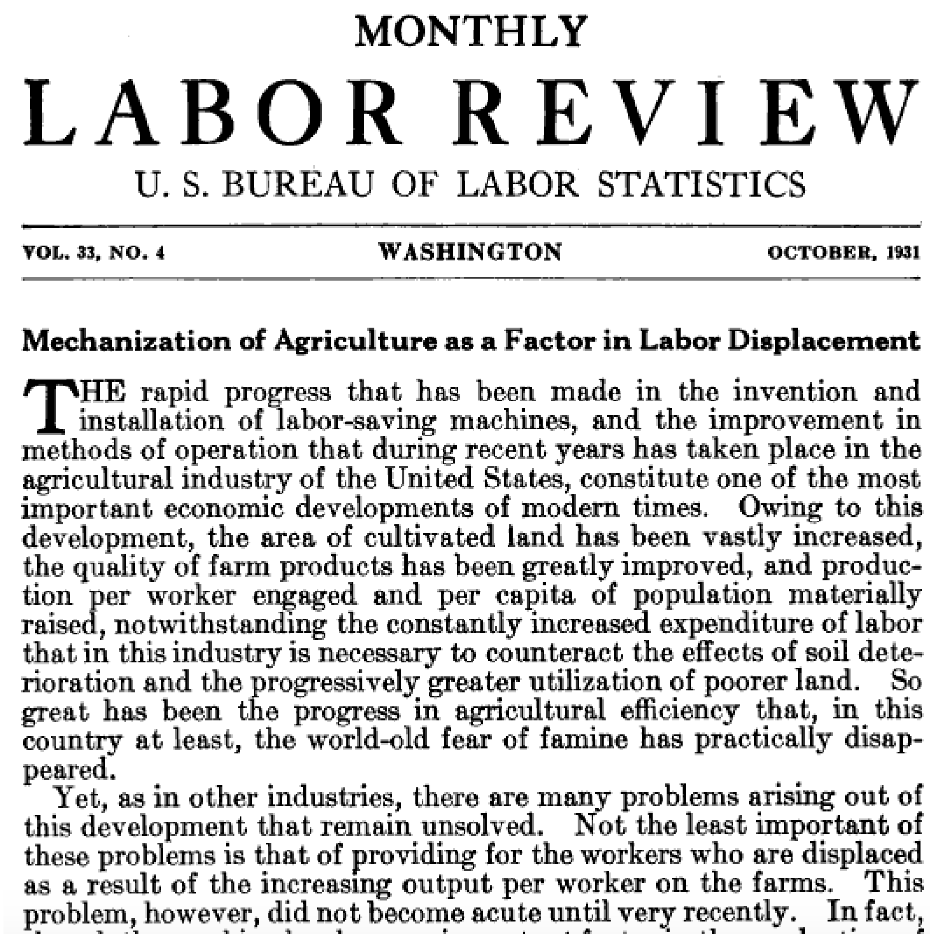


Figure 2.4: Density Chart for Midwest Region



MONTHLY
LABOR REVIEW
U. S. BUREAU OF LABOR STATISTICS

VOL. 33, NO. 4 WASHINGTON OCTOBER, 1931

Mechanization of Agriculture as a Factor in Labor Displacement

THE rapid progress that has been made in the invention and installation of labor-saving machines, and the improvement in methods of operation that during recent years has taken place in the agricultural industry of the United States, constitute one of the most important economic developments of modern times. Owing to this development, the area of cultivated land has been vastly increased, the quality of farm products has been greatly improved, and production per worker engaged and per capita of population materially raised, notwithstanding the constantly increased expenditure of labor that in this industry is necessary to counteract the effects of soil deterioration and the progressively greater utilization of poorer land. So great has been the progress in agricultural efficiency that, in this country at least, the world-old fear of famine has practically disappeared.

Yet, as in other industries, there are many problems arising out of this development that remain unsolved. Not the least important of these problems is that of providing for the workers who are displaced as a result of the increasing output per worker on the farms. This problem, however, did not become acute until very recently. In fact,

Figure 2.5: Labor Review Article from 1931 Discussing Labor Displacement and Farm Mechanization

	1917 - 1925		1925 onward	
	Small Grains	Row Crops	Small Grains	Row Crops
Plowing:	Tractor	Tractor	Tractor	Tractor
Planting:	Horse/Human	Horse/Human	Tractor	Tractor
Cultivating:	N/A	Horse/Human	N/A	Tractor
Reaping/ Harvesting:	Tractor	Horse/Human	Tractor	Horse/Human
Source: This information is aggregated from accounts found in Olmstead and Rhode (2001) and Martini and Silberberg (2006). Small grains include wheat, hay, and oats. Row crops include corn, cotton, tobacco, and potatoes. Although some tractors existed before 1917, it wasn't until 1917 that Henry Ford developed the Fordson tractor that was widely available to farmers around the United States.				

Figure 2.6: Farm tasks by crop type and technological needs

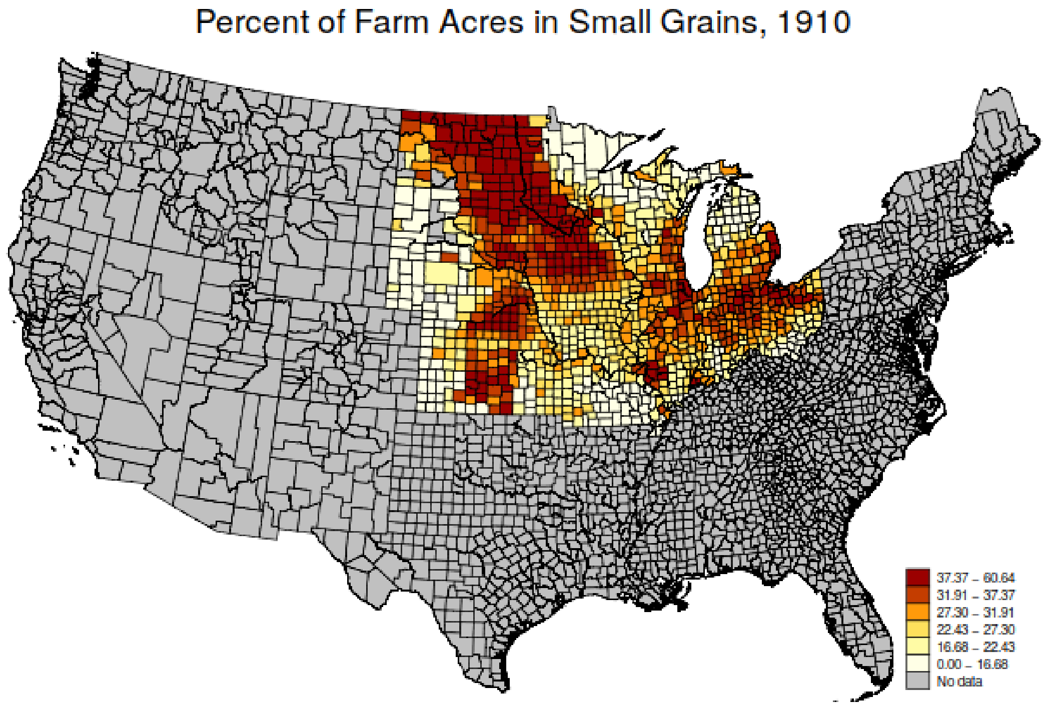


Figure 2.7: Map showing the variation used in the instrument for identifying tractor uptake.

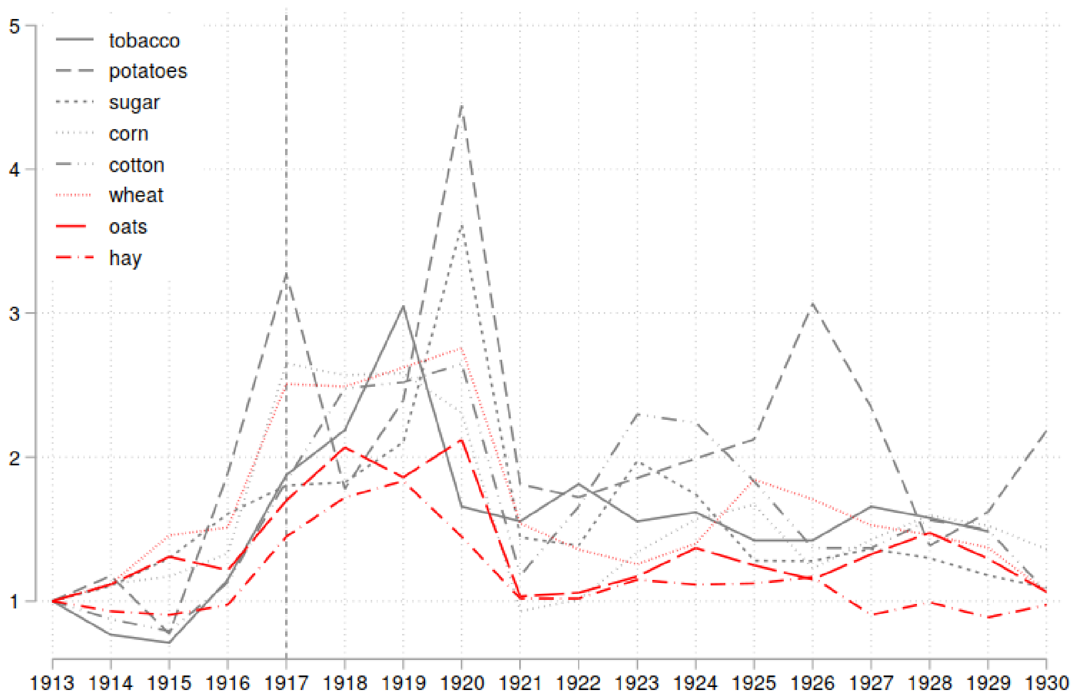


Figure 2.8: Wholesale US Crop Prices

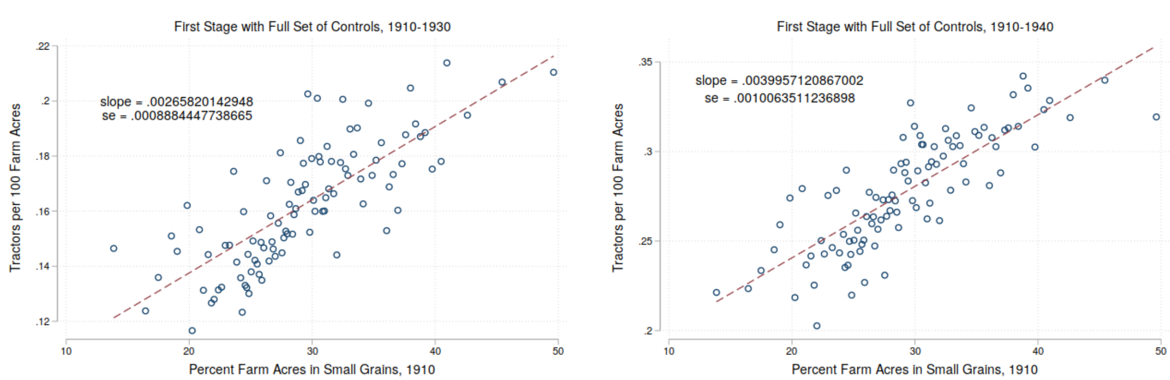


Figure 2.9: First Stage Regressions

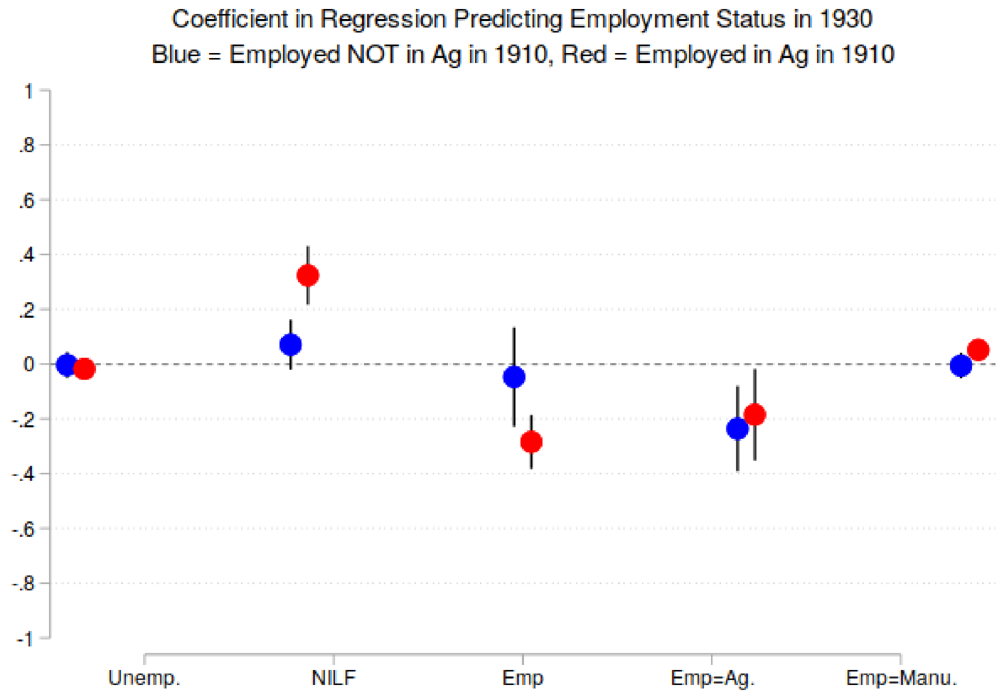


Figure 2.10: Coefficients on Tractor Adoption for Employed Males

Figure 2.11: Coefficients on Tractor Adoption for Farm and Non-Farm Residents

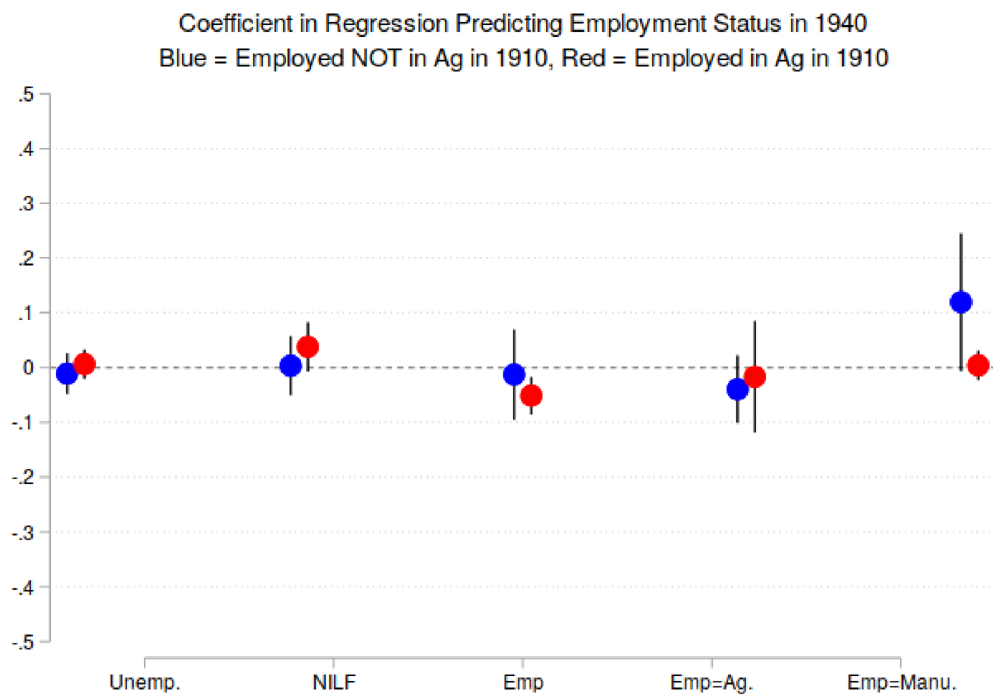


Figure 2.12: Impact of Tractor Adoption Midwestern Men Employed in 1910

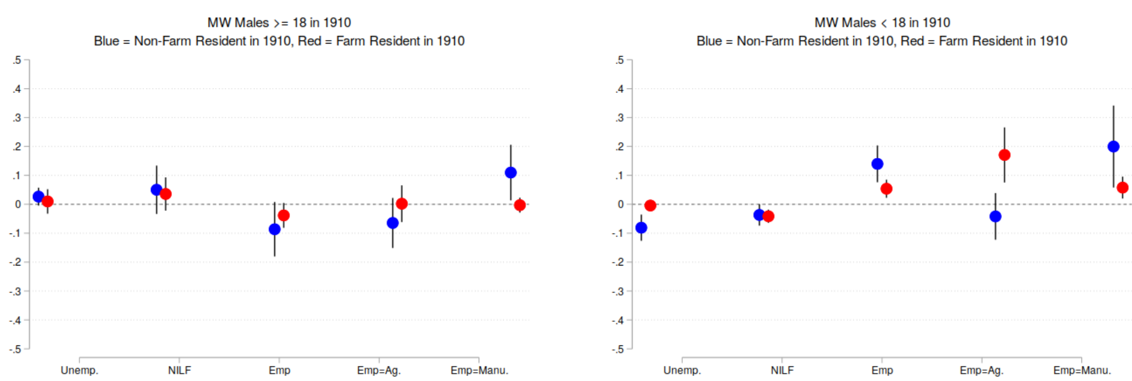


Figure 2.13: Impact of Tractor Adoption Midwestern Males

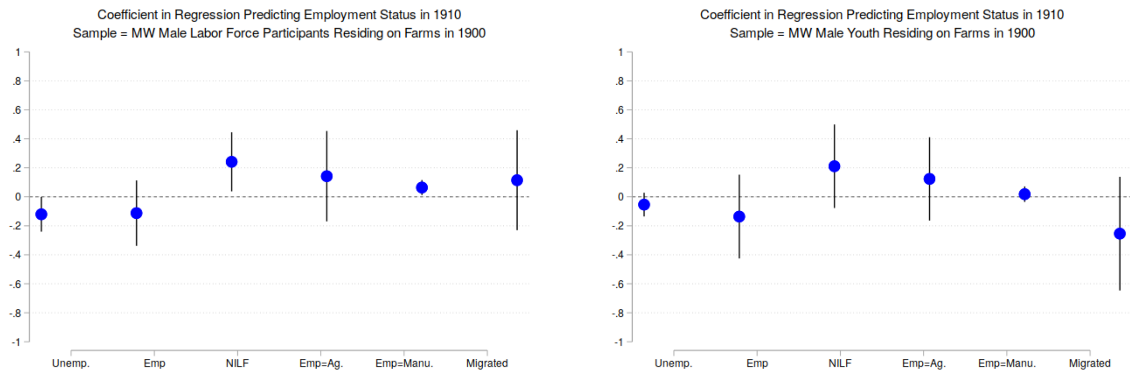


Figure 2.14: Placebo test using 1900-1910 linked data.

Chapter 3

Does Seasonal Adjustment Obscure Early Signs of Recession?

Joint work with Chris Cotton, Emi Nakamura, and Jón Steinsson

This chapter of my dissertation is taken from a working paper I have with three other coauthors. The paper focuses on the measurement of payroll employment as it is released in real time. The chapter uses vintage data to demonstrate how subtle choices in seasonal adjustment methodology can contribute to large revisions to the data. The paper contributes to my overall thesis by demonstrating the data and measurement work to be done in macro-labor so that we can improve real-time policy making.

3.1 Introduction

The US employment data announced by the Bureau of Labor Statistics on the first Friday of each month has immense impact on how economists and the public view the state of the economy. But these data are not without error. The employment data are revised multiple times after their initial release, sometimes by large amounts.

Figure 3.1 presents a graph of initial BLS employment data, announced in real-time the first Friday of the month, alongside the most up-to-date estimate of employment growth each month (which we will call "final" data moving forward). The figure demonstrates that there exists a large gap between the initially announced employment growth and the final data on employment growth over the period. This systematic error results in an underestimation of the severity of a recession as the recession takes hold - an issue that is of top importance for policy makers designing recession relief.

Revisions to the monthly employment data reflect two sources: first, the BLS gathers updated information about the state of the economy, and second, the BLS computes re-estimates of the seasonal adjustment factors that gets applied to the data. In this paper, we investigate the role of the seasonal adjustment re-estimates in introducing bias to the real-time employment data. Our research suggests updates to the seasonal adjustment methodology that can mitigate any real-time bias, particularly during a burgeoning recession.

Table 3.1 demonstrates that the revision bias observed during the heart of the Great Recession was driven primarily by the seasonal adjustment filter used by the BLS. Row (1) of Table 3.1 presents the initial BLS release of the seasonally adjusted (SA) employment data; that is, in October 2008, the initial release for the SA employment change in September 2008 was -159k. Row (2) presents the most up-to-date data (as of May 2020) for the SA employment change in a given month; that is, the May 2020 BLS release for SA employment change in September 2008 was -460k. Row (3) of Table 3.1 presents the total revision by the BLS to the SA employment data, while row (4) presents the total revision to the non-seasonally adjusted (NSA) data released by the BLS. The difference between rows (3) and (4), presented in row (5), shows the amount of the total revision that is attributable to the seasonal adjustment procedure. Clearly, during the Great Recession, the seasonal adjustment procedure employed by the BLS explains a large proportion of the systematic overestimate of employment (i.e. the revision due to seasonal adjustment was large and negative in each month presented and was larger than the revision due to nonseasonal adjustment in each month except September 2009).

Our paper's analysis begins by formally establishing the systematic relationship between intially announced employment growth and final employment growth. Under current seasonal adjustment methodology, the real-time employment growth data systematically overestimates true employment growth during a recession. We exploit a 2003 reform in the seasonal adjustment methodology to shed light on how the seasonal adjustment methods contribute to attenuation bias in the real-time employment data.

After establishing the attenuation of the initial employment growth data relative to true employment growth, we contribute evidence that specific features of the seasonal adjustment method used by the BLS - implemented using the X13-ARIMA software (hereafter "X13") - drives the attentuated relationship between the initial and final employment data. We test whether the seasonal adjustment procedure used by the BLS introduces bias and show that features of the X13 seasonal adjustment algorithm may in fact underly a large proportion of the systematic bias we observe.

We focus on three features of the X13 seasonal adjustment methodology: concurrent estimation of seasonals, forecasting before estimating seasonals, and the bandwidth used in the moving average estimation. We evaluate the impact of each of these parameters on the real-time data. To do this, we use the vintage real-time data provided by the BLS combined with our own X13 implementation to test the impact of these three X13-ARIMA features.

We find that all three of the aforementioned features contribute to the correlation between revisions and final employment growth data. Together, the concurrent estimation and forecast step create a mechanism through which a large shock to the data (for example, during the start of a recession) is propagated forward into the forecasted data, making the shock appear to be seasonally driven rather than a true shock to the economy. When this happens, the component of the shock that X13 attributes to the seasonal factor is amplified, and the apparent magnitude of the true economic shock is thereby diminished. The specified bandwidth can help mitigate this problem if the bandwidth is set to be large enough so that any individual shock is not contributing as much to the seasonal estimates.

The analysis in this paper is not only important for understanding real-time economic data and the mechanisms through which we produce and publish data. It is important because under the current methods of seasonal adjustment, real-time employment data systematically underestimates the severity of a burgeoning recession. If policy-makers do not correct for this bias, they may under-react to the contemporaneous recession and under-supply support to the economy and exacerbate the severity of the recession further. By understanding the mechanisms of the seasonal adjustment methodology used by the BLS, we learn how to better predict recession dynamics and promote recovery in real-time.

The paper proceeds as follows. First, in Section 3.2 we give a brief overview of the current seasonal adjustment methodology used by the BLS. Section 3.3 discusses related research along with some history of the development of the current seasonal adjustment methodology. Section 3.4 describes the data that we use. In Section 3.5 we describe the accuracy and efficiency of the real-time employment growth data released by the BLS for predicting the finalized, true employment growth data. Section 3.6 develops our argument and explanation of how the specific features of X13 contribute to the described pattern of revisions to the employment data and discusses possible alternatives to the current methodology. Finally, Section 3.7 describes why we have not seen a major impact of seasonal factors in the estimation of the severity of the most recent Covid-19 recession. Section 4.5 concludes.

3.2 The BLS's Seasonal Adjustment Procedure

X13-ARIMA is the seasonal adjustment program used to seasonally adjust most US data, including aggregate and industry level measures of total non-farm payroll employment. The procedure involves two key steps.

First, the non-seasonally adjusted employment series is augmented with a "forecast step." During the forecast step, the X13 software estimates a seasonal ARIMA model with all available non-seasonally adjusted data. Using the estimated ARIMA model, the software constructs a forecast of the non-seasonally adjusted series 12 or 24 months into the future (where the exact length of the forecast is set by the user before applying X13). Second, the X13 software applies a series of moving averages to the augmented series to extract the seasonal component, a non-seasonal trend component, and white noise. The seasonal component is removed from the main series to obtain the seasonally adjusted series. This iterated moving average step is called the X11 filter.

The ARIMA model used in the "forecast step" by the X13-ARIMA software and estimated on the data takes the form

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^D z_t = \theta(B)\Theta(B^S)a_t \tag{3.1}$$

where B is the lag operator ($Bz_t = z_{t-1}$), S is the seasonal operator, $\phi(B) = (1 - \phi_1 B - ... - \phi_p B^p)$ is the nonseasonal AR operator, $\Phi(B^S) = (1 - \Phi_1 B^S - ... - \phi_p B^{PS})$ is the seasonal AR operator, $\theta(B) = (1 - \theta_1 B - ... - \theta_q B^q)$ is the nonseasonal MA operator, $\Theta(B^S) = (1 - \Theta_1 B^S - ... - \Theta_Q B^{QS})$ is the seasonal MA operator. The a_t captures i.i.d. white noise with mean zero and variance σ^2. The $(1-B)^d(1-B)^D$ indicates nonseasonal differencing of order d and seasonal differencing of order D.

Even more generally, the X13-ARIMA software allows for a regARIMA to be estimated on the data, in which the model allows for the z_ts to be defined by the residuals of a linear regression. The regARIMA model is written as

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^D(y_t - \sum_i \beta_i x_{it}) = \theta(B)\Theta(B^S)a_t \tag{3.2}$$

where the x_{it} are observed contemporaneously with y_t. In reality, the BLS applies slightly different time series models to the different industry components of the employment data before aggregating and releasing total nonfarm payroll employment [1].

The goal of the forecast step is to improve the seasonal factor estimation at the tail end of the series. More specifically, by adding 24 forecasted months to the end of the series, the X11 seasonal filter is closer to symmetric. That is, the moving average uses data from both sides even at the end of the series.

After the ARIMA estimation and forecast step, the X13-ARIMA procedure uses an iterative moving average method called X11 to estimate the seasonal factors. One of the parameters that plays an important role in the X11 step is the chosen bandwidths of the iterated moving averages. In the BLS's implementation, the precise parameters of the X11 filter vary by industry.[2] Once the X11 filter has been applied to all industries, the BLS aggregates the data to a total non-farm employment seasonally adjusted series.

Important to note is that the estimation of the seasonal factor may or may not include all contemporaneously available data. Starting in May 2003, the BLS began using "concurrent seasonal factor estimation," meaning that they used all available data, including the most recent month's data, to estimate seasonal factors. Before May 2003, the BLS estimated seasonal factors twice annually and used the estimated factors for the following six months, thereby not using all available data for the five months following the seasonal factor estimation. We explore the role of concurrent seasonal factor estimation as a key contributor to the observed real-time bias of the data.

3.3 Related Literature

We are not the first to highlight the important relationship between seasonal factor estimation and recessions. Wright, 2013 argues that the seasonal adjustment filters in use by

[1]We present results that constructed data by applying X13 to the total non-farm employment data rather than to each industry separately and then aggregating. However, we attain similar results when we apply X13 to each disaggregated component and then aggregate.

[2]Our main construction of the data uses a 3x5 filter for all industries so that we can better test the role of the bandwidth filter in the observed real-time bias of the seasonally adjusted data.

US statistical agencies can cause “cyclical variation [to be] misattributed to seasonal factors and vice versa.” His work represents an increasingly large literature establishing the difficulty of disentangling seasonal factors from cyclical variation. In the present paper, we contribute new evidence of the pattern that Wright presents and we investigate new channels and mechanisms of the seasonal adjustment filter that contribute to the systemic bias in real-time data. Our work contributes to the greater literature on statistical uncertainty of real-time economic data, which goes back as far as Morgenstern et al., 1963 and has been discussed more recently by Manski, 2015, who advocates for increased discussion of uncertainty and bias in real-time data releases.

The issue of systematic bias in real-time data is of no small consequence for policy makers. For example, Ghysels, Horan, and Moench, 2018 find that real-time data has much less predictive power for future bond prices than final data, and that data revisions themselves hold predictive information about future prices in the bonds market. These results suggest that for policy makers making decisions in real-time that depend on the state of the economy, real-time data does not give a strong unbiased view of the future state of the world. In addition to posing a significant concern for policy-makers, the puzzling nature of the real-time data motivates a renewed evaluation of a long-running statistical question regarding the accuracy and efficiency of real-time data releases. The question of whether the initial announcement of data indicates “news” or “noise” has been discussed by scholars and policy makers including, notably, Mankiw and Shapiro, 1986. In the framework of Mankiw and Shapiro, 1986, the relationship between initial data and final data can be understood under two different statistical models. The first is a model in which the initial data is a measurement of the final data but with classical measurement error, or “noise.” Under this model, initial data would have higher volatility than final data due to the reduced noise in the final data. The second model stipulates that the initial data is an “efficient” forecast of the final data, and thus smoother - less volatile - than the final data. In their study of GNP, the authors argue that the real-time data that they study falls under the second characterization - that of an efficient forecast of the finalized data. In their paper, Mankiw and Shapiro, 1986 study the US Gross National Product (GNP) data and argue that the initially produced real-time GNP data represents an “efficient” forecast of true GNP given that at the time the real-time data is released, the data are the “best available estimate of the final value.” The initial data GNP data is less volatile than the final GNP data, indicating that the revisions to the data constitute new “news” about the economy, rather than reduced “noise.”

In this paper, we investigate whether the initial release of the BLS employment data qualifies as either an efficient or a noisy prediction of the true, final data. In particular, we evaluate the seasonal adjustment method applied to the BLS employment data in real time to establish whether or not the seasonal adjustment method introduces bias to the data in real time, or whether it purges the data of valuable information that would help better inform policy makers of the state of the economy. We find that the US employment statistics released each month have been systematically biased towards zero since May 2003, when the BLS updated its seasonal adjustment algorithm. That is, the true employment growth (as measured by the finalized data released by the BLS) is greater in magnitude than the initial release suggests. This is particularly a problem during a recession. In a month in which employment growth is negative, the data is revised downward; the true state of the economy is even worse than the initial data indicates. This relationship echoes the relationship that Mankiw and Shapiro, 1986 find in the GNP data. In the case of employment data, the initially released data is attenuated relative to the final, revised data. This attenuation does not necessarily mean that the initial employment growth data does not have strong forecasting power for true employment growth. However, there is not a one-to-one relationship between the initially observed data and the final data. This attenuation may in fact be instrumental to the efficiency of the real-time data.

Statistics Literature on Seasonal Adjustment

An older literature from the 1970s and 1908s evaluates the usage of X-11 to de-seasonalize data, suggesting strategies to give the best predictive real-time data. This older literature reflects advancements in seasonal adjustment methodology that remain highly influential today. Dagum, 1975 introduced the method of augmenting real-time data with forecasted data for future periods before applying X-11 to seasonally adjust the data, dubbed X-11-ARIMA. The X-11- ARIMA model was first adopted in the mid-1970s by Canada's national statistical office. Its predecessor, X-13-ARIMA remains in use by most US statistical agencies today. Kenny and Durbin, 1982 expanded on the method to include more forecasting models to be used in conjunction with X-11. In the same period, researchers debated the merits of concurrent versus non-concurrent seasonal factor estimation. In the early 1980s, it was common practice to estimate seasonal factors once every year (typically in December) and then apply those factors to the subsequent year of real-time data. Kenny and Durbin, 1982 argued that the changing costs of computation should make possible more regular estimation of seasonal factors, taking into account all available data (including contemporaneous data). Their study found that concurrent seasonal factor estimation performs better with the caveat that each seasonally adjusted contemporaneous observation is provisional when it first enters the series since the largest revision to the seasonally adjusted value occurs one month later after the first revision of the NSA data. McKenzie, 1984 and Pierce and McKenzie, 1987 confirmed the view that concurrent seasonal factor estimation improves the predictive power of real-time data by reducing the size of subsequent revisions. Pierce and McKenzie, 1987 investigate the performance of seasonal adjustment factor estimation when the preliminary not seasonally adjusted data is subject to revision. In their study, they find that the theoretical gain, measured as the reduction in root mean square error of seasonal adjustment revisions, of concurrent seasonal adjustment depends on how many revisions the NSA observation undergoes after its first release. In the case of only a one-month revision until the final data is known, the gain decreases as the magnitude of the revision increases. The situation is more complex when revisions persist for more than one month since the gain from concurrent seasonal adjustment depends on several noisy observations. Concurrent seasonal factor estimation began being used on the real-time employment data published monthly by the BLS in 2003 and continues to be the dominant model today.

3.4 Data

This paper uses a number of different data sources related to the monthly employment statistics from the Bureau of Labor Statistics (BLS). In fact, most of the data used in this paper is a version of the monthly BLS employment data. There are four main datasets used in the paper: final seasonally adjusted (SA) employment growth, final non-seasonally adjusted (NSA) employment growth, real-time SA employment growth, and real-time NSA employment growth. In this section, we describe these datasets and their relation to one another.

The first, and most commonly cited employment data in economics research is the final release of seasonally adjusted (SA) total non-farm payroll employment.[3] These data are used by researchers to measure "true" employment growth in the economy each month. In accordance with its popularity, the dataseries is widely available from the BLS website or from the popular St. Louis Fed's FRED. For every month τ, this data series includes an employment level $E^{SA}_{\tau,T}$ where the superscript indicates that the data has been seasonally adjusted and the second subscript T indicates that the data is the most up-to-date estimate of employment in month τ.

In addition to the seasonally adjusted data, the BLS also releases non-seasonally adjusted (NSA) employment data. The seasonal adjustment filter that the BLS applies is called X-13-ARIMA and uses a standard set of parameters to de-seasonalize the employment data. We discuss the parameters of the BLS X-13-ARIMA at length in Section 3.2. Thus for each month τ, we observe $E^{NSA}_{\tau,T}$ in the dataset of "final" NSA data from the BLS.[4]

Lastly, we use the real-time vintages of the SA and NSA data to better understand how the seasonal adjustment filter interacts with observed revisions to the data. For each month τ, there is an initial release of the data at time $\tau + 1$, denoted $E_{\tau,\tau+1}$. This datapoint is revised the following two months, and again in each of the subsequent two years (for NSA data) or each of the subsequent five years (for SA data) [5]. Occasionally, there are benchmark updates to NSA employment levels going back many years that also change SA employment levels. Thus for a given month τ's employment situation, we observe a full series of vintages for that datapoint, $E_{\tau,\tau+1}, E_{\tau,\tau+2}, ..., E_{\tau,t}, ..., E_{\tau,T}$. Revisions to the data reflect both changes in available information at the BLS, as well as updated estimation of the seasonal factors using all available data (up to and beyond time τ).

[3] U.S. Bureau of Labor Statistics, All Employees, Total Nonfarm [PAYEMS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PAYEMS, May, 2020.

[4] U.S. Bureau of Labor Statistics, All Employees, Total Nonfarm [PAYNSA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PAYNSA, May, 2020.

[5] The nonseasonally adjusted (tri_000000_NSA) and seasonally adjusted (tri_000000_SA) vintage data can be downloaded in a zip file from the CES website using the following link on the CES website: https://www.bls.gov/web/empsit/cesvindata.htm . These data begin in May 2003. We also use real-time seasonally adjusted data from the Philadelphia Fed that provides realtime data going back further in time: https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/real-time-data-set-for-macroeconomists

We use the real-time vintages of the NSA data to re-construct real-time seasonally adjusted data under different seasonal adjustment methods. Construction of the real-time SA data that we use is described in more detail in Section 3.4.

Construction of X13 Seasonally Adjusted Data

To study the X13 seasonal adjustment process, we construct our own seasonally filtered series by applying the X13-ARIMA software to the non-seasonally adjusted real-time BLS employment data. We vary the exact specification used by the X13 program to test the significance of each specification choice in the estimated bias of the initial SA release of the data.

To construct the data, we apply the X13 algorithm to the real-time NSA data available for each month. For each month τ, we start with a series of aggregate NSA employment data for months t_0 (the start of the available data is May 2003) up to τ, where each month's data is the data that was available from the BLS at time $\tau + 1$ (the first available observation of data for month τ). That is, we start with the NSA real-time data from month $\tau + 1$, which consists of the datapoints $\{E^{NSA}_{t_0,\tau+1}, ..., E^{NSA}_{\tau,\tau+1}\}$ where each datapoint refers to a different month observed at time $\tau + 1$. To produce the real-time seasonally adjusted data with our own X13 algorithm, we put this series through the X13 filter with a chosen set of X13 parameters. The resultant seasonally adjusted data gives us our approximation of the "initial X13 data" for month τ, the single datapoint $\tilde{E}^{SA}_{\tau,\tau+1}$. To get the "initial X13 data" for the next month, $\tau + 1$, we use the the full NSA real-time data series that was available at time $\tau + 1$, $\{E^{NSA}_{t_0,\tau+2}, ..., E^{NSA}_{\tau+1,\tau+2}\}$. The data observed at time $\tau + 2$ but referencing employment in month τ is often different from the data for month τ that was available at time $\tau + 1$. This is due to the fact that in their first three months, data are revised to reflect better information and seasonal adjustment estimation. Thus it is important and relevant that we use the whole series of real-time data that is contemporaneously available each month to compute the "first release" of each seasonally adjusted series.

The real-time data published by the BLS is not available at the industry level, only in aggregate. For this reason, we apply X13 to the aggregate series rather than applying it to each individual industry-level series and then aggregating up, as the BLS does. While this may pose a threat to our analysis, we show in Figures B1 and B2 that the results of applying X13 to the aggregate real-time NSA data tracks very closely to the BLS's SA data which is based off of the disaggregated NSA data. In Figure B1, we see that the initial data we create by running the baseline X13 specification on the real-time data is extremely close to the BLS real-time SA data. Likewise, Figure B2 shows that the data we simulate as the final data under our X13 baseline specification on aggregate data closely tracks the BLS final data series. We take these as evidence that our X13 method using aggregate real-time NSA data is realistic and not introducing new noise or biases.

The real-time NSA data begin to be available in May 2003, thus all analysis that uses the real-time NSA data is focused only on the post-2003 period. This means that we cannot test how changing the seasonal adjustment method would impact the pre-2003 period, during which time the BLS did not use concurrent seasonal factor estimation. It also limits the analysis to primarily the Great Recession period. In our reported regression results, all series end in Dec. 2018 so that we can use at least 24 months of subsequent data to construct "final" data estimates (as the BLS continues to revise data following the initial release). Nonetheless, our results are robust to changing the end-date of the data we use.

3.5 Predicting True Employment Growth in Real-Time

The disparity between the real-time employment data and the final "true" employment data during the Great Recession motivates a deeper investigation into the workings of the seasonal adjustment method and how to best interpret real-time data in the midst of a recession. In particular it appears that a large proportion of the discrepancy between the initial data and the final data originates from the seasonal adjustment procedure. One way to further investigate this claim is to examine changes made to the seasonal adjustment method used by the BLS to see whether these changes altered the pattern between true employment growth and data revisions.

In May 2003, the BLS made a number of changes to their seasonal adjustment methodology, including that they began employing "concurrent seasonal adjustment." In this section we present results showing that the seasonal adjustment method used by the BLS since May 2003 appears to attenuate the data so that during a recession, the real-time data under-predicts the severity of the employment loss.

Table 3.2 looks at the accuracy of the initial real-time data by regressing the most up-to-date release of month-to-month employment growth (the closest estimate we have to the "truth") on the initial real-time BLS release for the respective month. Specifically, columns (1) and (3) of Table 3.3 present the results of regressing the "true" final employment growth numbers published by the BLS on the initial release of that month's data by the BLS.[6] The regression takes the form

$$\Delta e^{SA}_{\tau,T} = \beta \Delta e^{SA}_{\tau,\tau+1} + \mu_{\text{month}} + \epsilon_{\tau} \tag{3.3}$$

where $\Delta e^{SA}_{\tau,T}$ is up-to-date (month T) SA employment growth data for employment growth in month τ and $\Delta e^{SA}_{\tau,\tau+1}$ is SA employment growth in month τ initially published by the BLS (which we denote as time $\tau+1$, since the data for month τ first comes out at the beginning of the following month). We include a month-of-year fixed effect μ_{month}.[7]

If the initial data perfectly forecast the true employment growth, we should see a one-to-one relationship between the initial data release and the finalized data. That is, if data collection were perfect, the first release of the data would be equal to the final release of the data and no adjustments would need to be made. Given random measurement error

[6] Employment data was downloaded from FRED as of May 8, 2020.

[7] Results are robust to excluding the month-of-year fixed effects. See Appendix Table B2.

in the regressor, we might expect that rather than a coefficient of 1, we'd see a coefficient below one. Prior to the introduction of concurrent adjustment in May 2003, we estimate a coefficient of 0.807, as presented in column (1) of Table 3.2. This finding corresponds to the "noise" model postulated by Mankiw and Shapiro, 1986. There are two possible channels that may result in an estimate less than one. First, classical measurement error creates attenuation bias, and thus we'd expect that a certain amount of i.i.d. white noise in the initial estimates of employment growth would result in a coefficient less than 1. The second, and perhaps more concerning reason, is that the seasonal adjustment procedure introduces this bias by absorbing the recessionary shock into the final seasonal factors, but not the initially computed seasonal factors. This could be the case with non-concurrent seasonal adjustment, as will be discussed further below.

In contrast, in the post-May 2003 period, the coefficient estimated and presented in column (3) is 1.177, meaning that the initial release of employment growth understates the true change in employment growth (as measure by the most up-to-date version of the data). A coefficient of 1.177 should be interpreted to mean that an initial release that is 1 percentage point below the mean suggests the final release will be 1.177 percentage points below the mean. Thus, the initial release understates the degree to which the final release deviates from the mean. Because the systematic under-estimation in the initial data only emerges in May 2003, it appears that the concurrent seasonal adjustment method may be associated with systematic bias towards zero in the BLS's initial estimate of employment growth.

Figure 3.2 plots the coefficients estimated by running the regression specified in Equation (3.3) as a rolling regression on the data with a sliding time window used for each month's estimate. The window of data used for estimation in each month includes the two years before each month and the two years after each month. This means that there are 49 months of data included in each estimated coefficient. The red line in Figure 3.2 plots the estimates in columns (1) and (3) of Table 3.2. We see that the estimated coefficient does appear to rise around May 2003, though there is significant variation between time-windows both before and after May 2003. Overall, the results of this exercise suggest that May 2003 does appear to be a breaking point in the relationship between initial seasonally adjusted employment data and final employment data in the US. Further, this figure shows that the main increase in the coefficient relating initial to final data occurs during the Great Recession. This fact raises concerns regarding whether future recessions will see the same pattern and therefore realtime data will underestimate the severity of recessions on an ongoing basis. We discuss the COVID-19 recession below.

Our results so far suggest that, starting in May 2003, the error in initial employment growth data is best understood as "efficient" rather than "noisy." While the statistical relationship between the initial real-time data and the final data may be "efficient" in the Mankiw and Shapiro, 1986 sense, having an attenuated real-time data release is concerning if policy makers don't make proper adjustments and thus don't respond appropriately to an economic downturn. In particular, columns (2) and (4) of Table 3.2 demonstrate that the majority of the under-estimation of the real-time data loads on times when final employment growth is negative, i.e. employment is contracting. In columns (2) and (4), we report the results from a regression of the form

$$\Delta e_{t,T} = \beta_1 \Delta e_{t,t+1} + \beta_2 \Delta e_{t,t+1} 1\{\Delta e_{t,T} > 0\} + \mu_{month} + \epsilon_{\tau} \tag{3.4}$$

where we've now included an interaction term to see how the relationship between the initial and final data differs when employment growth is negative. The results demonstrate that in times of employment contraction, the attenuated relationship between the initial real-time data and the final data is amplified further. This is most likely because these times see sharp and abrupt changes in the employment rate, and thus are more likely to alter the seasonal factor estimation. This result also serves as motivation for why understanding the seasonal adjustment method and how to forecast true employment growth using real-time data is of particular importance at the onset of economic recessions.

3.6 Alternative X13 Methods

In this section, we examine how changing the X13 procedure would alter the accuracy of the initial release of employment relative to the final release. We consider four methods: the standard X13 protocol currently in use at the BLS (which includes both the forecast step and concurrent adjustment), X13 without the forecast step, X13 without concurrent adjustment, and X13 with a wider bandwidth used in the iterated moving average estimation. To study these variants on the X13 seasonal adjustment procedure, we simulate real-time seasonally adjusted data using vintage NSA data from the BLS. That is, we run the real-time vintage NSA data through the X-13-ARIMA seasonal adjustment software under different set parameters of the model. We construct the data without the forecast step, without concurrent seasonal factor estimation, and with the highest possible bandwidth.

The simulated data for each scenario is pictured in Section 3.9. Note that in the top left of the figure, we plot the BLS initial data along with our re-construction of the BLS's seasonal adjustment methodology. Neither case matches the final "true" employment growth data. In contrast, we see that all three of our alternative X13 procedures get much closer to accurately predicting the final employment growth data in real time. In the remainder of the section, we describe our alternative methods in detail and discuss why they perform better, particularly during a recession. We evaluate the performance of each seasonal adjustment procedure using the estimated coefficient relating the initial and final employment growth data as well as computing the root-mean-squared-error (RMSE) on the estimated relationship between the initial and final data. While the coefficient gives a measure of bias in the initial data, the RMSE gives us a measure of the precision or efficiency of the initial data. Each are important to understanding the value of the real-time data under our different seasonal adjustment procedures.

Method 1: BLS Replication

Since May 2003, the BLS has employed both concurrent seasonal adjustment and a forecasting step. That is, the currently used BLS method utilizes all non-seasonally adjusted employment data up to and including month τ to estimate the seasonal factors for this release. This is referred to as concurrent seasonal adjustment because the non-seasonally adjusted employment for month τ impacts the seasonal factor for month τ. The idea behind using concurrent seasonal adjustment is that seasonal factors can change over time, and concurrent seasonal adjustment helps to ensure they remain accurate and up-to-date.

The forecast step extends the series by estimating a seasonal ARIMA time series and using this estimated model to forecast future data-points before applying the seasonal filter. The forecast step is used to ensure that even at the tail of the series (the most recent data), a relatively symmetric moving average can be applied to estimate seasonal factors.

As discussed in Section 3.5, the BLS's current seasonal adjustment procedure understates seasonally adjusted employment changes in initial BLS releases. The regressions in columns (2) and (4) of Table 3.2 reflect the regressions corresponding to Equation (3.3) and Equation (3.4). That is, Table 3.2 presents the results of regressing the final BLS release of seasonally adjusted employment growth on the initial BLS release of seasonally adjusted employment growth. Ideally, the reported coefficients would be 1, implying that the initial release of seasonally adjusted employment growth provides a perfectly accurate picture of the final release. The regression in column (2) examines the relationship from May 2003 onwards when the BLS was using its current seasonal adjustment methodology with concurrent adjustment and a forecast step. The coefficient for regression (2) is significantly above 1, which implies that the initial release of seasonally adjusted employment growth is actually significantly understating the final release of seasonally adjusted employment growth. In other words, when the final release is positive, the initial release will be positive but not as high; and when the final release is negative, the initial release will be negative but not as low.

Our replication of the BLS seasonal adjustment methodology supports the result that the current seasonal adjustment methodology understates employment growth in initial releases. Regression (1) in Table 3.3 looks at the same regression as eq. (3.3), but rather than use BLS seasonally adjusted employment data, we use our replication of the BLS seasonally adjusted employment growth series, which we construct using the underlying BLS non-seasonally adjusted data and our own application of X13-ARIMA. We estimate a coefficient of 1.159, again finding that the initial release of seasonally adjusted employment growth understates the final release. [8] The RMSE on this regression is 0.682, which we take as the baseline against which we will compare our other methods.

[8]The reason we don't find exactly the same numbers as in regression (2) of Table Table 3.2 is largely because we only have realtime aggregate nonseasonally adjusted data on which to apply the seasonal adjustment procedure. The BLS instead applies its seasonal adjustment procedure to disaggregated series, but historical realtime data are not available for these.

Method 2: No Forecast Step

We now consider what happens if we remove the forecast step from the X13-ARIMA procedure. That is, we run real-time vintage data from the BLS through a variant of the X13-ARIMA procedure in which we do not allow the algorithm to forecast data forward into the future before applying the X11 moving average filter. Column (2) of Table 3.3 displays the results for a regression of the true final data on real-time SA data constructed using our alternative X13 procedure without the forecast step.

Simulated results show that turning off the forecast step corrects the observed bias in the initial real-time data, so that the initial release of seasonally adjusted employment growth accurately forecasts the final release (an estimated coefficient close to 1). We find a coefficient of 1.018, which is not significantly different from 1. Therefore, we do not reject the hypothesis that when we turn off the forecast step the initial release accurately forecasts the final release. Consequently, we find that turning off the forecast step appears to yield initial releases of the data that are less biased than with those computed using the forecast step.

We also find that the RMSE of this regression is slightly lower than the RMSE from our regression testing the BLS's seasonal adjustment procedure (column (1) of Table 3.3). This suggests that without the forecast step, the predictive power of the real-time data is slightly better than the real-time data that was constructed using the forecast step.

To understand why the forecast step contributes to attenuation in real-time seasonally adjusted data, consider the following example:

The simplest possible monthly seasonal ARIMA model with a unit root in the trend and seasonal series, but without any autocorrelation or moving averages, takes the form in eq. (3.5) for monthly data. [9] This states that the change in employment in the current month should equal the change in employment in the same month one year ago plus some residual. In this case, the forecast of the change in employment 12 months and 24 months from now are given by eq. (3.6). We observe that these forecasts are exactly the same as the change in employment in the current month.

$$\Delta X_t = \Delta X_{t-12} + \epsilon_t \tag{3.5}$$

$$\mathbb{E}_t[\Delta X_{t+24}] = \mathbb{E}_t[\Delta X_{t+12}] = \Delta X_t \tag{3.6}$$

Now, consider the case where there is a large fall in employment in the latest month for which employment data is available, denoted t: In this case, the forecast given by eq. (3.6) will also show a large fall at times $t+12$ and $t+24$. Then, the seasonal adjustment procedure will observe large falls at times $t, t+12, t+24$, so it will look like the seasonal pattern is to have large falls in the calendar month associated with time t. Thus, since the moving average suggests there are typically large falls in the calendar month of t and annually thereafter, the seasonal factor is estimated such that these large declines are considered seasonal. Thus, the seasonally adjusted release of data for month t and time t will be seasonally adjusted upwards (as if the seasonal factor that month is a large negative number).

[9] The ARIMA models used by the BLS all allow for a unit root in the trend and seasonal components.

However, once the data for times $t + 12$ and $t + 24$ are actually released over the next two years, it is likely that they will not show the same steep decline in employment that was observed in month t. That is, it will become clear that what the model assumed was seasonal contemporaneously was actually a one-month aberration from the trend. In later releases, the data used to construct the seasonal adjustment factor at time t will only show a fall in employment in month t and not months $t + 12$ and $t + 24$, so it will not look like the seasonal pattern is to have large falls in the calendar month associated with times t, $t + 12$, and $t + 24$. With more accurate, non-forecasted information about times $t + 12$ and $t + 24$, the revised employment growth data for time t will no longer be seasonally adjusted upwards as it was in real-time. The employment growth data for month t will be revised downward as it becomes clear that the originally observed shock was not due to seasonal factors. Consequently, the initial real-time release of the data understates the true negative change in employment growth that is eventually revealed in the final release.

Figure 3.4 illustrates the forecasted data for the simulated real-time seasonal adjustment procedure used in September of 2008, at the start of the Great Recession. We see that the forecasted data does not accurately capture the extent of the unfolding recession. Instead, the forecasted data predicts a more moderate decline in employment. The estimated error in this prediction does largely include the true path of employment. However, this error is not reported in the real-time data and thus the real-time data that is reported relies on this underlying assumption that the recession will be more moderate. By including this forecast in the estimated seasonal factors, the X11 procedure absorbs some of the "real" decline in employment as part of the seasonal factors. By removing the forecast step, the X13 procedure reduces the chance of imbedding large assumptions over the future path of employment in the estimates of the seasonal factors.

Method 3: Non-concurrent Adjustment

Before May 2003, the BLS applied non-concurrent seasonal adjustment. Non-concurrent seasonal adjustment means that if the latest release is for employment in April 2003, the seasonal factor for April 2003 is computed without using non-seasonally adjusted employment in April 2003. [10] In this subsection, we look at whether using non-concurrent seasonal adjustment would also produce initial releases of seasonally adjusted employment growth that give unbiased predictions of the final release. We find that non-concurrent seasonal estimation results in biased initial data, but that the bias in this case is towards overstated shocks in real-time rather than understated shocks. This results appears to fall under the "noise" model of real-time data.

As discussed in Section 3.5, the non-concurrent seasonal adjustment procedure used prior to May 2003 overstated seasonally adjusted employment changes in the initial BLS releases. Regression (1) in Table 3.2 shows a regression of the final BLS release of seasonally adjusted

[10] In practice, the BLS would recompute projected seasonal factors every 6 months for 6 months into the future and would use these factors to seasonally adjust the subsequent six months.

employment growth on the initial BLS release of seasonally adjusted employment growth when a non-concurrent seasonal adjustment procedure was used. The coefficient of 0.822 is significantly less than 1, implying that the real-time release of seasonally adjusted employment growth significantly overstates the true employment growth. A coefficient less than 1 indicates that the initial release may contain random measurement error in the form of white noise, that is eventually discarded as the releases of the data get more accurate (the "noise" model of Mankiw and Shapiro, 1986). Another possible cause of a coefficient less than 1 might be that with non-concurrent adjustment, the seasonal factor estimation may not capture a true contemporaneous change in the seasonal factor. In this case, the seasonal factor estimation will catch the change in the seasonal factor in the final data but not in the real-time data and thus the real-time data will overestimate changes to employment growth. In either case, we would observe a coefficient less than 1 in our estimation of Equation (3.3) and the real-time data will be biased relative to the final data.

Our own seasonal adjustment analysis supports the result that the non-concurrent seasonal adjustment methodology overstates employment growth in initial releases. We apply our own non-concurrent seasonal adjustment approach to real-time non-seasonally adjusted data for May 2003 onwards.[11] Column (3) in Table 3.3 reports the results from a regression of the final release on the initial release with our non-concurrent seasonally adjusted employment data. We find a coefficient of 0.948, which is again less than 1, though in this case only with marginal statistical significance.[12] The RMSE of this regression is 0.714, which is larger than the RMSE using either of the two previous models. This result suggests that while using non-concurrent seasonal adjustment does improve the real-time bias of the data, it also reduced the amount of relevant information that the real-time data uses to predict the final data. That is, the real-time data has less precision in predicting the final data than if we were to use contemporaneous seasonal factor estimation.

Note that with non-concurrent adjustment, it does not matter as much whether or not the forecast step is included in the seasonal adjustment procedure, since the release for the current month is not used in the analysis. We can see this by observing column (5) of Table 3.3, which reports the regression using real-time data constructed with non-concurrent seasonal factor estimation and without the forecast step. We see that with this specification of the seasonal adjustment method, we get nearly the same coefficient as we did with non-concurrent estimation and the forecast step (column (3)). However, without the forecast step, we see a significant increase in the RMSE, from 0.714 to 0.830, suggesting that in this case, the forecast step improves the informativeness of the real-time data.

[11] Realtime nonseasonally adjusted employment data is not available before May 2003.

[12] It is not surprising that this coefficient is different from the BLS coefficient, since we are considering a different time period and our adjustment procedure is applied to aggregate rather than disaggregate employment series due to a lack of available data.

Method 4: Wide Bandwidth

Our final method examines the role of the moving average bandwidth employed by the X11 seasonal factor estimator. After the ARIMA estimation and potential forecast step, the seasonal factor estimation procedure employs an iterated weighted moving averages over the data. Specifically, the X11 procedure takes a symmetric n-term moving average of m-term averages. This is referred to as an $n \times m$ seasonal filter (Wright, 2013). In its final iteration over the data, the X11 algorithm chooses a bandwidth, $n \times m$ for the final computation of seasonal factors. In the BLS's employment of X13-ARIMA, the bandwidth is chosen dynamically based on the volatility of the data. However, in our analysis so far, we have fixed the chosen bandwidth to be 3×5, which is the baseline bandwidth used by X13. To test the importance of the bandwidth parameter, we construct the data using a 3×15 bandwidth, the maximum possible bandwidth. In column (4) of Table 3.3 we report the regression result using data constructed using the 3×15 bandwidth. We find a coefficient of 1.063 which is slightly larger than 1. In this case, we are still using the forecast step and concurrent estimation. We see that by using a wider bandwidth, the bias in the prediction of the final data using the real-time data is reduced relative to the BLS's methodology. However, we also see that the RMSE is 0.723, larger than the previously explored models. Therefore, by using a wider bandwidth, the informational content of the initial data is reduced relative to methods using a smaller bandwidth.

Wright, 2013 describes the bandwidth choice in depth. Wright, 2013 argues that "The choice of filter is effectively the bandwidth choice in a nonparametric statistical problem, and the choice of bandwidth involves a bias-variance trade-off. If seasonal patterns fluctuate a great deal, then a small choice of bandwidth will be appropriate to reduce the problem of changing seasonals being incorrectly attributed to cyclical variation (bias)."

Using this logic, a large bandwidth will mitigate the problem of cyclical variation being attributed to fluctuating seasonals. Because we stipulate that the current BLS seasonal adjustment procedure attributes true cyclical fluctuations to the seasonal factor in the initially released real-time data, applying a larger bandwidth will help mitigate this fact.

Combining the larger bandwidth with the no-forecast and non-concurrent estimation results in a coefficient of 0.940 and a RMSE of 0.818. These results suggest that under the no-forecast-non-concurrent model, a wider bandwidth improves on both the bias dimension and the information dimension.

3.7 How this might matter for the COVID-19 recession?

The ongoing Covid-19 recession demonstrates an enormous change in the employment data between the months of March and May, 2020. The change has been large enough that one might expect the X13 procedure employed by the BLS to introduce large bias to the real-time data. If the procedure behaves as it did during the Great Recession, we'd expect the seasonal factor for April 2020 to absorb a large portion of the shock to employment growth, and a series of downward revisions to the data.

In fact, the shock to employment growth has been so anomalous during the Covid-19 recession that the BLS made explicit comments regarding the impact of the shock to their seasonal adjustment methodology. In a published statement, the BLS wrote that for the payroll employment data, the BLS did not make any manual adjustments to the X13 algorithm, but noted that the algorithm includes active use of outlier detection. On the other hand, for the household survey that feeds into the unemployment numbers (not studied in this paper), the BLS did make manual changes to the seasonal adjustment method, making sure to mark recent months as outliers in the data. [13] Clearly, the Covid-19 pandemic marks an extreme-enough event as to warrant direct examination by the BLS as to how the seasonal adjustment method responds to large shocks in the economy. Fortunately, the X13 algorithm does appear to have some capacity to detect and adjust for large outliers.

The X13 outlier detection procedure tags months that it deems to be outliers in the data using a user-determined sensitivity (critical value). Once an outlier is detected, the algorithm excludes the datapoint from the ARIMA estimation prior to the forecast step and excludes the datapoint from the X11 seasonal filter estimation. We examine the role of the outlier detection using data from the 2020 recession. Running a simulation of the real-time employment data under our alternative X13 specifications, we test whether the outlier detection that became active in April 2020 impacts our results. We run the same simulations as we ran in the rest of this paper, and now add a simulation that removes all outlier detection. The time paths of the real-time initially announced employment data under these different X13 specifications are plotted in Figure 3.5. We see that seasonal adjustment without the outlier-detection step produces a value of employment nearly 50 percentage points in annualized growth rates above what we see when using the BLS outlier detection sensitivity. The data produced without outlier detection remains attenuated relative to the data produced with the outlier detection, lending credence to our result that under normal circumstances, when the outlier step is non-binding, the seasonally adjusted real-time data is attenuated relative to the true data.

This result validates our proposed mechanism - that the seasonal adjustment process over-reacts to extreme changes in the data (unless, of course, the change is officially detected as an outlier by the X13 algorithm, as it was in April 2020). Without the outlier-detection step, the negative shock is captured in the seasonal factor so that the seasonal factor is more negative than it would be otherwise and thus the seasonally adjusted real-time data is greater than it should be (and will be in the final data). Once releases from future years don't confirm this seasonal pattern, the seasonal factor gets revised upward causing the final release of the employment data to be revised downward. When the outlier detection step is included in X13, the initial release is much closer to the final release and we don't see the

[13] See the BLS's statements regarding the employment and unemployment data during the Pandemic here: https://www.bls.gov/covid19/effects-of-covid-19-pandemic-and-response-on-the-employment-situation-news-release.htm

same downward revision to the data.

The global pandemic has induced a recession of enormous proportion, but the seasonal factor estimation done by the BLS has not been dramatically impacted due to the appropriately sensitive outlier detection. In fact, the seasonal adjustment process we've seen during the pandemic further proves the importance and validity of the mechanism we highlight by illustrating how the seasonal factor estimation plays a role in the observed revision bias.

3.8 Conclusion

In this paper, we have reviewed the role of the commonly used X13-ARIMA seasonal adjustment methodology, most notably employed by the BLS in the seasonal adjustment procedure applies to real-time employment data each month. We find that the X13 methodology introduces attenuation bias in real time which causes underestimation of recession severity in real time. We trace the source of this bias to the forecasting procedure and concurrent estimation protocol specified in the BLS's preferred X13 method. When we remove either of these steps from the X13 procedure applied to real-time NSA data, we get better real-time predictions of the true data.

Considering the framework of Mankiw and Shapiro, 1986, we see that the current methodology used by the BLS appears to fall in the model of "news" rather than "noise." That is, the early releases of employment data act as efficient predictions of final data. Revisions to the preliminary data reflect newly acquired "news" about the state of the economy, rather than reductions in "noise." While this may be a good model in general, it does create a persistent pattern of under-estimations of true economic contraction in real time. We argue that this is an important issue for policy makers to consider in order to best understand how to structure appropriate recession relief strategies.

3.9 Tables and Figures

Table 3.1: Revisions during the Great Recession

	2008M09	2008M10	2008M11	2008M12
BLS Data (Standard Seasonal Adjustment)				
(1) Δ SA Employment, Initial Release	-159	-240	-533	-524
(2) Δ SA Employment, Final Release	-460	-481	-727	-706
(3) Revision of Δ SA Employment (2)—(1)	-301	-241	-194	-182
(4) Revision of Δ NSA Employment	-192	-8	-48	-74
(5) Revision due to Seasonal Adjustment (3)—(4)	-109	-233	-146	-108

Note: All data is measured in thousands. Δ SA Employment measures the final (most recent) release of the change in seasonally adjusted employment. Revision of Δ SA Employment = Final Release of Δ SA Employment - Initial Release of Δ SA Employment. Revision of Δ NSA Employment = Final Release of Δ NSA Employment - Initial Release of Δ NSA Employment. Revision due to seasonal adjustment = Revision in Δ SA Employment - Revision in Δ NSA Employment.

Table 3.2: Predicting final employment growth using initially-announced employment growth.

	(1)	(2)	(3)	(4)
	BLS final	BLS final	BLS final	BLS final
BLS initial SA	0.822	1.187	0.782	1.085
	(0.0245)	(0.0283)	(0.0333)	(0.0609)
BLS initial SA * $1(e\_\tau, T < 0)$			0.127	0.174
			(0.0721)	(0.0923)
Constant	0.0663	-0.0944	0.195	0.0439
	(0.234)	(0.160)	(0.245)	(0.175)
N	462	188	462	188
Period	Pre-5/03	Post-5/03	Pre-5/03	Post-5/03
Month FEs	Y	Y	Y	Y
RMSE	1.422	0.613	1.419	0.609

Note: This table presents the results of regressing the most up-to-date employment growth data released by the BLS on the initially released real-time employment growth data (also from the BLS). Columns (1) and (3) show the relationship in the data before May 2003 at which point the BLS updated their seasonal adjustment method to include "concurrent seasonal factor estimation." Columns (2) and (4) show the relationship after May 2003. In columns (3) and (4) the regression includes the intitial real-time data interacted with a dummy for whether employment growth was negative that month as an additional regressor. All regressions include month-of-year fixed effects. Data in columns (1) and (3) begins in Nov. 1964. Data in columns (2) and (4) ends in Dec. 2018. Standard errors are reported in parentheses.

Table 3.3: Regressions predicting final employment growth using alternative X-13 ARIMA specifications.

	(1) BLS final	(2) BLS final	(3) BLS final	(4) BLS final	(5) BLS final	(6) BLS final
BLS Replication	1.159 (0.0312)					
No forecast step		1.030 (0.0275)				
Non-concurrent			0.948 (0.0268)			
3x15 Bandwidth				1.063 (0.0305)		
Non-concurrent, No forecast					0.928 (0.0313)	
No forecast, non-concurrent, 3x15 BW						0.940 (0.0312)
Constant	-0.0328 (0.178)	0.0302 (0.177)	0.108 (0.185)	-0.218 (0.189)	0.130 (0.215)	-0.196 (0.214)
N	188	188	188	188	188	188
Month FEs	Y	Y	Y	Y	Y	Y
RMSE	0.682	0.679	0.714	0.723	0.830	0.818

Note: This table presents the results of regressing the most up-to-date employment growth data released by the BLS on the initially released real-time employment growth data under a variety of different X-13 specifications. The table uses data from May 2003 to December 2018. Column (1) reports the regression using the re-constructed real-time seasonally adjusted data meant to replicate the BLS. Column (2) reports the regression using a version of the real-time data constructed with X-13 in which the bandwidth of the moving average estimator is increased to 3x15. Column (3) uses data produced from running X-13 without concurrent estimation of the seasonal factors. Column (4) uses data produced from running X-13 without the forecasting step of the seasonal adjustment process. Column (5) uses data produced from running X-13 without the forecasting step of the seasonal adjustment process and also without the concurrent estimation. Column (5) uses data produced from running X-13 without the forecasting step, without the concurrent estimation, and with a bandwidth of 3x15. All regressions include month-of-year fixed effects. Standard errors are reported in parentheses.

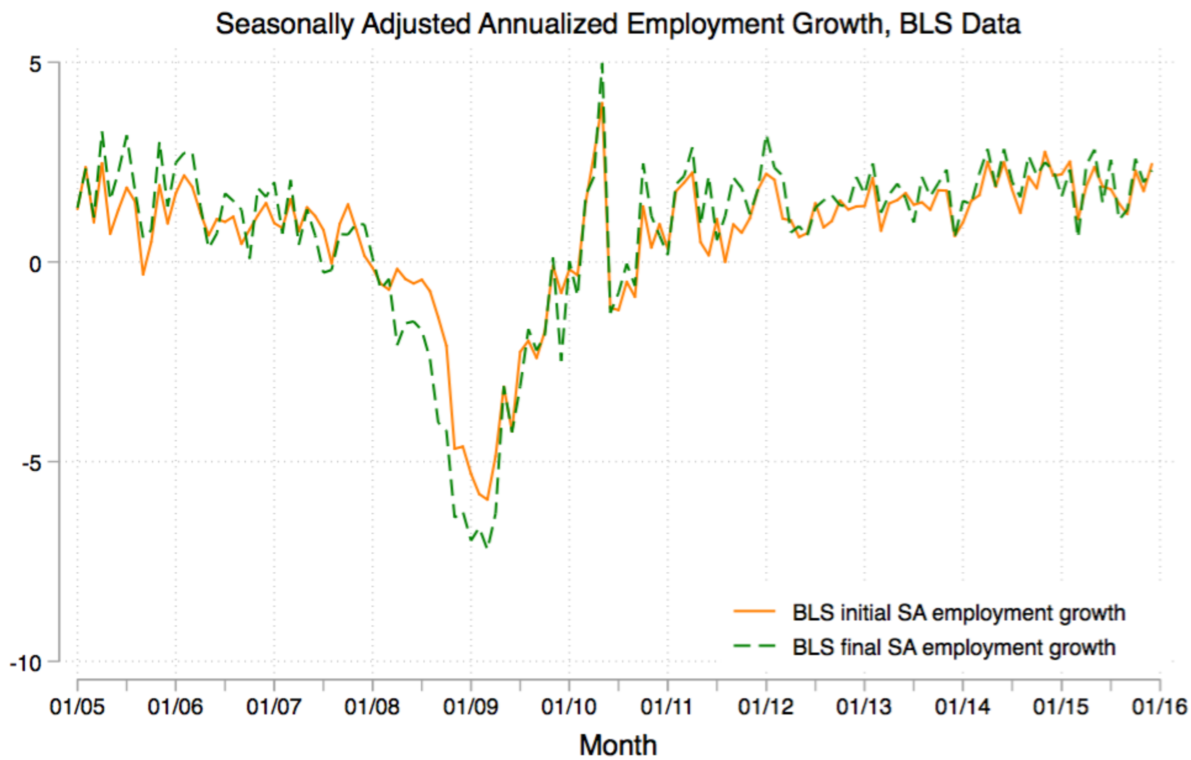


Figure 3.1: Employment Growth Data from BLS

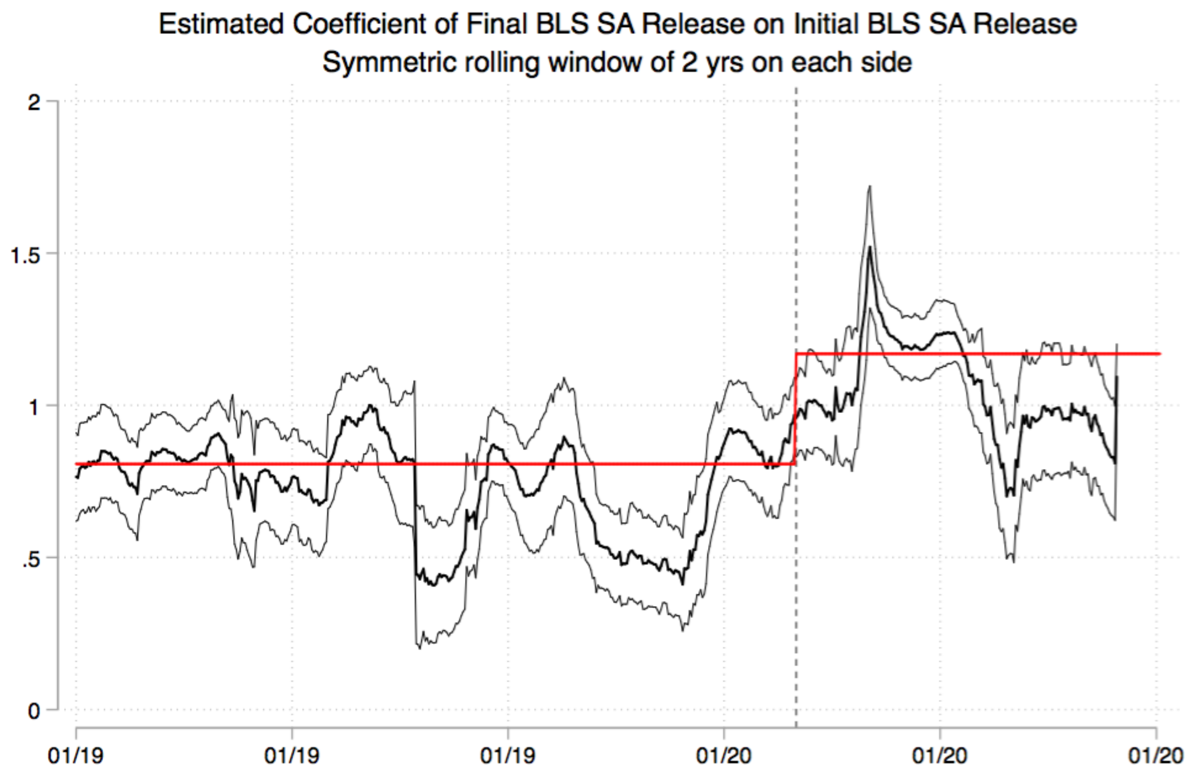


Figure 3.2: Estimated coefficients from rolling regression of Final Employment Growth on Intitially Released BLS employment growth.

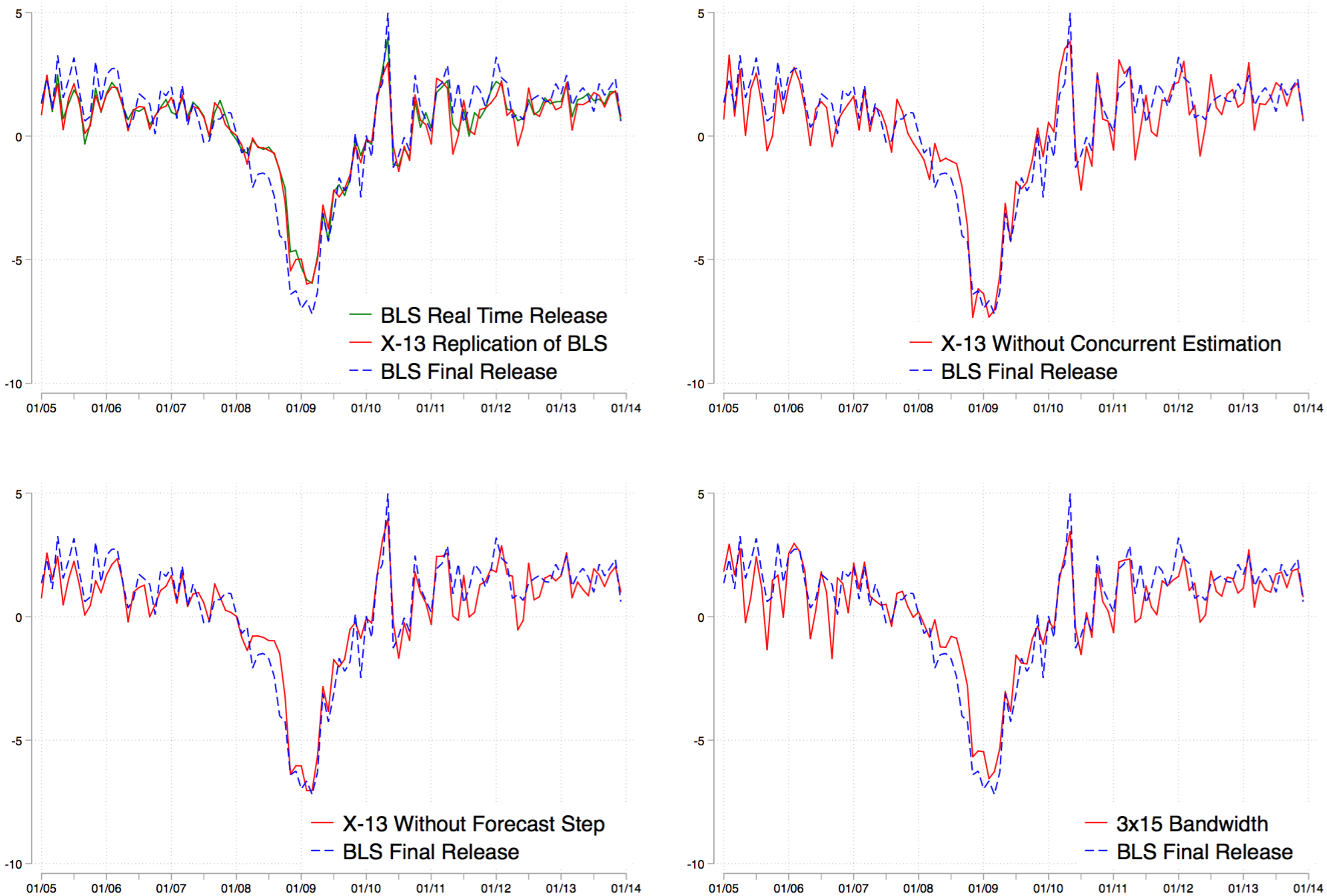


Figure 3.3: Real-time BLS employment data under alternate X-13 specifications

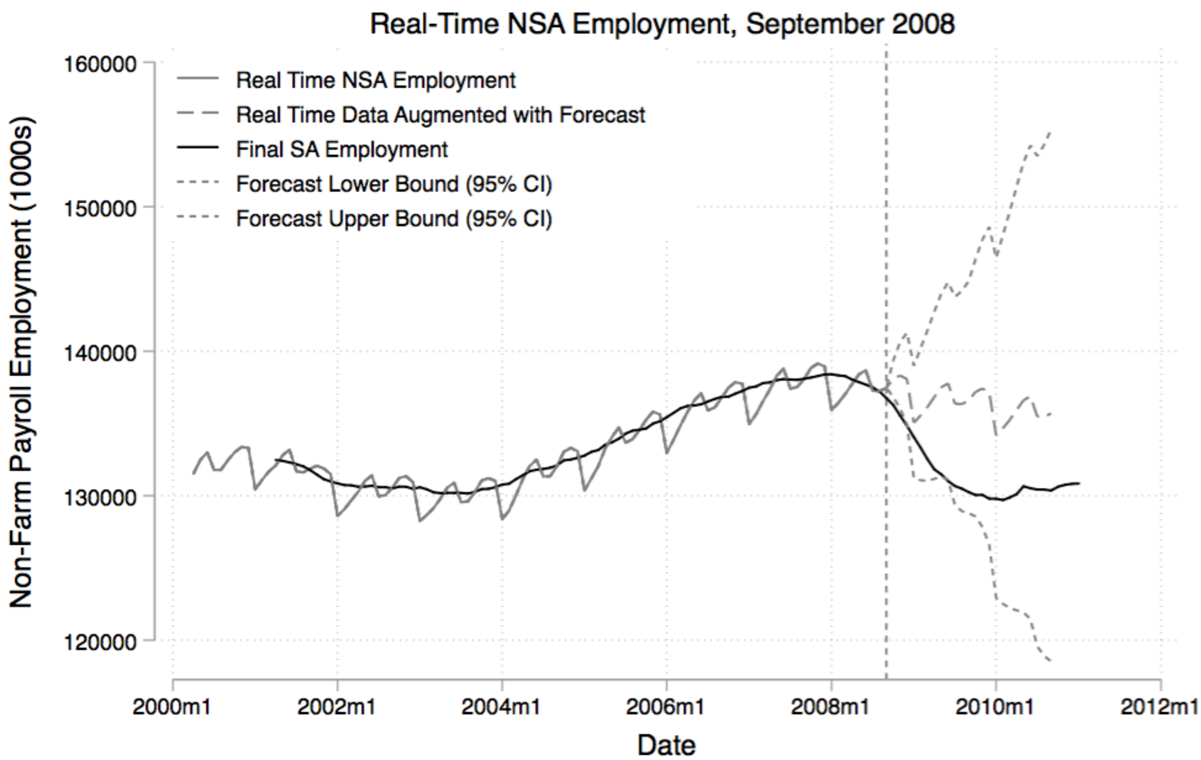


Figure 3.4: Illustration of the forecast step from the X13-ARIMA output on real-time data from September 2008.

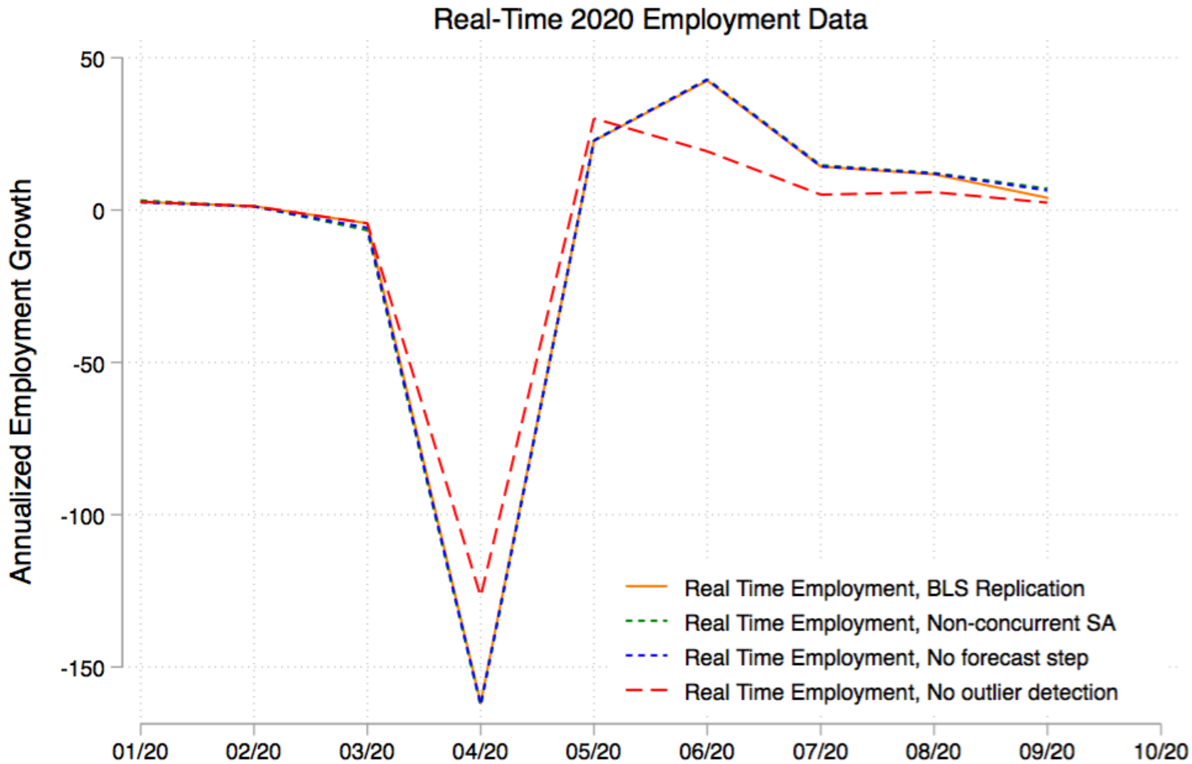


Figure 3.5: Plot showing the initially announced employment data where the seasonal adjustment method either does or does not have an outlier-detection step.

Chapter 4

Home-Production Over a Recession

This chapter of my dissertation is taken from a paper I wrote early in my graduate school career. The paper focuses on the theoretical and empirical patterns in time use within a household over the course of a recession. The paper contributes to my overall thesis by digging into some of the conventional assumptions used by macroeconomists regarding home production and its role in business cycles. I offer a more nuanced approach to understanding how individuals divide time and labor within a household and relate my empirical findings to the standard theoretical models.

4.1 Introduction

The trade off between leisure and consumption is at the heart of economic research. However, in classical macro models, the modeling of time use frequently omits the choice to consume home-produced goods or market-produced goods, and ignores the role of substitution in time use between two co-habitating income earners.

In this paper I will study how individuals adjust household production and childcare over the business cycle and how including household production in models of the business cycle affects propagation and amplification of shocks to production. In contrast to previous work, I will specifically look at how behavior differs by sex of householder. This is a novel contribution to the literature, and is an important feature to study as there has been dramatic change in the composition of the labor force historically and around the globe. This topic is particularly of interest in the wake of the COVID-19 pandemic which altered work and childcare patterns for households everywhere.

I begin by building an RBC-style model with home production and a two-earner household. In the model, the two earners maximize household utility which includes shared consumption and disutility in each earner's labor in the home and the market. The two earners are employed in different sectors that can receive correlated but distinct shocks. This framework is based on the initial RBC model with home production introduced by Benhabib, Rogerson, and Wright, 1991. My model contributes a two-income earner dimension which may be increasingly important in a world where many households have two earners and in a labor market with considerable gender segregation between sectors. This paper does not estimate the model or calibrate the model - instead I use the model to think about the dimensions of time trade-off between the two earners in the home. I view the model as a useful benchmark to begin to think about how households may allocate time.

Once establishing a model, I use micro-level time use data from the American Time Use Survey (ATUS) to learn about how the data can inform the our understanding of the model in both functional form and parameter size. Because the ATUS only goes back to 2003, I exploit regional variation over the Great Recession.

I first use the analytical framework from Aguiar, Hurst, and Karabarbounis, 2013. I disaggregate their results by gender and marital status to see if there are any large differences in how men and women substitute between activities. I find that the largest differences between men and women appear to be in how people substitute time between market work and childcare. This is intuitive and a useful fact, but their research design does not take into consideration actual labor market conditions, nor does it consider the trade-offs that may be made between two income earners in a household.

I fill the gap left by Aguiar, Hurst, and Karabarbounis, 2013 by looking at how time use adjusts to real labor market conditions, as measured by the state level unemployment rate. I find that while men's market work decreased substantially during the recession in states with high unemployment rates, women's market work did not. Similarly, men's time in home production and childcare increases with local labor market conditions whereas women's does not. This is in line with the model and reflective of the fact that industries that were hit hard during the Great Recession employed primarily men. Surprisingly, and contrary to the findings of previous researchers, there does not seem to be a strong adjustment in leisure corresponding to local labor market conditions.

Finally, I explore how time use by men and women relate, another feature that has been ignored at this point in the literature. The model predicts that time spent in home production for each earner will co-move. I find that this co-movement does not appear in the data. If anything, it appears that time spent in the home is negatively correlated between men and women. This is informative as it suggests that the model needs to be rethought so account for the fact that men and women may act as substitutes in home production.

My work contributes to a small but important literature. Including household production in real business cycle models had a brief moment of glory in the 1990s. Benhabib, Rogerson, and Wright, 1991 wrote a simple RBC model that included household production to show that relative productivity between household labor and market labor could drive movements into and out of the labor force. Greenwood, Guner, and Vandenbroucke, 2017 demonstrated further that household production plays a significantly large and important role in the macroeconomy.

Rupert, Rogerson, and Wright, 1995 use the PSID to estimate labor supply elasticities for various demographics and argue that the labor supply elasticities between market work and household work for single women and couples are high enough to justify including household production in macro models.

More recently, Karabarbounis and Neiman, 2014 includes household production in an open economy framework and showed that his model was better able to match cyclical movements in an international macro model when domestic production was included.

My empirical work is most closely related to that of Aguiar, Hurst, and Karabarbounis, 2013 who looked at time use over the Great Recession. This paper uses the American Time Use Survey similarly to how I will use the survey. They document an increase in leisure and an increase in non-market work during the recession. They do not do a careful analysis of these substitutions by gender, however.

Finally, while may paper does not currently look closely at sector-specific shocks and sectoral composition in the propagation of economic downturns, my paper dovetails with this recently growing literature. Olney and Pacitti, 2017 argue that the shift towards the service sector may explain this persistence in unemployment. Their work abstracts from the fact that the service sector employs more women than other sectors. If women substitute time into and out of the labor force differently than do men, it will be important to consider how a service economy may respond to shocks. Atalay, 2017 shows that sector-specific shocks play an important role in business cycles. Again, this work abstracts from any motion of home production or gender difference in time substitution, though it is an undeniable feature of the US labor market that sectors have a large amount of gender segregation. My hope is that eventually my work will be able to speak to changing pattern in business cycles as women became are now active in the labor force and two-income earner homes have become dominant.

The paper proceeds as follows: Section 4.2 builds the model, Section 4.3 describes the data, Section 4.4 presents the empirical work, and Section 4.5 concludes.

4.2 Model

The classic model of home production in RBC models was introduced by Benhabib, Rogerson, and Wright, 1991 and uses a simple representative consumer model with preferences of consumption and leisure that are determined by choices about work in the home and work in the labor market. Here I will extend the model slightly to include two members of the household employed in separate sectors in the labor market. The household makes a joint decision over each partner's time spent in the labor market, time spent working at home, and consumption of home/market produced goods. This model nests that of a representative agent, but includes additional dimensions of the leisure-labor trade-off to more closely model the dynamics inside a household with two income earners.

Model Set-Up: Households

Households have a husband (m) and a wife (w) who each supply labor - h_m and h_w - but in two separate sectors, sectors m and w, respectively. We can think of the sectors as manufacturing and healthcare, for example. While this simplifying assumption may seem stark and unrealistic (as workers could switch sectors), the aim of the model is to test the importance of complementarities/substitutability of agents' time spent in the home. Including separate sectors can be justified through the fact that there is large gender segregation in labor markets by industry (BLS). Recent macro literature suggests that sector-specific shocks play an important role in overall macroeconomic dynamics Atalay, 2017.

Flow utility of the household is given by

$$U(c, h_m, h_w) = \alpha \ln(C) - (1-\alpha)\lambda_1 \frac{(h_w)^{1+1/\xi_1}}{1+1/\xi_1} - (1-\alpha)\lambda_2 \frac{(W_h)^{1+1/\xi_2}}{1+1/\xi_2} \tag{4.1}$$

$$- (1-\alpha)\lambda_3 \frac{(h_m)^{1+1/\theta_1}}{1+1/\theta_1} - (1-\alpha) \frac{(M_h)^{1+1/\theta_2}}{1+1/\theta_2} \tag{4.2}$$

where $C = (aC_m^e + (1-a)C_h^e)^{1/e}$, consumption of home-produced goods and market-produced goods. Market-produced goods are bought by the household with their labor income. Home-produced goods are produced in the home with time from both partners under Cobb-Douglas production

$$C_h = Z_h W_h^{\nu} M_h^{1-\nu} \tag{4.3}$$

where W_h is the wife's time working at home and M_h is the husband's time working in the home.

Whatever time is left over after market work and home production work is allocated towards leisure. For example, the wife's leisure time is given by $1 - h_w - W_h$.

The modeling of the household closely follows that used in Benhabib, Rogerson, and Wright, 1991. In my model, like in Benhabib, Rogerson, and Wright, 1991, total household consumption is a CES aggregate of market-bought consumption and home-produced consumption. Unlike Benhabib, Rogerson, and Wright, 1991, my model uses disutility of labor in the home and in the market, rather than modeling utility in leisure. The Benhabib model uses a log-log utility in leisure functional form which abstracts from any differences in the elasticity of substitutions between leisure and market work and between leisure and home production. My model therefore generalizes the Benhabib model to try to match more closely the observed elasticities in data.

Finally, production in the home is modeled as Cobb-Douglas function for simplicity. However, this choice is not inconsequential as the elasticity of substitution between spousal time spent in the home will be a key question in the data. Unfortunately in this paper I am not able to fully develop a theory for how we should model elasticity of substitution between spouses' work in the home. This is a preliminary model to begin thinking about the issue.

Model Set-Up: Firms

On the firm side, there are two sectors each with Cobb-Douglas production.

$$Y_{m,t} = Z_{m,t} K^{\gamma}_{m,t-1} h^{1-\gamma}_{m,t} \tag{4.4}$$
$$Y_{w,t} = Z_{w,t} K^{\eta}_{w,t-1} h^{1-\eta}_{w,t} \tag{4.5}$$

The choice to use two different sectors is again a simple way of accounting for the fact that the US labor market displays a heavy amount of gender segregation. I am not attempting to explain this fact - I simply want to create a model that can generate some of the business cycle dynamics that we may expect in a world where there are two separate sectors employing household members. Shocks to each sector may be highly correlated. Further, the labor share of income in each sector may be different. Again this is a dimension of flexibility in the model that I do not attempt to pin down with data, but that I think offers an interesting view into how business cycle dynamics may have changed over time as there have been long historical sectoral shifts in US labor markets.

To close the model we have the normal conditions:

$$Y_t = Y_{m,t} + Y_{w,t} = C_{m,t} + I_{m,t} + I_{w,t} \tag{4.6}$$
$$K_{m,t} = (1-\delta)K_{m,t-1} + I_{m,t} \tag{4.7}$$
$$K_{w,t} = (1-\delta)K_{w,t-1} + I_{w,t} \tag{4.8}$$
$$Z_{m,t} = e^{\epsilon_{m,t}} Z^{\rho_m}_{m,t-1} \tag{4.9}$$
$$Z_{w,t} = e^{\epsilon_{w,t}} Z^{\rho_w}_{w,t-1} \tag{4.10}$$
$$Z_{h,t} = e^{\epsilon_{h,t}} Z^{\rho_h}_{h,t-1} \tag{4.11}$$

Model Implications

Agents optimize according to the Lagrangian given by

This gives us Lagrangian

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \Bigg[\alpha \ln(C) - (1-\alpha)\lambda_1 \frac{(h_w)^{1+1/\xi_1}}{1+1/\xi_1} - (1-\alpha)\lambda_2 \frac{(W_h)^{1+1/\xi_2}}{1+1/\xi_2} \tag{4.12}$$
$$- (1-\alpha)\lambda_3 \frac{(h_m)^{1+1/\theta_1}}{1+1/\theta_1} (1-\alpha) \frac{(M_h)^{1+1/\theta_2}}{1+1/\theta_2} \tag{4.13}$$
$$+ \mu_{1,t}[Z_{m,t} K^{\gamma}_{m,t-1} h^{1-\gamma}_{m,t} + Z_{w,t} K^{\eta}_{w,t-1} h^{1-\eta}_{w,t} - C_{m,t} - I_{m,t} - I_{w,t}] \tag{4.14}$$
$$+ \mu_{2,t}[(1-\delta)K_{m,t-1} + I_{m,t} - K_{m,t}] \tag{4.15}$$
$$+ \mu_{3,t}[(1-\delta)K_{w,t-1} + I_{w,t} - K_{w,t}] \tag{4.16}$$
$$+ \mu_{4,t}[Z_{h,t} W^{\nu}_{h,t} M^{1-\nu}_{h,t} - C_{h,t}] \Bigg] \tag{4.17}$$
$$\tag{4.18}$$

The full derivation of the first order conditions and log-linearization around the steady state are presented in the Appendix.

Using log-linearization equations guides the business cycle dynamics of the system. We can see that the model generates very strong predictions regarding time spent working in the home and time spent in the labor force. Combining the two labor-supply conditions for each spouse results in the equation:

$$(1/\xi_1)\check{h}_{t,w} - (1/\theta_1)\check{h}_{t,m} = \check{Y_{w,t}} - \check{h_{w,t}} - (\check{Y_{m,t}} - \check{h_{m,t}}) \tag{4.19}$$

$$\implies (1 + 1/\xi_1)\check{h}_{t,w} = (\check{Y_{w,t}} - \check{Y_{m,t}}) + (1 + 1/\theta_1)\check{h}_{t,m} \tag{4.20}$$

This equation implies that, conditional on the relative production in each sector staying fixed, spousal labor market decisions will move together (positive correlation).

Similarly, combining the home-production-supply curves of the two spouses results in the relationship:

$$(1/\xi_2)\check{W}_h - (1/\theta_2)\check{M}_h = -\check{W}_{h,t} + \check{M}_{h,t} \tag{4.21}$$

$$\implies (1 + 1/\xi_2)\check{W}_h = (1 + 1/\theta_2)\check{M}_h \tag{4.22}$$

Again, this equation implies that spousal time spent on home production will move with positive correlation. While this may seem a simple derivation, it also may contradict the intuition that spouses could would substitute time on home production with each other. Each of these results depends crucially on the fact that in the model, spouses share consumption completely (consumption is a public good within the household).

Model Business Cycle Dynamics

The log-linearized form of the model is presented in the appendix. To see how the various parameters effect market and home production hours for each spouse, I present simulated impulse response functions.

Figure 4.1 presents impulse response functions to a negative shock to each sector (separately) for when 'men' and 'women' have completely symmetric behavior both in the home and in the workplace. For a negative shock to either sector, market consumption declines and home-produced consumption increases. This is an intuitive result (thought not one that I will test in data, since I don't have consumption data). As anticipated, a shock to the men's sector will decrease men's time in the labor market but increase women's time in the labor market (and vice versa). However, a negative shock to either sector will increase the time that both spouses spend in home production. This illustrates that the model predicts that the time each spouse spends in home production should be highly positively correlated with the other spouse.

Figures 4.2 and 4.3 show how varying the parameters of the model impact the severity and longevity of the shock. Figure 4.2 shows that the elasticities of substitution between labor (both home and market) and consumption guide the magnitude of the shock to labor market hours and home production hours. Figure 4.3 demonstrates that if the labor market sectors are different in how much income goes to labor, then which sector the shock occurs in will have different effects on total time family members spend in home production and market production.

In none of these models do we see the spouses substitute time against each other in the home. In the next section, we test this prediction.

4.3 Data

The main data source used in this paper is the American Time Use Survey. The data include over 110,000 observations of individuals between 2003 and 2015. For each observation the data includes detailed information about the demographics of the individual in question and detailed information about time use for a single day of that individual's life. Time use information is categorized in fine detail for each minute of an individual's day.

In accordance with work by Aguiar, Hurst and Karabarbounis, I limit the sample to individuals between ages 18 and 65 (working age) and remove individuals who report any 'unclassified' time. Time use categories are then aggregated as in Aguiar, Hurst and Karabarbounis (2013) [1]. The broadest mutually exclusive categories are market work, other income generating activities, job search, childcare, non-market work, leisure, and other. Within non-market work, leisure and other are a number of sub-categories (such as core home production, sleep, tv watching, education, etc.). For information on what each of these categories encompasses, see Aguiar, Hurst, Karabarbounis (2013).

Table 4.1 presents summary statistics for men and women by their marital status and labor force status. Average weekly hours spent on each type of time use are computed taking into consideration the sampling weights provided by the ATUS data (through the BLS). The summary statistics give us a first order sense of what people spend their time doing. These statistics span the entire sample, thus do not account for trends over time. Instead, it's meant to show the differences between groups to get a sense of how different people use time and why that might be important for business cycle dynamics.

While of course no causal inference can be made about how a person substitutes time between categories through this chart, it is interesting to not the variation in time use around child care and leisure between the employed and unemployed group. While unemployed women on average spend a greater amount more time taking care of children than employed women, unemployed men spend somewhat less time taking care of children relative to women and instead seem to spend more time on leisure and education.

[1]The online Appendix from Aguiar, Hurst and Karabarbounis (2016) contains full documentation on how time use categories are constructed. The appendix can be found at this link: `https://scholar.princeton.edu/maguiar/publications/time-use-during-great-recession`

4.4 Empirical Results

The remainder of this paper revolves around understanding how time use data from the American Time Use Survey can inform our understanding of the trade-off between market work, home production, and leisure in the US. The data only covers the period between 2003-2015, thus only contains information about one business cycle. While this is not ideal for studying time use over a business cycle, I harness geographical variation to study how time use adjusts over the business cycle.

AHK Empirical Model

First, I employ the empirical model used by Aguiar, Hurst and Karabarbounis (2013). These authors use the same dataset that I use, the American Time Use Survey. At the time that their paper was written, they only had access to data between 2003 and 2010. Thus, my replication extends the dataset until 2015 which allows us to see whether their results hold during a recovery period.

The empirical model that the authors use to test time use substitutions during the great recession is given by

$$\Delta\tau_{st}^{j} = \alpha^{j} + \beta^{j}\Delta\tau_{st}^{market} + \epsilon_{st} \tag{4.23}$$

where τ denotes time use (in minutes) and j denotes a category of time use (such as 'leisure' or 'home production'). This model exploits state-year level variation in aggregate time use after taking weighted averages of individual-level time use in a state-year. The authors use a first-difference model in order to better control for state-level differences in time use that might generate a correlation in aggregate time use that is not attributable to time use substitution at the individual level. Thus the authors argue that their model produces an unbiased estimate of time-use substitution between time use categories - specifically how individuals adjust time use with a change in the number of hours they spend in market work.

I replicate their results for a number of sub-populations to see if we could observe any differences by population. For example we may expect married women to substitute more of their foregone market work time into home production or childcare compared to single men. In fact, we see fewer differences in time use substitution than we might expect. Tables 4.2-4.4 present the replicated results using data from 2003 to 2015. Each column represents a sample restriction before aggregating to the state-year level. Aggregate state-year observations are weighted by population size in the reported regression to account for variance in the averages stemming from smaller samples.

Table 4.2 shows that substitution between market and non-market work (home production excluding child care) is remarkably similar across demographics with women substituting slightly more than do men into home production for each hour of market work foregone. Table 4.3 shows that for childcare there are large differences between how men and women substitute between market work and childcare though in all populations the trade-off is estimated to be small (around 3 minutes more of childcare for every one hour of market work foregone) in the aggregate. Finally, Table 4.4 shows that for both men and women, married and single, the largest component of time saved when labor market hours are decreased is allocated towards leisure.

These results shed some light on how people substitute between activities in their own life. However, the research design employed by Aguiar, Hurst and Karabarbounis does not offer any structural interpretation for the estimates, nor do they pair their estimates with any sort of local labor market conditions to see how time use actually responds to economic conditions. Instead, we see how people have substituted time in the aggregate, over the last 15 years. This does not necessarily correspond to how time use changes over the business cycle.

Local Unemployment and Time Use

To get at how time use adjusts over the business cycle, we begin by simply observing how regional variation in the unemployment rate corresponds to time use.

The first model I use to get at this uses the ATUS data at the individual level and looks at how time use in each category varies with an individual's state-level unemployment rate. The idea is that the state unemployment rate captures the health of the local labor market. This is obviously an imperfect measure of local labor market conditions, but it serves the purpose of showing how and individual's time use may vary with the health local economy. The model is given by

$$\tau^j_{i,s,t} = \beta_1 UR_{s,t} + \beta_2 female_i + \beta_3 UR_{s,t} \times female_i + \beta_4 age_i + \gamma_t + \eta_s + \epsilon_i \tag{4.24}$$

where $\tau^j_{i,s,t}$ is time (in minutes) spent on activity j by individual i in year t and state s. The variable $UR_{s,t}$ is the state unemployment rate on average in year t, and $female_i$ is a dummy variable indicating whether individual i self-reported as female. The regression includes state and year fixed effects to absorb any trend over time and any regional differences that might bias the results. I control for age of the respondent because time use varies significantly over the life cycle.

We see in Table 4.5 the results from this model. Column 1 demonstrates that market work for men is significantly lower in states with higher unemployment rates. In contrast, women's market work does not seem to be lower. Column 2 demonstrates that higher state unemployment rates correspond with more time spent working in home production for men, but not for women. Columns 3 and 4 show less significant results, but also hint at a difference in time use between men and women around places that had larger labor market shocks.

These results are specific to the Great Recession, which impacted male employment much more severely than it did female employment. The COVID-19 recession would provide a setting to evaluate intra-household time use decisions in a recession that disproportionately impacted women's employment.

To try to test how time use of individuals responded over time to the Great Recession unemployment shock, I regress time use categories by individual on a series of dummies for the year and whether or not there was a large unemployment shock to the state in 2009. This method has been used by Yagan, 2017 and Blanchard and Katz, 1992. As a first step, I use an AR(5) process to model the time path of the unemployment rate for each state. Using the estimated model, I generate prediction errors for each period. I focus specifically on the prediction error in 2009. I then take the top 25 percent of states based on the size of the model prediction error in 2009 and call these the 'shocked' states. I take the bottom 25 percent of states based on the size of the model prediction error in 2009 and call these the 'unshocked' states. I limit the sample to individuals who are in one of these two groups of states, and run the following regression:

$$\tau^{j}_{i,t,s} = \alpha + \beta_1 1(s \in shock) + \sum_{y=2006}^{2015} \beta_y 1(y = t) + \sum_{z=2006}^{2015} \beta_z 1(z = t) 1(s \in shock) + \epsilon_{i,t,s} \quad (4.25)$$

where $\tau^{j}_{i,t,s}$ is time use in minutes of category j time use for individual i who reported time use when living in state s and year t. The indicator function $1(s \in shock)$ will be 1 when state s is one of the states with a high unemployment shock in 2009 and 0 otherwise.

The coefficients β_z tell us how time use for individuals in 'shocked' states evolved differently from individuals who were in 'unshocked' states over the sample period.

Figures 4.4-4.7 present graphical representations of the data over the Great Recession period. Figure 4.4 demonstrates the unemployment rate over time in the states that I've categorized by 'high UR shock' and 'low UR shock.' The figure demonstrates that the states have closely linked parallel trends before 2009 but that the 'high shock' states have higher jumps in their unemployment rates starting in 2009.

Figure 4.5 demonstrates that the rate of unemployment within the sample of ATUS participants (limited to the two groups of states based on the unemployment shock) looks a bit different from the aggregate unemployment rates. Reporting of unemployment among ATUS respondents begins to rise in 2007 for shocked states relative to unshocked states. This is a bit different from at the aggregate macroeconomic level and perhaps speaks to imperfect sampling.

Figure 4.6 shows that the number of participants reporting that they worked 0 hours in their diary day looks fairly different from and much larger than the rate of unemployed ATUS participants. This is because there are a large number of people out of the labor force and the ATUS data includes weekends (which I do not drop because some people work on the weekends). It is a bit unclear what is going on in the sample of people living in states without large unemployment shocks. Overall this plot tells us that the number of unemployed people in the dataset is very marginal relative to the number of people not working for other reasons.

Figure 4.7 plots average time spent on market work for the two groups of states. The difference between the two groups is primarily driven by the number of people reporting zero market hours (as shown in Figure 4.6).

These plots show why research with this data is difficult and sensitive to specification. The large swing in market hours that we see in the data does not align with the Great Recession, as we'd expect it to. Nonetheless, I proceed with formal econometric tests to confirm the results.

Table 4.6 shows the formal tests that correspond to Figures 4.4-4.7. I report the coefficients β_z from the regression model presented above. Column 1 of Table 4.6 shows that the difference in the path of the unemployment rate for the two groups of states (as depicted in Figure 4.4) is statistically significant and begins in 2009 as desired. This is simply a validation of the test I designed. Column 2 corresponds to Figure 4.5 and shows that the increase in reported unemployment in the ATUS sample begins earlier than the Great Recession - an undesired result. Similarly, Column 3 (corresponding to Figure 4.6) shows that the number of people reporting 0 labor market hours is larger for the shocked states starting before the Great Recession shock. Finally, as anticipated, the average reported time spent working in the labor market is lower before the Great Recession shock and also significantly lower (by 40 minutes) in 2010, after the shock.

Table 4.7 reports results from the same regression model run on other time use categories. Column 1 of Table 4.7 limits the sample to only those people who report positive labor market hours. This test corresponds to how the intensive margin adjusted during the Great Recession and shows no significant results. Column 2 reports the test using Non-Market Work (primarily home production excluding child care) and shows that home production was slightly elevated in 2008 in states with large labor market shocks. However this slight elevation does not persist and is not located at the precise timing of the unemployment shock, thus it is not a strong result. Similarly, child care looks slightly elevated in 2008 for states with large unemployment shocks. However, the result does not persist and we do not see the elevation occur when unemployment is peaking. Thus the results do not seem like compelling evidence of time substitution. Running the test on leisure time results in even weaker and less informative results.

These results are surprising and perhaps due to the noisy nature of the ATUS data. While Aguiar, Hurst and Karabarbounis were able to make some strong statements about time susbtitutions of individuals using this data, the data does not present strong evidence that people truly shift their time a lot during the Great Recession.

Empirical Tests of the Model

As described in Section 4.2, my model of labor/leisure tradeoffs in a household predicts that spousal time spent in home production will be positively correlated. Given that the vast majority of married-couple households in the US are heterosexual (US Census, 2010), I test this prediction of the model by observing the aggregate movements of men's time spent in home production against the aggregate movements of women's time spent in home production. To do this, I employ the methods used by Aguiar, Hurst and Karabarbounis to aggregate time use data to the state-year level. In this case, I keep the data disaggregated by gender to test the following model, derived from my RBC-style model described in Section 2:

$$(1/\xi_2)\check{W}_h - (1/\theta_2)\check{M}_h = -\check{W}_{h,t} + \check{M}_{h,t} \tag{4.26}$$

$$\implies \ln(\tau^j_{women,s,t}) = \alpha + \frac{1+1/\theta_2}{1+1/\xi_2}\ln(\tau^j_{men,s,t}) + \epsilon_{s,t} \tag{4.27}$$

where $\tau^j_{women,s,t}$ is the average time spent per day on activity j by women in state s and year t. I test this model for childcare and all remaining home production separately since we saw previously that time use in these two activities is substituted into differently by men and women.

Table 4.8 tests the model looking specifically at time spent on childcare. We see that, contrary to the prediction of the model, there appears to be no strong relationship between women's time spent on childcare and men's time spent on childcare. In fact, the sign on the estimated coefficient is negative for each variant of the model, suggesting that, if anything, men and women seem to be substituting time against each other rather than adjusting their time spent taking care of children in sync with one another. This finding motivates a significant change in our modeling of intra-household work allocation.

Table 4.9 presents a similar result when looking at home production excluding child care. In fact, the fixed effects models presented in columns 2 and 4 of Table 4.9 both show a significant negative correlation between men's time spent in home production and women's time spent in home production. Again, this presents evidence against the prediction of the classical model presented in this paper which predicts that time use in home production for two spouses should co-move. Overall these results suggest that the model of time use in home production presented in Section 4.2 does not correctly capture the time use allocations that we see in data.

Table 4.10 reports the results for running a similar regression, but instead looking at labor market time use by men and women. The model presented in Section 2 predicts that market work of the two spouses should evolve according to the relationship

$$(1+1/\xi_1)\check{h}_{t,w} = (\check{Y_{w,t}} - \check{Y_{m,t}}) + (1+1/\theta_1)\check{h}_{t,m} \tag{4.28}$$

$$\implies \check{h}_{t,w} = \frac{1}{(1+1/\xi_1)}(\check{Y_{w,t}} - \check{Y_{m,t}}) + \frac{(1+1/\theta_1)}{(1+1/\xi_1)}\check{h}_{t,m} \tag{4.29}$$

I test this relationship by running the regression model

$$\ln(\tau^{market}_{women,s,t}) = \alpha + \beta_1\ln(Y_{w,t}) + \beta_2\ln(Y_{m,t})) + \beta_3\ln(\tau^{market}_{men,s,t}) + \epsilon_{s,t} \tag{4.30}$$

where $\tau^{market}_{women,s,t}$ is the average time (in minutes) that women in state s and year t spent on labor market work and $Y_{w,t}$ is the output level averaged over the reported industries of the women respondents in the ATUS data in state s and year t.

The results are reported in Table 4.10 and show that while women's work is responsive to the output level in women's industries (as one would expect based on our model and our general notion of wage setting and labor supply), women's hours in the labor market are not responsive to men's hours in the market, nor the output level of industries that men work in. This again poses a challenge to the model that I presented in Section 4.2. And begs the question of how we may better model home production and leisure in macro models when we have two working members of the household.

4.5 Conclusion

This paper offers some preliminary ideas about how to better model households and home production in macroeconomic models. The paper is motivated by the idea that the changing demographics of the labor force in terms of gender and household composition and the sectoral shifts in the US must have a role in the changing nature of business cycles. Especially in the wake of the COVID-19 recession which uprooted our childcare sector and altered the nature of work, intra-household time use decisions will contribute increasingly to labor supply channels.

The paper builds a simple model which can be used as a benchmark to think about the dimensions of time use and household time use and consumption decisions that matter and that may effect business cycle dynamics.

The paper then demonstrates some empirical results that show that the time use data available in the US is noisy and hard to learn from, but that there does appear to be substitution between market work and home production during a recession. The data also show that the modeling of substitution between men and women's time in the home is incorrect - instead of seeing co-movement between time spent in the home, there is practically no relationship at all.

Finally, this paper serves as a starting point to think further about home production and sector specific shocks in macro models. In the future, the hope would be to think more carefully about how these household decisions translate into macroeconomic dynamics and to better nest this work in the theory of labor supply and business cycles.

4.6 Tables and Figures

Table 4.1: Average Weekly Hours For Regular Activities

	Employed				Unemployed				Out of LF			
	Single		Married		Single		Married		Single		Married	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Market Work	43.9	39.0	45.8	35.9	0.7	0.5	0.7	0.5	0.3	0.1	0.4	0.2
Other Income Generating Activities	0.1	0.1	0.1	0.2	0.6	0.8	0.3	0.2	0.2	0.3	0.2	0.3
Job Search	0.1	0.1	0.1	0.1	4.3	3.0	4.3	2.4	0.2	0.1	0.2	0.0
Child Care	1.1	2.9	3.2	5.0	2.1	7.0	4.1	7.2	1.3	5.9	2.5	9.6
Non-Market Work	11.5	15.9	12.6	19.4	16.8	23.9	20.2	29.4	14.6	22.2	18.3	30.5
–Home Production	5.2	9.0	5.0	11.7	7.7	15.1	9.2	19.2	7.2	14.6	7.6	20.1
–Home Ownership Acitivies	1.7	1.0	3.1	1.3	3.2	1.1	4.6	1.8	2.7	1.4	4.9	2.2
–Obtaining Goods and Services	3.9	5.4	3.9	5.8	4.2	6.1	5.1	7.4	3.6	5.0	4.6	6.9
–Others Care	0.7	0.6	0.6	0.7	1.7	1.6	1.3	1.1	1.1	1.2	1.1	1.3
Leisure	106.9	103.9	101.3	101.0	134.5	120.9	128.0	118.7	140.5	127.9	134.9	117.1
–TV Watching	17.8	14.8	16.1	12.8	30.3	22.6	24.7	20.4	36.4	28.0	31.8	20.0
–Socializing	7.5	7.1	6.1	7.0	13.6	9.9	13.3	10.9	10.1	9.9	9.7	9.5
–Sleep	59.3	59.9	57.0	58.7	66.4	66.5	65.2	64.8	67.3	67.6	64.6	63.2
–Eating and Personal Care	12.8	13.8	13.3	14.4	10.8	11.9	11.6	12.5	10.9	11.7	12.5	13.6
–Other Leisure	9.6	8.3	8.8	8.1	13.4	10.0	13.3	10.2	15.9	10.8	16.4	10.9
Other	2.9	4.3	3.4	4.5	6.0	9.1	7.4	7.0	9.7	9.8	9.9	7.7
–Education	1.5	2.0	1.0	1.5	3.8	5.2	4.4	3.7	5.5	4.3	4.5	2.6
–Civic	1.1	1.5	1.9	2.2	1.6	2.7	2.6	2.5	1.7	2.5	2.7	3.3
–Own Medical	0.4	0.8	0.5	0.8	0.6	1.2	0.4	0.8	2.5	2.9	2.7	1.9
Obs.	15665.0	20308.0	27067.0	23585.0	1315.0	1765.0	1194.0	1360.0	3779.0	6328.0	3224.0	9797.0

Table 4.2: Change in Non Market Work

	(1) All	(2) Women	(3) Men	(4) Single W.	(5) Single M.	(6) Married W.	(7) Married M.
Δ marketwork	-0.238***	-0.290***	-0.206***	-0.246***	-0.181***	-0.290***	-0.253***
	(-10.28)	(-11.08)	(-11.29)	(-10.69)	(-9.06)	(-9.73)	(-10.85)
N	663	663	663	661	663	661	659

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.3: Change in Childcare

	(1) All	(2) Women	(3) Men	(4) Single W.	(5) Single M.	(6) Married W.	(7) Married M.
Δ marketwork	-0.0459***	-0.0556***	-0.0180*	-0.0386**	-0.00469	-0.0918***	-0.0260**
	(-3.83)	(-3.67)	(-2.28)	(-3.06)	(-0.80)	(-6.24)	(-2.94)
N	663	663	663	661	663	661	659

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.4: Change in Leisure

	(1) All	(2) Women	(3) Men	(4) Single W.	(5) Single M.	(6) Married W.	(7) Married M.
Δ marketwork	-0.599***	-0.560***	-0.654***	-0.612***	-0.694***	-0.492***	-0.604***
	(-20.35)	(-15.55)	(-25.97)	(-22.59)	(-30.19)	(-16.22)	(-19.53)
N	663	663	663	661	663	661	659

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.5: Time Use and State Unemployment by Sex (with state and year FEs)

	(1) Mkt Wk	(2) NonMkt Wk	(3) Child Care	(4) Leisure
UR	-4.888**	1.908*	0.101	1.087
	(1.628)	(0.891)	(0.336)	(1.331)
female	-117.0***	76.96***	31.66***	-8.578
	(8.818)	(3.935)	(1.955)	(7.518)
UR x female	3.945**	-1.671**	-0.888**	-1.014
	(1.272)	(0.495)	(0.258)	(1.303)
age	-0.334*	1.745***	-0.800***	0.243
	(0.166)	(0.0541)	(0.0311)	(0.175)
Observations	115387	115387	115387	115387

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.6: States that Suffered Worse/Better UR Shocks: Validation Test

	(1) UR	(2) UR, ATUS	(3) Mkt Wrk==0	(4) Mkt Wrk
2004 × shock==1	-0.0566	0.0172*	0.0486*	-21.01
	(0.123)	(0.00921)	(0.0252)	(17.12)
2005 × shock==1	-0.211	-0.00830	0.0544**	-30.49*
	(0.234)	(0.0120)	(0.0198)	(15.63)
2006 × shock==1	-0.00237	0.00466	0.0751***	-48.07***
	(0.308)	(0.0107)	(0.0231)	(16.56)
2007 × shock==1	0.170	0.0193***	0.0929***	-47.14**
	(0.289)	(0.00610)	(0.0237)	(17.73)
2008 × shock==1	0.517*	0.0325***	0.0898***	-47.82**
	(0.301)	(0.0108)	(0.0287)	(19.54)
2009 × shock==1	2.462***	0.0179	0.0166	-9.805
	(0.428)	(0.0163)	(0.0438)	(32.50)
2010 × shock==1	1.823***	0.0241*	0.0832***	-39.29**
	(0.383)	(0.0122)	(0.0268)	(17.34)
2011 × shock==1	1.317***	0.0122	0.0298	-28.78
	(0.434)	(0.00992)	(0.0308)	(19.25)
2012 × shock==1	0.747*	0.00237	0.0590**	-22.12
	(0.436)	(0.0113)	(0.0255)	(17.17)
2013 × shock==1	0.581	0.0139	0.0622**	-38.87**
	(0.351)	(0.0166)	(0.0225)	(15.41)
2014 × shock==1	0.0327	0.0134	0.0669**	-41.48**
	(0.386)	(0.0122)	(0.0320)	(19.18)
2015 × shock==1	-0.159	-0.00716	0.0118	-19.73
	(0.441)	(0.0113)	(0.0281)	(17.02)
Observations	49259	49259	49259	49259

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.7: States that Suffered Worse/Better UR: Comparing Time Use

	(1) Mkt Wrk (Mkt Wrk != 0)	(2) NonMkt Wrk	(3) Child Care	(4) Leisure
2004 × shock==1	5.439	11.44	1.180	8.536
	(14.51)	(9.363)	(4.155)	(18.54)
2005 × shock==1	-6.484	1.339	4.315	18.56
	(18.80)	(6.185)	(3.304)	(15.65)
2006 × shock==1	-18.04	13.93	4.993*	27.99*
	(14.78)	(8.882)	(2.817)	(15.12)
2007 × shock==1	-2.053	-1.157	4.531	33.85*
	(18.37)	(6.587)	(3.316)	(18.74)
2008 × shock==1	-6.442	13.74*	8.075**	14.73
	(14.82)	(7.389)	(3.855)	(19.46)
2009 × shock==1	-2.842	-4.765	4.700	11.97
	(30.80)	(10.70)	(3.480)	(24.49)
2010 × shock==1	4.529	-5.892	6.005	32.96*
	(13.49)	(8.815)	(4.531)	(16.24)
2011 × shock==1	-26.42	4.549	1.928	20.95
	(17.39)	(9.085)	(4.006)	(25.93)
2012 × shock==1	14.05	4.136	13.53***	11.31
	(17.12)	(7.642)	(4.290)	(22.70)
2013 × shock==1	-14.54	8.109	6.646*	12.38
	(15.96)	(8.455)	(3.799)	(16.72)
2014 × shock==1	-13.70	15.88*	11.26***	12.42
	(18.91)	(8.577)	(3.576)	(19.21)
2015 × shock==1	-26.37	-9.936	6.219	16.21
	(19.83)	(6.118)	(4.092)	(16.21)
Observations	22078	49259	49259	49259

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.8: Men vs. Women's Time spent on Childcare

	(1) FD, levels	(2) FE, levels	(3) FD, logs	(4) FE, logs
Childcare, men	-0.0136	-0.0222	0.00487	-0.0222
	(0.0661)	(0.0656)	(0.0289)	(0.0656)
Observations	612	663	592	663

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.9: Men vs. Women's Time spent on Home Production

	(1) FD, levels	(2) FE, levels	(3) FD, logs	(4) FE, logs
Home prod, men	-0.0981	-0.144**	-0.0427	-0.0604*
	(0.0538)	(0.0546)	(0.0256)	(0.0259)
Observations	612	663	610	662

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.10: Men vs. Women's Time Spent in Mkt Work

	(1) FD, levels	(2) FE, levels	(3) FD, logs	(4) FE, logs
Y, women	0.000606**	0.000515**	0.360***	0.298***
	(0.000198)	(0.000181)	(0.0863)	(0.0801)
Y, men	-0.0000599	-0.0000555	-0.0597	-0.0373
	(0.000182)	(0.000141)	(0.0853)	(0.0773)
Mkt. Work, men	0.0487	0.0480	0.0768	0.0617
	(0.0380)	(0.0362)	(0.0523)	(0.0499)
Observations	510	561	510	561

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figures

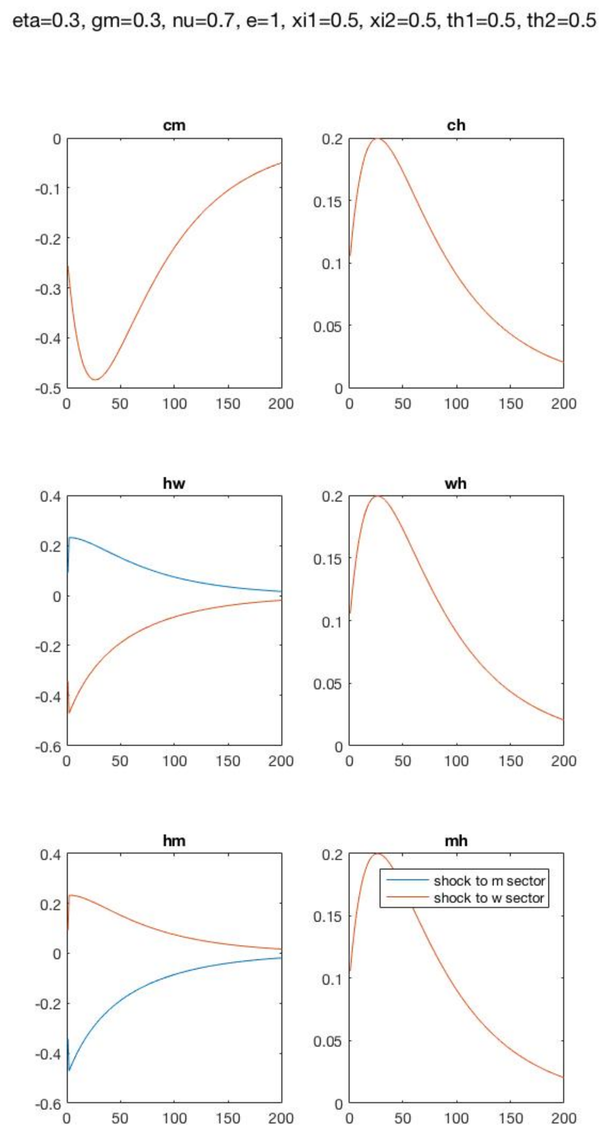


Figure 4.1: Simulated model with symmetric spouse behavior

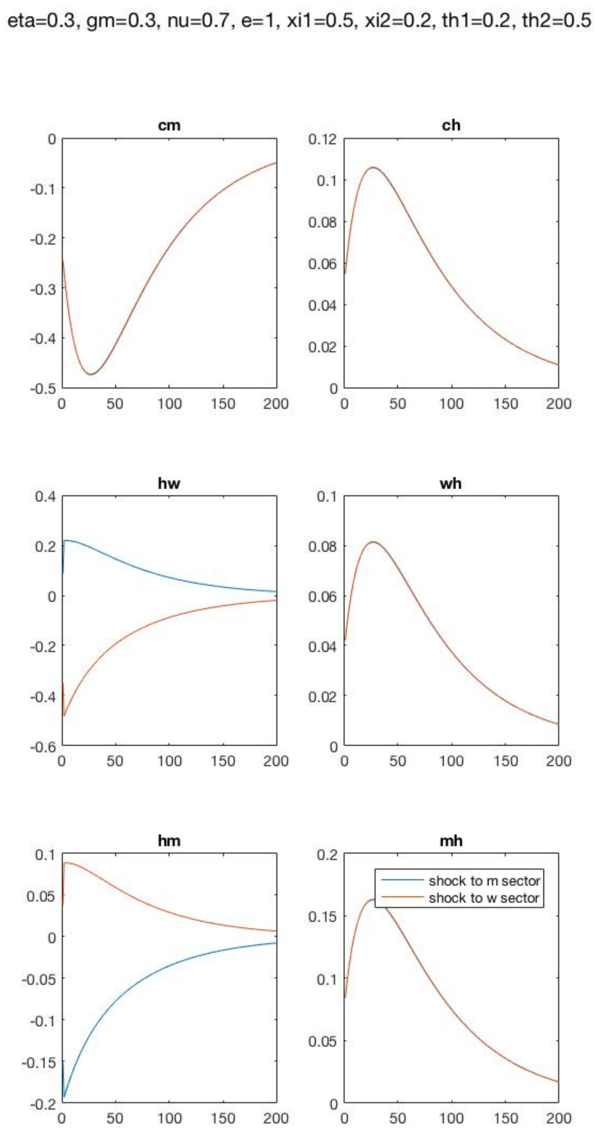


Figure 4.2: Simulated model, men have higher elasticity of substitution in home and women have higher elasticity of substitution in market

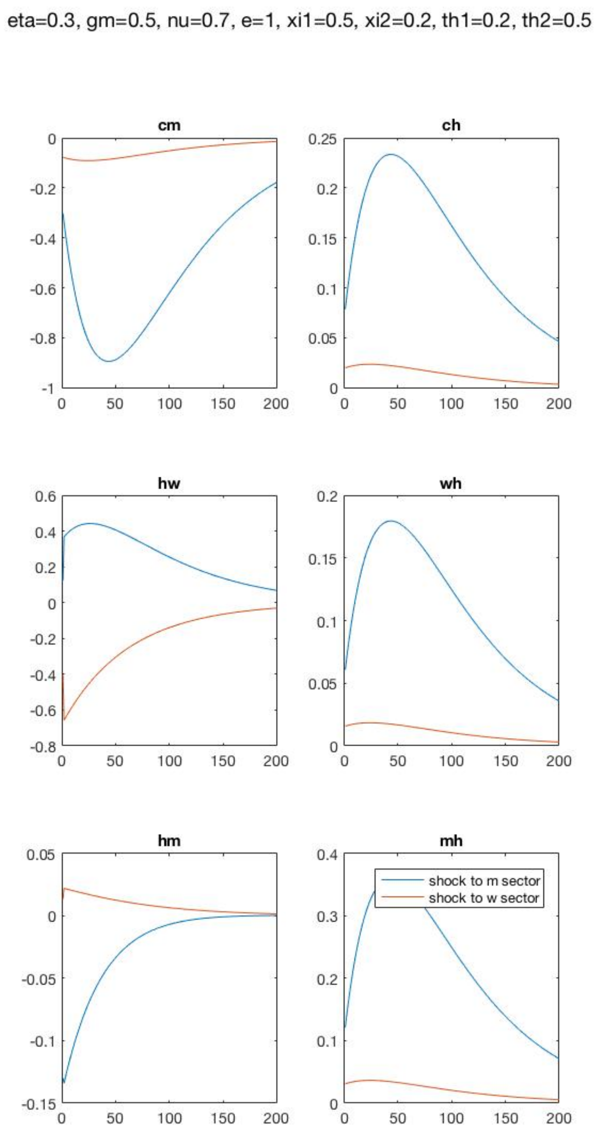


Figure 4.3: Simulated model, women's sector has higher labor share of income

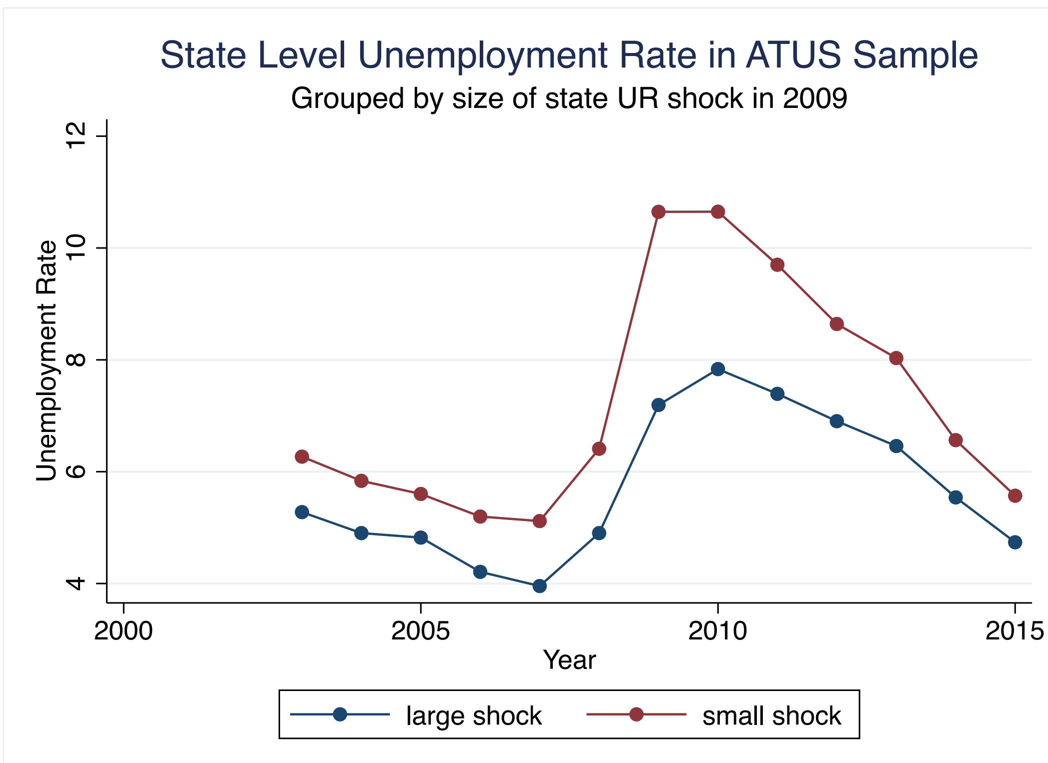


Figure 4.4: State Unemployment Rate for Individuals in ATUS Data

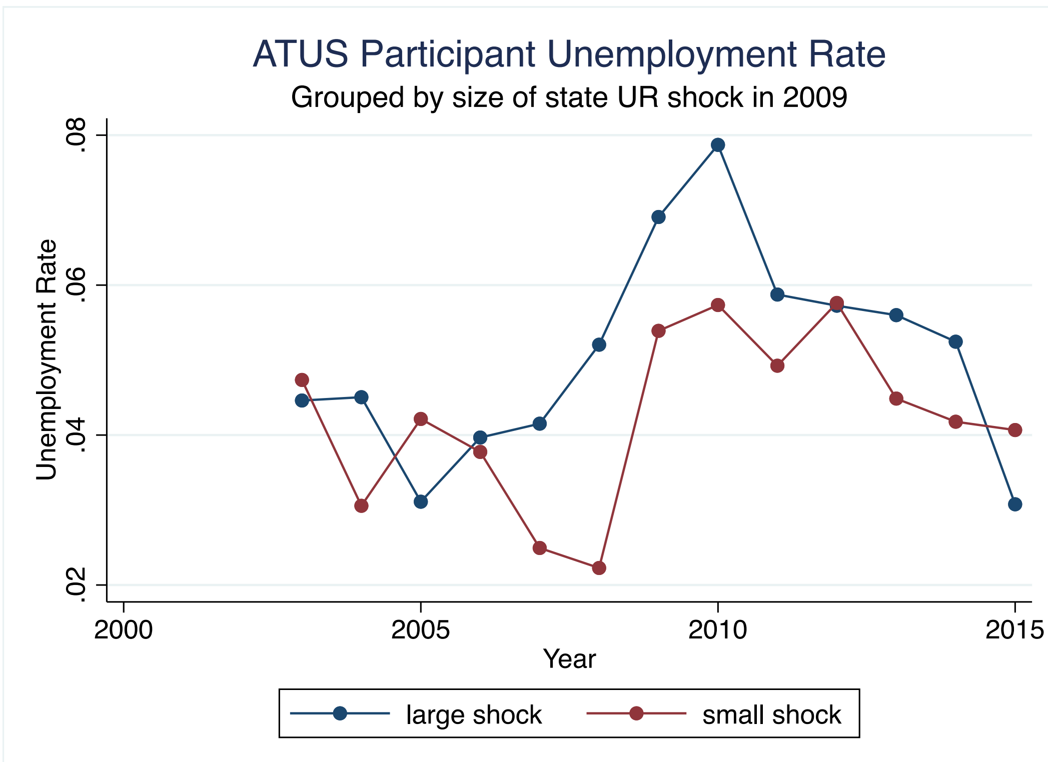


Figure 4.5: Rate of Unemployed ATUS participants

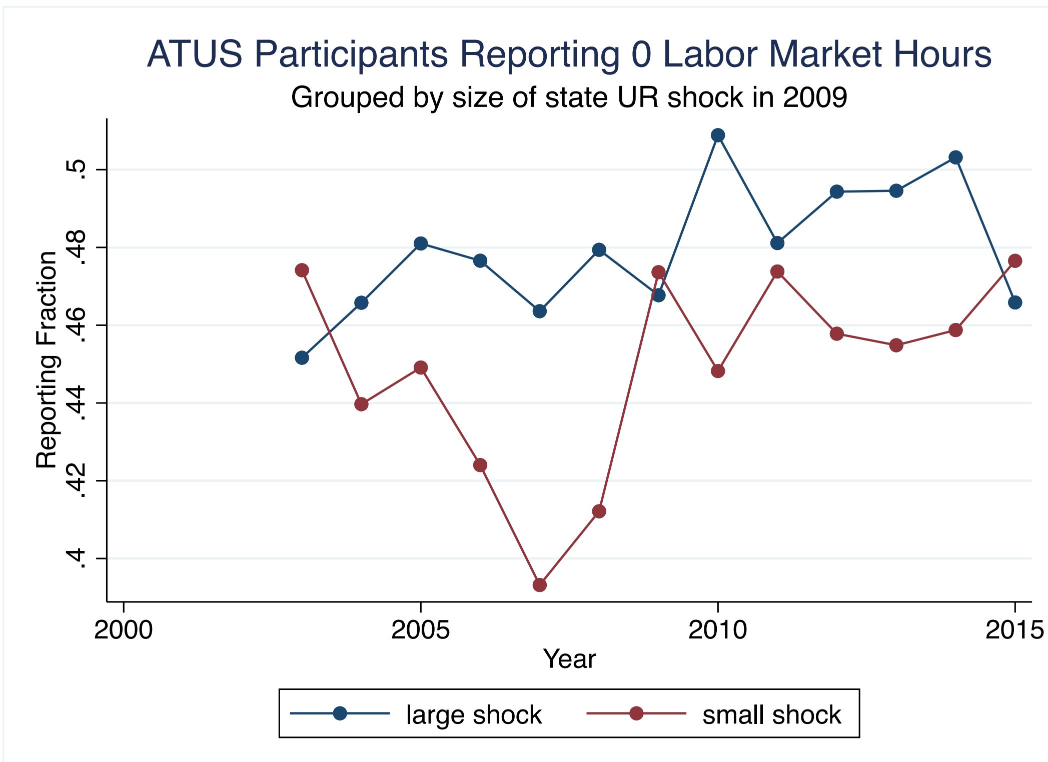


Figure 4.6: Rate of Response Stating Labor Market Hours = 0 in ATUS Data

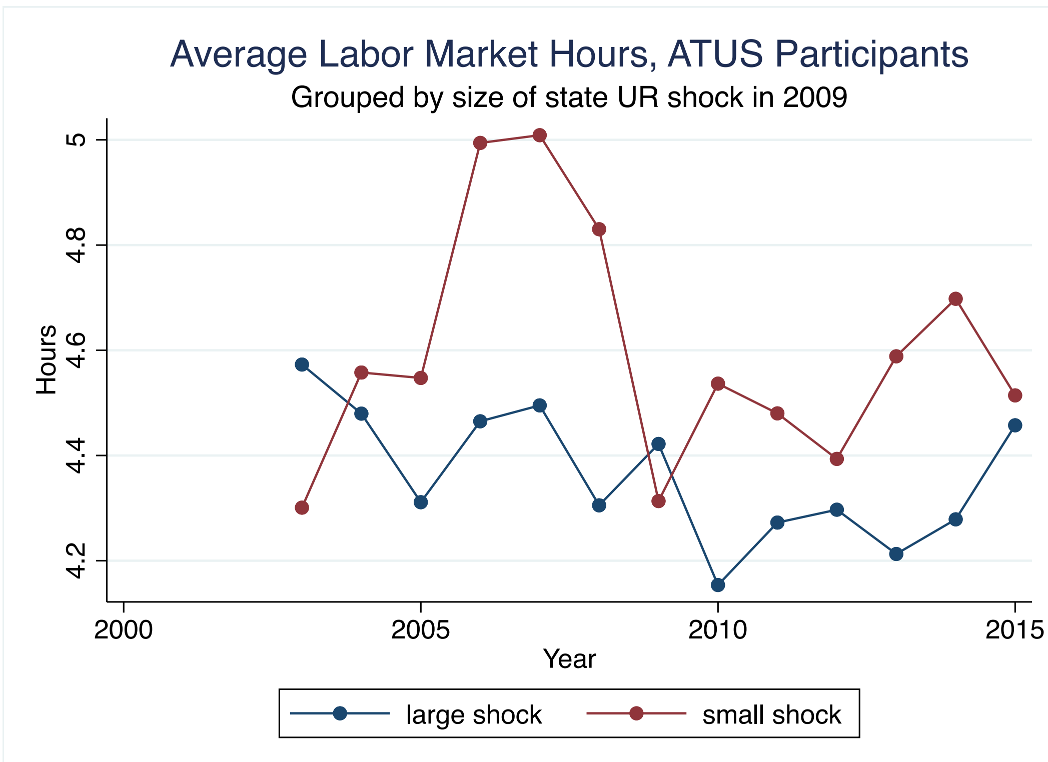


Figure 4.7: Average Labor Market Hours Reported by ATUS Participants

Chapter 5

Conclusion

This dissertation develops three in-depth studies of labor markets in the macroeconomy. Each study addresses a unique question relevant to policy-makers and academics regarding the precise impact of macroeconomic shocks on individual lives.

Throughout my work, I take the stance that data analysis and measurement require close attention to understand the depth and richness of economic phenomena. Further, my research reflects my deep desire to contribute to the wellbeing of individuals in our society. I focus on three key issues in today's society: technological change, recession measurement, and home production. Under today's rapidly evolving economy, each of my studies reflects the heightened need for work that can be employed to improve policy-making and better shape labor markets to serve our entire community.

By examining how technological change impacts individuals and their progeny, I contribute to the current conversation on automation and the changing nature of work. While my work does not directly address modern policy-making on this question, I develop insight into the difficulty individuals face as labor markets shift under their feet. My work sheds appropriate light on the challenges to come, as well as raises hope for future generations that will adapt better to ongoing change.

In my coauthored work on real-time aggregate employment data, my coauthors and I contribute to immediate concerns over data measurement that feed directly into policy decisions. This work has been timely as the COVID-19 pandemic drove employment to a crashing halt and uprooted the lives of many Americans (and others around the globe). By providing tools to better measure recession severity, this work contributes directly to improving the daily lives of individuals.

Finally, my work on the intra-household division of labor has become more relevant in the wake of the COVID-19 recession, as families' daily lives have been turned on their heads with no clear view of what the future will hold. How individuals within a home share the work of home-maintenance and childcare has profound implications for human wellbeing as well as economic recovery.

Overall, my dissertation contributes to the macro-labor literature and macroeconomic policy regarding the changing nature of work. It is my sincerest hope that this work not only advances knowledge but also provides a basis for improving our social fabric.

Bibliography

[1] Yasser Abdih and Stephan Danninger. *What Explains the Decline of the US Labor Share of Income? An Analysis of State and Industry Level Data*. International Monetary Fund, 2017.

[2] Ran Abramitzky, Roy Mill, and Santiago Pérez. “Linking individuals across historical sources: a fully automated approach”. In: *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 53.2 (2020), pp. 94–111.

[3] Daron Acemoglu and Pascual Restrepo. *Automation and New Tasks: The Implications of the Task Content of Technology for Labor Demand*. Tech. rep. mimeo, 2018.

[4] Daron Acemoglu and Pascual Restrepo. “Robots and jobs: Evidence from US labor markets”. In: *Journal of Political Economy* 128.6 (2020), pp. 2188–2244.

[5] Daron Acemoglu and Pascual Restrepo. “The race between man and machine: Implications of technology for growth, factor shares, and employment”. In: *American Economic Review* 108.6 (2018), pp. 1488–1542.

[6] Philippe Aghion et al. “What Are the Labor and Product Market Effects of Automation? New Evidence from France”. In: *CEPR Discussion Paper No. DP14443* (2020).

[7] Mark Aguiar, Erik Hurst, and Loukas Karabarbounis. “Time use during the great recession”. In: *American Economic Review* 103.5 (2013), pp. 1664–96.

[8] Marcella Alsan and Marianne Wanamaker. “Tuskegee and the Health of Black Men”. In: *The Quarterly Journal of Economics* 133.1 (2018), pp. 407–455.

[9] Francisco Alvarez-Cuadrado and Markus Poschke. “Structural change out of agriculture: Labor push versus labor pull”. In: *American Economic Journal: Macroeconomics* 3.3 (2011), pp. 127–58.

[10] A Aneja and Carlos F Avenancio-Leon. “The effect of political power on labor market inequality: Evidence from the 1965 Voting Rights Act”. In: *Unpublished working paper* (2019).

[11] Enghin Atalay. “How important are sectoral shocks?” In: *American Economic Journal: Macroeconomics* 9.4 (2017), pp. 254–80.

[12] David H. Autor, David Dorn, and Gordon H. Hanson. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States". In: *American Economic Review* 103.6 (2013), pp. 2121–68. DOI: 10.1257/aer.103.6.2121. URL: http://www.aeaweb.org/articles?id=10.1257/aer.103.6.2121.

[13] Martha Bailey et al. "How well do automated methods perform in historical samples? Evidence from new ground truth". In: *National Bureau of Economic Research Working Papers* (2017).

[14] Samuel Bazzi, Martin Fiszbein, and Mesay Gebresilasse. "Frontier culture: The roots and persistence of "rugged individualism" in the United States". In: *Econometrica* 88.6 (2020), pp. 2329–2368.

[15] Jess Benhabib, Richard Rogerson, and Randall Wright. "Homework in macroeconomics: Household production and aggregate fluctuations". In: *Journal of Political Economy* 99.6 (1991), pp. 1166–1187.

[16] Olivier Jean Blanchard and Lawrence F Katz. "Regional Evolutions". In: *Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution* 23 (1992), p. 1.

[17] Sally Clarke. "New Deal Regulation and the Revolution in American Farm Productivity. A Case Study of the Diffusion of the Tractor in the Corn Belt, 1920–1940". In: *The Journal of Economic History* 51.1 (1991), pp. 101–123.

[18] Estela B Dagum. "Seasonal factor forecasts from ARIMA models". In: *Bulletin of the International Statistical Institute* 46.3 (1975), pp. 203–216.

[19] Esther Duflo, Michael Kremer, and Jonathan Robinson. "Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya". In: *American Economic Review* 101.6 (2011), pp. 2350–90.

[20] Fabian Eckert and Michael Peters. "Spatial Structural Change and Agricultural Productivity". In: (2017).

[21] Michael WL Elsby, Bart Hobijn, and Ayşegül Şahin. "The decline of the US labor share". In: *Brookings Papers on Economic Activity* 2013.2 (2013), pp. 1–63.

[22] James Feigenbaum and Daniel P Gross. "Automation and the Fate of Young Workers: Evidence from Telephone Operation in the Early 20th Century". In: *National Bureau of Economic Research Working Papers* (2020).

[23] Price Fishback. "How Successful Was the New Deal? The Microeconomic Impact of New Deal Spending and Lending Policies in the 1930s". In: *Journal of Economic Literature* 55.4 (2017), pp. 1435–85.

[24] Eric Ghysels, Casidhe Horan, and Emanuel Moench. "Forecasting through the rearview mirror: Data revisions and bond return predictability". In: *The Review of Financial Studies* 31.2 (2018), pp. 678–714.

[25] Georg Graetz and Guy Michaels. "Robots at work". In: *Review of Economics and Statistics* 100.5 (2018), pp. 753–768.

[26] Jeremy Greenwood, Nezih Guner, and Guillaume Vandenbroucke. "Family economics writ large". In: *Journal of Economic Literature* 55.4 (2017), pp. 1346–1434.

[27] Daniel P Gross. "Scale versus scope in the diffusion of new technology: evidence from the farm tractor". In: *The RAND Journal of Economics* 49.2 (2018), pp. 427–452.

[28] G.B. Gunlogson. "General Purpose Tractor Needed for American Farm Market". In: *Automotive Industries* 47.1 (1922), pp. 4–7.

[29] Richard Hornbeck and Suresh Naidu. "When the levee breaks: black migration and economic development in the American South". In: *American Economic Review* 104.3 (2014), pp. 963–90.

[30] Loukas Karabarbounis and Brent Neiman. "The global decline of the labor share". In: *The Quarterly Journal of Economics* 129.1 (2014), pp. 61–103.

[31] Peter B Kenny and James Durbin. "Local trend estimation and seasonal adjustment of economic and social time series". In: *Journal of the Royal Statistical Society: Series A (General)* 145.1 (1982), pp. 1–28.

[32] Michael Kremer, Gautam Rao, and Frank Schilbach. "Behavioral development economics". In: *Handbook of Behavioral Economics: Applications and Foundations 1*. Vol. 2. Elsevier, 2019, pp. 345–458.

[33] N Gregory Mankiw and Matthew D Shapiro. "News or noise? An analysis of GNP revisions". In: *National Bureau of Economic Research Working Papers* (1986).

[34] Charles F Manski. "Communicating uncertainty in official economic statistics: An appraisal fifty years after Morgenstern". In: *Journal of Economic Literature* 53.3 (2015), pp. 631–53.

[35] Rodolfo E Manuelli and Ananth Seshadri. "Frictionless technology diffusion: The case of tractors". In: *American Economic Review* 104.4 (2014), pp. 1368–91.

[36] Dinah Duffy Martini and Eugene Silberberg. "The diffusion of tractor technology". In: *The Journal of Economic History* 66.2 (2006), pp. 354–389.

[37] Sandra K McKenzie. "Concurrent seasonal adjustment with census X-11". In: *Journal of Business & Economic Statistics* 2.3 (1984), pp. 235–249.

[38] Benjamin Moll, Lukasz Rachel, and Pascual Restrepo. "Uneven Growth: Automation's Impact on Income and Wealth Inequality". In: *National Bureau of Economic Research Working Paper* (2021).

[39] Oskar Morgenstern et al. *On the accuracy of economic observations*. Princeton University Press, 1963.

[40] Alan L Olmstead and Paul W Rhode. "Reshaping the landscape: the impact and diffusion of the tractor in American agriculture, 1910–1960". In: *The Journal of Economic History* 61.3 (2001), pp. 663–698.

[41] Alan L Olmstead and Paul W Rhode. “The Diffusion of the Tractor in American Agriculture: 1910-60”. In: *National Bureau of Economic Research Working Papers* (2000).

[42] Martha L Olney and Aaron Pacitti. “The rise of services, deindustrialization, and the length of economic recovery”. In: *Economic Inquiry* 55.4 (2017), pp. 1625–1647.

[43] David A Pierce and Sandra K McKenzie. “On concurrent seasonal adjustment”. In: *Journal of the American Statistical Association* 82.399 (1987), pp. 720–732.

[44] Raghuram Rajan and Rodney Ramcharan. “The anatomy of a credit crisis: The boom and bust in farm land prices in the United States in the 1920s”. In: *American Economic Review* 105.4 (2015), pp. 1439–77.

[45] Peter Rupert, Richard Rogerson, and Randall Wright. “Estimating substitution elasticities in household production models”. In: *Economic Theory* 6.1 (1995), pp. 179–193.

[46] Richard H Steckel and William J White. “Engines of growth: farm tractors and twentieth-century US economic welfare”. In: *National Bureau of Economic Research Working Papers* (2012).

[47] Tavneet Suri. “Selection and comparative advantage in technology adoption”. In: *Econometrica* 79.1 (2011), pp. 159–209.

[48] Reynold M Wik. “Henry Ford’s tractors and American agriculture”. In: *Agricultural History* 38.2 (1964), p. 79.

[49] Jonathan H Wright. “Unseasonal seasonals?” In: *Brookings Papers on Economic Activity* 2013.2 (2013), pp. 65–126.

[50] Danny Yagan. *Employment hysteresis from the great recession*. Tech. rep. National Bureau of Economic Research, 2017.

Appendix A

Appendix for Chapter 2

A.1 Linked Population Census Records

To understand the impact of technological adoption on vulnerable individuals, I employ the US Population Censuses for 1900-1940. The US Census was not collected as a panel with individual identifiers between decades. Thus, I must use computational methods to match census records between decades.

This is an exercise that is still evolving in the literature and is being worked on by a number of research groups. I follow the recent census linking literature, but add a metric that allows me to maintain a notion of 'distance' for each link that I form. By preserving a measure of 'distance,' I can test whether the quality of linked data varies with other variables of interest. To understand the linking methodology, consider the case of linking the 1910 census data to the 1930 census data. The algorithm performs as follows:

1. First, I clean the string names in both the 1910 and 1930 census data. The string cleaning that I perform ensures that all letters are the same case (upper or lower) and removes any non-alphabetic characters. This is different from many other matching algorithms people use that rely heavily on cleaning out some of the variation in spelling that may occur across censuses.

2. Once string names are cleaned, I group the data by first letter of first name, first letter of last name, and birthplace (state, or country if outside the US). I do this grouping for both the 1910 and 1930 data so that each group includes records from both census. This step creates approximately 100,000 groups ranging in size from a single census record to tens of thousands of observations. For a given group, g, there are $N_{g,10}$ records from 1910 and $N_{g,30}$ records from 1930.

3. For each group, g, I compute three $N_{g,10} \times N_{g,30}$ matrices comparing each of the 1910 records in the group to the 1930 records in the group. The first matrix gives the absolute distance between reported birthyears between records. The second and third matrices give the Jaro-Winkler String Distances for the first and last names in the

records, respectively. Each of these matrices represents a series of 'distance' metrics between records in the 1910 and 1930 censuses. To give a sense of scale, this step can take over 100G of memory for a single group, g.

4. Combine the three distance measures for each 1910-1930 records pair into a single distance metric for the record pair. At the moment, this step weights each of the three metrics equally. This step produces a single matrix of distances for each group, g.

5. For each individual in 1910, choose the minimum distance match to a 1930 record. Because there may be repeated values of 1930 records in the set of minimized distance matches, repeat this step for individuals in 1930. This results in a dataset of unique matches where each 1910 record is matched to a single 1930 record and vice versa. Not all records are matched.

6. For cases in which there is more than one "exact" match (the names and birth years exactly match between two datasets), it is not possible to take a single minimum match. In this case, I choose one match at random. Econometrically, this preserves information and adds only random, unbiased noise.

Finally, I take the best 25% of matches to reduce the incidence of poor quality data. This procedure results in a set of unique matches that have an associated 'distance' that can be used to test for bias due to my matching algorithm. The relationship between link quality and crop mix or tractor adoption is depicted in Figures A1 and A2. We can wee that the link quality is better (lower distance metric) in places that adopted more tractors or grew more small grains. However, the statistical significance is weak. At worst, this shows that for outcomes such as migration, which are correlated with link quality, my estimates may be biased downward. That is, there may be more migration as a result of tractor adoption than my estimates reflect.

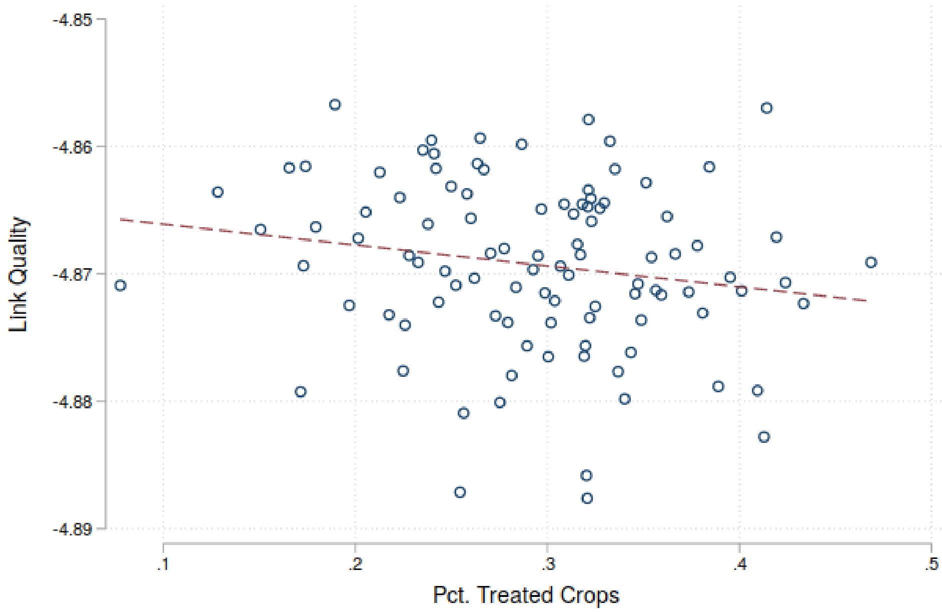


Figure A1: Computed Link Quality (1910-1930) vs. Pct. Treated Crops in 1910

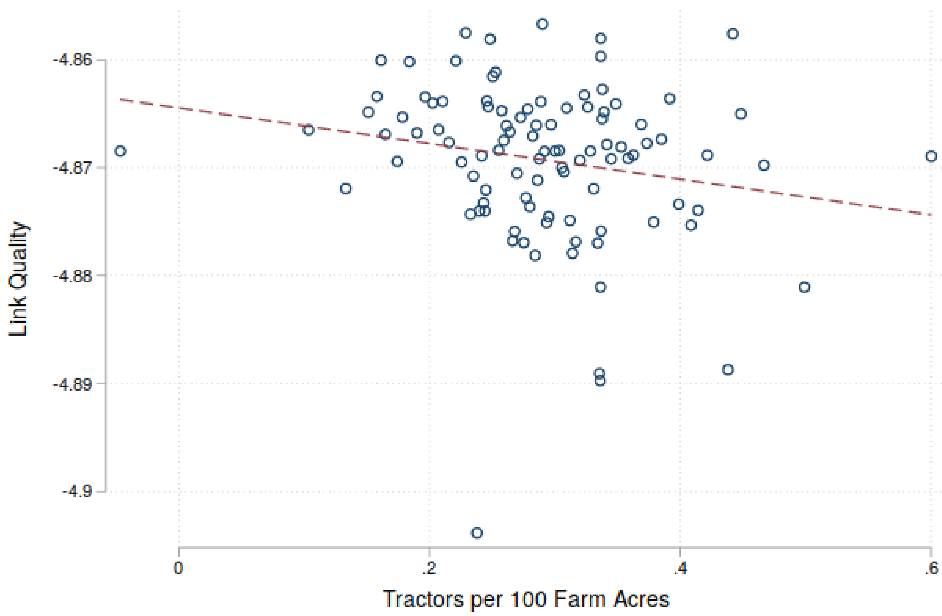


Figure A2: Computed Link Quality (1910-1930) vs. Tractors per Acre 1930

Appendix B

Appendix for Chapter 3

Table B1: Robustness Check. X13-constructed data as dependent variable.

	(1) X13 final	(2) X13 final	(3) X13 final	(4) X13 final	(5) X13 final	(6) X13 final
BLS Replication	1.169 (0.0289)					
3x15 Bandwidth		1.103 (0.0279)				
Non-concurrent			0.959 (0.0245)			
No forecast step				1.045 (0.0248)		
Non-concurrent, No forecast					0.941 (0.0293)	
No forecast, non-concurrent, 3x15 BW						0.979 (0.0297)
Constant	-0.148 (0.165)	-0.438 (0.173)	0.00901 (0.169)	-0.109 (0.159)	0.0365 (0.202)	-0.209 (0.203)
N	188	188	188	188	188	188
Month FEs	Y	Y	Y	Y	Y	Y
RMSE	0.632	0.660	0.652	0.611	0.776	0.778

Note: This table presents the results of regressing the most up-to-date employment growth data released by the BLS on the initially released real-time employment growth data under a variety of different X-13 specifications. The table uses data from May 2003 to December 2018. In each regression, the left hand side variable is the most up-to-date version of the data as computed using our alternative X-13 specifications. Column (1) reports the regression using the re-constructed real time seasonally adjusted data meant to replicate the BLS. Column (2) reports the regression using a version of the real-time data constructed with X-13 in which the bandwidth of the moving average estimator is increased to 3x15. Column (3) uses data produced from running X-13 without concurrent estimation of the seasonal factors. Column (4) uses data produced from running X-13 without the forecasting step of the seasonal adjustment process. Column (5) uses data produced from running X-13 without the forecasting step of the seasonal adjustment process and also without the concurrent estimation. Column (5) uses data produced from running X-13 without the forecasting step, without the concurrent estimation, and with a bandwidth of 3x15. All regressions include month-of-year fixed effects. Standard errors in parentheses.

Table B2: Robustness Check. Regressions without month-of-year fixed effect.

	(1) BLS final	(2) BLS final	(3) BLS final	(4) BLS final	(5) BLS final	(6) BLS final
BLS Replication	1.143 (0.0316)					
3x15 Bandwidth		1.004 (0.0333)				
Non-concurrent			0.925 (0.0276)			
No forecast step				1.018 (0.0278)		
Non-concurrent, No forecast					0.905 (0.0316)	
No forecast, non-concurrent, 3x15 BW						0.844 (0.0347)
Constant	0.0276 (0.0565)	0.131 (0.0648)	0.161 (0.0589)	0.0928 (0.0550)	0.189 (0.0669)	0.238 (0.0758)
N	188	188	188	188	188	188
Month FEs	N	N	N	N	N	N
RMSE	0.700	0.817	0.748	0.691	0.852	0.970

Note: This table presents the results of regressing the most up-to-date employment growth data released by the BLS on the initially released real-time employment growth data under a variety of different X-13 specifications. The table uses data from May 2003 to December 2018. Column (1) reports the regression using the re-constructed real time seasonally adjusted data meant to replicate the BLS. Column (2) reports the regression using a version of the real-time data constructed with X-13 in which the bandwidth of the moving average estimator is increased to 3x15. Column (3) uses data produced from running X-13 without concurrent estimation of the seasonal factors. Column (4) uses data produced from running X-13 without the forecasting step of the seasonal adjustment process. Column (5) uses data produced from running X-13 without the forecasting step of the seasonal adjustment process and also without the concurrent estimation. Column (5) uses data produced from running X-13 without the forecasting step, without the concurrent estimation, and with a bandwidth of 3x15. These regressions do not include month-of-year fixed effects. Standard errors in parentheses.

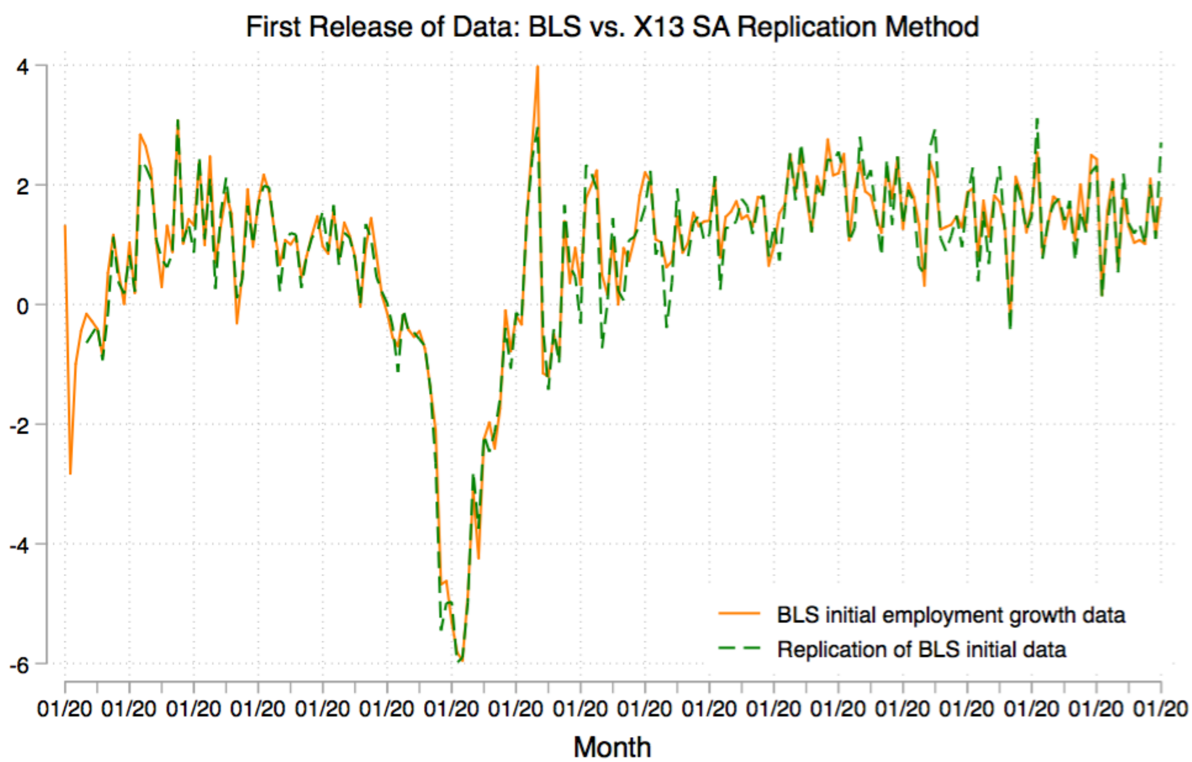


Figure B1: Initial data released by BLS compared to our replication using real-time NSA data and X-13.

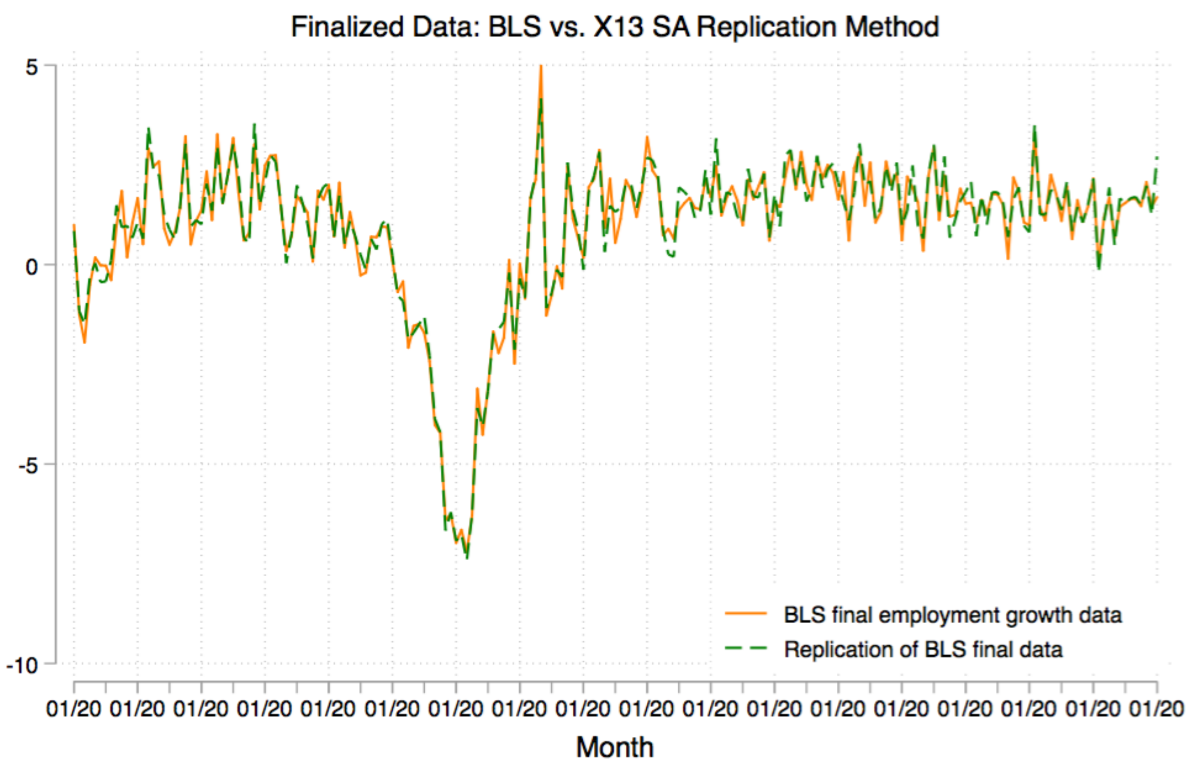


Figure B2: Final data released by BLS compared to our replication using real-time NSA data and X-13.

Appendix C

Appendix for Chapter 4

C.1 Solution to the Model

The model presented in section 2 results in the following first order conditions:

$$\alpha\frac{\partial \ln C_t}{\partial C_{m,t}} = \mu_{1,t} \tag{C.1}$$

$$\alpha\frac{\partial \ln C_t}{\partial C_{h,t}} = \mu_{4,t} \tag{C.2}$$

$$(1-\alpha)\lambda_1(h_w)^{1/\xi_1} = \mu_{1,t}(1-\eta)\frac{Y_{w,t}}{h_{w,t}} \tag{C.3}$$

$$(1-\alpha)\lambda_3(h_m)^{1/\theta_1} = \mu_{1,t}(1-\gamma)\frac{Y_{m,t}}{h_{m,t}} \tag{C.4}$$

$$(1-\alpha)\lambda_2(W_h)^{1/\xi_2} = \mu_{4,t}\nu\frac{C_{h,t}}{W_{h,t}} \tag{C.5}$$

$$(1-\alpha)(M_h)^{1/\theta_2} = \mu_{4,t}(1-\nu)\frac{C_{h,t}}{M_{h,t}} \tag{C.6}$$

$$\beta\mu_{1,t+1}\gamma(Y_{m,t+1}/K_{m,t}) - \mu_{2,t} + \beta\mu_{2,t+1}(1-\delta) = 0 \tag{C.7}$$

$$\beta\mu_{1,t+1}\eta(Y_{w,t+1}/K_{w,t}) - \mu_{3,t} + \beta\mu_{3,t+1}(1-\delta) = 0 \tag{C.8}$$

$$\mu_{1,t} = \mu_{2,t} \tag{C.9}$$

$$\mu_{1,t} = \mu_{3,t} \tag{C.10}$$

Where we know that $\frac{\partial \ln C_t}{\partial C_{m,t}} = \frac{\partial \ln(aC^e_{m,t}+(1-a)C^e_{h,t})^{1/e}}{\partial C_{m,t}} = (1/C_t)aC^{e-1}_{m,t}$ and likewise $\frac{\partial \ln C_t}{\partial C_{h,t}} = (1/C_t)(1-a)C^{e-1}_{h,t}$.

The model simulation uses log-linearization around the steady state. To begin with, the steady state is given by:

$$\bar{Y}_m + \bar{Y}_w = \bar{C}_m + \bar{I}_m + \bar{I}_w$$
$$\bar{Z}_m = \bar{Z}_w = \bar{Z}_h = 1$$
$$\bar{Y}_m = \bar{K}_m^{\gamma} \bar{h}_m^{1-\gamma}$$
$$\bar{Y}_w = \bar{K}_w^{\eta} \bar{h}_w^{1-\eta}$$
$$\bar{C}_h = \bar{W}_h^{\nu} \bar{M}_h^{1-\nu}$$
$$\delta \bar{K}_m = \bar{I}_m$$
$$\delta \bar{K}_w = \bar{I}_w$$
$$\beta\gamma(\bar{Y}_m/\bar{K}_m) - 1 + \beta(1-\delta) = 0$$
$$\beta\eta(\bar{Y}_w/\bar{K}_w) - 1 + \beta(1-\delta) = 0$$
$$(1-\alpha)\lambda_1(\bar{h}_w)^{1/\xi_1} = (\alpha/\bar{C})a\bar{C}_m^{e-1}(1-\eta)\frac{\bar{Y}_w}{\bar{h}_w}$$
$$(1-\alpha)\lambda_3(\bar{h}_m)^{1/\theta_1} = (\alpha/\bar{C})a\bar{C}_m^{e-1}(1-\gamma)\frac{\bar{Y}_m}{\bar{h}_m}$$
$$(1-\alpha)\lambda_2(\bar{W}_h)^{1/\xi_2} = (\alpha/\bar{C})(1-a)\bar{C}_h^{e-1}\nu\frac{\bar{C}_h}{\bar{W}_h}$$
$$(1-\alpha)(\bar{M}_h)^{1/\theta_2} = (\alpha/\bar{C})(1-a)\bar{C}_h^{e-1}(1-\nu)\frac{\bar{C}_h}{\bar{M}_h}$$

We log-linearize the system around the steady state:

$$\check{Y}_{m,t} = \check{Z}_{m,t} + \gamma \check{K}_{m,t-1} + (1-\gamma)\check{h}_{m,t}$$
$$\check{Y}_{w,t} = \check{Z}_{w,t} + \eta \check{K}_{w,t-1} + (1-\eta)\check{h}_{w,t}$$
$$\check{C}_{h,t} = \check{Z}_{h,t} + \nu \check{W}_{h,t} + (1-\nu)\check{M}_{h,t}$$
$$(1/\xi_1)\check{h}_{t,w} = -\frac{(a\bar{C}_m^e \check{C}_{m,t} + (1-a)\bar{C}_h^e \check{C}_{h,t})}{(a\bar{C}_m^e + (1-a)\bar{C}_h^e)} + (e-1)\check{C}_{m,t} + \check{Y}_{w,t} - \check{h}_{w,t}$$
$$(1/\theta_1)\check{h}_{t,m} = -\frac{(a\bar{C}_m^e \check{C}_{m,t} + (1-a)\bar{C}_h^e \check{C}_{h,t})}{(a\bar{C}_m^e + (1-a)\bar{C}_h^e)} + (e-1)\check{C}_{m,t} + \check{Y}_{m,t} - \check{h}_{m,t}$$
$$(1/\xi_2)\check{W}_h = -\frac{(a\bar{C}_m^e \check{C}_{m,t} + (1-a)\bar{C}_h^e \check{C}_{h,t})}{(a\bar{C}_m^e + (1-a)\bar{C}_h^e)} + (e-1)\check{C}_{h,t} + \check{C}_{h,t} - \check{W}_{h,t}$$
$$(1/\theta_2)\check{M}_h = -\frac{(a\bar{C}_m^e \check{C}_{m,t} + (1-a)\bar{C}_h^e \check{C}_{h,t})}{(a\bar{C}_m^e + (1-a)\bar{C}_h^e)} + (e-1)\check{C}_{h,t} + \check{C}_{h,t} - \check{M}_{h,t}$$
$$-\frac{(a\bar{C}_m^e \check{C}_{m,t+1} + (1-a)\bar{C}_h^e \check{C}_{h,t+1})}{(a\bar{C}_m^e + (1-a)\bar{C}_h^e)} + (e-1)\check{C}_{m,t+1} + \frac{(\gamma \bar{Y}_m/\bar{K}_m)(\check{Y}_{m,t+1} - \check{K}_{m,t})}{\gamma(\bar{Y}_m/\bar{K}_m) + (1-\delta)}$$
$$= -\frac{(a\bar{C}_m^e \check{C}_{m,t} + (1-a)\bar{C}_h^e \check{C}_{h,t})}{(a\bar{C}_m^e + (1-a)\bar{C}_h^e)} + (e-1)\check{C}_{m,t}$$
$$-\frac{(a\bar{C}_m^e \check{C}_{m,t+1} + (1-a)\bar{C}_h^e \check{C}_{h,t+1})}{(a\bar{C}_m^e + (1-a)\bar{C}_h^e)} + (e-1)\check{C}_{m,t+1} + \frac{(\eta \bar{Y}_w/\bar{K}_w)(\check{Y}_{w,t+1} - \check{K}_{w,t})}{\eta(\bar{Y}_w/\bar{K}_w) + (1-\delta)} =$$
$$= -\frac{(a\bar{C}_m^e \check{C}_{m,t} + (1-a)\bar{C}_h^e \check{C}_{h,t})}{(a\bar{C}_m^e + (1-a)\bar{C}_h^e)} + (e-1)\check{C}_{m,t}$$
$$\check{Z}_{m,t} = \rho \check{Z}_{m,t-1} + \epsilon_{m,t}$$
$$\check{Z}_{w,t} = \rho \check{Z}_{w,t-1} + \epsilon_{w,t}$$
$$\check{Z}_{h,t} = \rho \check{Z}_{h,t-1} + \epsilon_{h,t}$$
$$\check{K}_{m,t} = (1-\delta)\check{K}_{m,t-1} + \delta \check{I}_{m,t}$$
$$\check{K}_{w,t} = (1-\delta)\check{K}_{w,t-1} + \delta \check{I}_{w,t}$$
$$\check{Y}_{w,t} = \frac{\bar{C}_m}{\bar{Y}_w}\check{C}_{m,t} + \frac{\bar{I}_m}{\bar{Y}_w}\check{I}_{m,t} + \frac{\bar{I}_w}{\bar{Y}_w}\check{I}_{w,t} - \frac{\bar{Y}_m}{\bar{Y}_w}\check{Y}_{m,t}$$