ABSTRACT

Title of Dissertation: ESSAYS ON THE MACROECONOMIC AND MEASUREMENT CONSEQUENCES OF GOVERNMENT SYSTEMS

Michael A. Navarrete
Doctor of Philosophy, 2024

Dissertation Directed by: Professor Judith Hellerstein
Department of Economics

In Chapter 2, I study the macroeconomic consequences to delaying a fiscal stabilizer. Specifically, I study how delays to unemployment insurance benefits during the pandemic recession (fiscal stabilizer) affected consumption (macroeconomic consequence). The United States experienced an unprecedented increase in unemployment insurance (UI) claims starting in March 2020. State UI-benefit systems were inadequately prepared to process these claims. In states that used an antiquated programming language, COBOL, to process claims, potential claimants experienced a larger increase in administrative difficulties, which led to longer delays in benefit disbursement. Using daily debit and credit card consumption data from Affinity Solutions, I employ a two-way fixed-effects estimator to measure the causal impact of having an antiquated UI benefit system on aggregate consumption. Such systems led to a 2.8-percentage-point decline in total credit and debit card consumption relative to card consumption in states with more modern systems. I estimate that the share of claims whose processing was delayed by over 70 days rose

by at least 2.1 percentage points more in COBOL states relative to non-COBOL states. Based on a back-of-the-envelope calculation using 2019 data, my results suggest that the decline in consumption in COBOL states in 2020 after the pandemic-emergency declaration corresponds to a real-GDP decline of at least $105 billion (in 2019 dollars).

In Chapter 3, Joonkyu Choi, Samuel Messer, Veronika Penciakova, and I study how business formation patterns in 2020 were affected by antiquated UI benefit systems. New business formation surged after the pandemic recession, but the causes of this surge are not well understood. The expansion of UI benefits under the CARES Act, coupled with the reduction of work search, provided unemployed potential entrepreneurs with the funds and time needed to develop business ideas. States that used an antiquated programming language, COBOL, to process claims experienced a lower growth rate in UI payments per unemployed than states with more modernized systems. Using business application data from the Business Formation Statistics, we employ a two-way fixed-effects estimator to measure the causal impact of having an antiquated UI benefit system on business formation. Such systems led to a 6.6 percent decline in business applications per capita in COBOL states relative to more modernized states from March 2020 to July 2020. We also find some evidence of business quality deterioration while the Federal Pandemic Unemployment Compensation program was in effect. Our findings highlight the potential role of UI policy in contributing to economic recoveries by fostering entrepreneurship.

In Chapter 4, the RESET team Gabriel Ehrlich, John Haltiwanger, David Johnson, Ron Jarmin, Seula Kim, Jake Kramer, Edward Olivares, R. Rodriguez, Mathew D. Shapiro, and I use point of sales (POS) data to construct real sales and compare these POS generated statistics to official statistics. Businesses, individuals, and government policymakers rely on accurate and timely measurement of nominal sales, inflation, and real output, but current official statistics face challenges on a number of dimensions. First, these key indicators are derived from surveys conducted by multiple agencies with different time frames, yielding a complex integration process. Second, some of the source data needed for the statistics (e.g., expenditure weights) are only available with a considerable lag. Third, response rates are declining, especially for high-frequency surveys. Focusing on retail trade statistics, we document important discrepancies between official statistics and measures computed directly from item-level transactions data. The long lags in key components of the source data delay recognition of economic turning points and lead to out-of-date information on the composition of output. We provide external data sources to validate the transactions data when their nominal sales trends differ importantly from official statistics. We then conduct counterfactual exercises that replicate the methodology that official statistical agencies use with the transactions data in the construction of nominal sales indices. These counterfactual exercises produce similar results to the official statistics even when the official nominal sales and item-level transactions data exhibit different trends.

ESSAYS ON THE MACROECONOMIC AND MEASUREMENT CONSEQUENCES OF GOVERNMENT SYSTEMS

by

Michael A. Navarrete

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:
Professor Judith Hellerstein, Chair/Advisor
Professor John Haltiwanger
Professor Ethan Kaplan
Professor Katharine Abraham
Professor Michael Faulkender

Dedication

To my parents, Hernaldo and Maria, who taught me perseverance.

Acknowledgments

I am deeply grateful to my advisor, Judy Hellerstein, for her mentorship and guidance. Her critical analysis, precision, and clarity helped shape me into the researcher that I am today. As my mentor, Judy was generous with her time and willing to meet with me to discuss my research projects. I have learned to delve into projects to get a deeper understanding and write more compelling research projects. Judy was considerate and always deeply cared for my wellbeing, which I truly appreciate. She alerted me to opportunities that I would not have otherwise known about such as becoming the inaugural Alice Rivlin Dissertation Fellow at the Hutchins Center of Fiscal and Monetary Policy. This broadened my connections and gave me the opportunity to meet a new set of economists that I would not have otherwise met. I still remember the lunch when Ben Bernanke brought his Nobel prize and David Wessel paraded it around for us all to see. Judy also taught me skills not related to research, but which are equally important: how to have a thick skin and how to present research projects.

I am also grateful to John Haltiwanger and Ethan Kaplan, who provided feedback on my research ideas. As a mentor and a coauthor, John helped me always keep thinking about the bigger picture in research questions, while simultaneously going into the weeds of our projects to convince the reader of our bigger picture claims. I am grateful to have worked with John Haltiwanger through my position at the U.S. Census Bureau and hope to continue working on future research projects with him through the RESET team. Through my experience working on these projects, I have become an expert in the Nielsen retail scanner dataset as well as comfortable with working with big data.

I have received feedback and guidance from other economists such as Katharine Abraham, Tara Watson, and Louise Sheiner who have been critical of my work and helped me always be my own biggest skeptic. I thank Michael Faulkender for contributing his time and expertise to my dissertation committee as well as for providing me with firsthand accounts of what was transpiring with policymakers at the beginning of the pandemic. I am thankful to John Shea and Daniel Reck for taking the time to read my job market paper and provide feedback. I am thankful for seminar participants at the University of Maryland such as Melissa Kearney as well as seminar participants at external conferences for providing feedback that has led me to improve my projects and to start new projects.

I would like to acknowledge support from the Peterson Foundation for funding my dissertation fellowship at the Brookings Institution, the Washington Center for Equitable Growth for awarding Seula and me a doctoral grant for our research project, and to the AEA Mentoring Program that has provided me with financial support to attend conferences as well as with a wonderful mentor, Luisa Blanco.

I am extremely grateful to my friends and classmates, including Mario Leccese, Ece Yegane, Daniel Chapman, Sueyoul Kim, Alvaro Silva, Rachel Nesbit, Seula Kim, Kanat Isakov, and Alessandra Palazzo. I thank my coauthors who helped me during the job market through mock interviews and support, Seula Kim, Andrew McCallum, and Veronika Penciakova. I am also thankful to my friends outside the Economics department at the University of Maryland who have supported me through the years, Guan Wang, Roger Vargas, Nick Nosce, Agata Farina, Chelsea Hunter, Karen Huan, and all the RAs in the Economic Studies program at the Brookings

Institution. I would also like to thank my late friend Jim Pagels.

I would like to thank all the professors at Williams College who supported me and nurtured my intellectual curiosity, Sara LaLumia, Dukes Love, and Lucie Schmidt. Susie Godlonton was a great mentor during my time as her research assistant, which was my first experience into economics research. Will Olney was a great undergraduate thesis advisor. Prior to attending Williams College, I knew I wanted to study economics, but I never considered pursuing a Ph.D. in economics. They helped me see the potentially large returns to pursuing a Ph.D. in economics and the diversity of opportunities outside of academia. I tremendously benefited from the liberal arts approach at Williams where professors were extremely generous with their time, both inside and outside the classroom.

Finally, I would like to thank my family for their unwavering support and belief in me. My parents, Maria and Hernaldo, shaped me into the person that I am today. They taught and encouraged me to take new opportunities, especially if it is an uncomfortable one that comes with uncertainty and possibility. They have been loving and supportive of me as has my extended family in Nicaragua and Costa Rica. They are all proud of my accomplishments and they are constantly reminding me of how far I have already come. I am thankful to the individuals that I mentioned in these acknowledgements as well as the countless others who supported me.

Table of Contents

List of Tables

List of Figures

Disclaimer

The fourth chapter of this dissertation uses confidential data from NielsenIQ, Circana, and the U.S. Census Bureau, as such, the following disclaimer applies. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. Only University of Maryland and University of Michigan affiliates have worked with the NielsenIQ data. This paper also uses data from Circana and NielsenIQ housed at the U.S. Census Bureau. The Census Bureau has reviewed the data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data (Project No. 7504725, Disclosure Review Board (DRB) approval numbers: CBDRB-FY19-122, CBDRB-FY21-074, CBDRB-FY23-067, CBDRB-FY23-0234, CBDRB-FY24-0203). There is no commingling of micro data from the different transactions-level data sources. Opinions and conclusions expressed are those of the authors and do not necessarily represent the view of the U.S. Census Bureau. We thank staff at the U.S. Census Bureau for help on the analysis with data at Census especially Jose Asturias. We also thank staff of NPD/Circana for their insights and comments on this research.

Chapter 1: Introduction

In this dissertation, I study the intricate relationship between government systems and their tangible impact on individuals and businesses, particularly against the backdrop of the unprecedented challenges posed by the COVID-19 pandemic. In the United States, government systems play a key role in various aspects of Americans' lives, including the distribution of government transfers. These polices and others are determined by government systems that measure how the economy is faring by producing official statistics such as inflation. In this dissertation, I study unemployment insurance (UI) during the pandemic and the macroeconomic consequences on aggregate consumption and business formation. I also study how official statistics of real sales are measured in the U.S. and the potential shortcomings of the current official methodology.

I begin with two chapters, Chapter 2 and Chapter 3, focusing on unemployment insurance during the pandemic recession in 2020. In 2020, large changes were made to UI through the CARES Act passed at the end of March. Among other things, the Act established the Federal Pandemic Unemployment Compensation (FPUC) program, which increased UI benefits by $600 per week through the end of July 2020, an unprecedented increase in UI generosity. Another important change was the introduction of Pandemic Unemployment Assitatance (PUA). Consistent with these large increases, Ganong, Noel & Vavra (2020a) estimate that 76% of unemployed individuals received more than 100% of their pre-unemployment income as UI benefits during this period. Despite policymakers' awareness of potential issues with the uniform $600 weekly increase leading to replacement rates greater than 100%, altering state UI systems would have been challenging and, due to the antiquated nature of these systems, would have substantially delayed benefit disbursement.

In Chapter 2, I explore the macroeconomic consequences of delaying a fiscal stabilizer. Specifically, I look at UI during the pandemic as the fiscal stabilizer and aggregate consumption as the macroeconomic consequence. In the United States each state has its own unemployment insurance benefit system. All states struggled processing an unprecedented number of initial claims starting in March 2020. The agencies that administered these UI benefits were ill-prepared to handle the surge in claims as well as changes to benefits (eligibility and UI generosity). One important feature of UI benefit systems was whether the UI benefit system was using an antiquated programming language, COBOL, to process claims. My primary data source to measure consumption is daily debit and credit card consumption at the state-level from the Affinity Solutions. States using COBOL experienced longer delays in the disbursement of benefits and potentially experienced an increase in discouraged filers. Using a two-way fixed-effects estimator, I estimate that the recovery in total credit and debit card consumption in COBOL states was 2.8 percentage point lower than in non-COBOL states.

In Chapter 3, my coauthors and I continue to study UI during the pandemic, looking at how antiquated UI benefit systems affected business formation in 2020. There was a sharp increase in business formation during the pandemic and business formation remains elevated above prepandemic levels even years after the pandemic. Decker & Haltiwanger (2023) conduct a deep analysis of the business formation patterns observed during this period and its potential implications for growth. One potential reason business formation surged in 2020 was due to UI that led claimants to have additional time and financial resources to start a business. Building on the COBOL variation identified in Chapter 2, we examine the effect of COBOL usage on business formation using the Business Formation Statistics (BFS) as the primary data source. Using a two-way fixed-effects estimator, we estimate that the increase in business formation in COBOL states was 6.6 percentage points lower than in non-COBOL states.

Transitioning from the realm of UI, Chapter 4 shifts focus to the measurement of real sales in the United States. We decompose real sales into price indices and nominal sales indices. First, we construct real sales using point of sales (POS) data. Then, we we compare these POS series to official ones produced through government systems, where latter get used by businesses, individuals, and policymakers such as the Federal Reserve Board when determining monetary policy. Leveraging insights from Ehrlich, Haltiwanger, Olivares, Shapiro & Zhao (2023), we highlight differences between official statistics and the POS data. The official statistics mostly rely on surveys, which can sometimes be released with a considerable lag. For example, the Bureau of Labor Statistics (BLS) produces the Chained Consumer Price Index (CCPI) and the Consumer Price Index (CPI). Even though the CCPI has advantages over the CPI from a theoretical price index perspective, in practice the CPI is more used by policymakers because the CPI is released before the CCPI.

Not only are there issues with the lag of producing official statistics, but there all also issues with multiple government agencies being involved. The three government agencies involved in creating official real sales statistics are the U.S. Census Bureau, the BLS, and the Bureau of Economic Analysis (BEA). One emblematic example of issues arising from this issue are that concordances are required to map Economic Census (EC) categories (U.S. Census Bureau) and PCE categories (BEA). This would not be problematic if there was a one to one mapping, but

BEA must rely on the sometimes more aggregated EC data to update its disaggregated PCE nominal sales series. We find the reliance on survey data and complex integration processes poses significant challenges, particularly in accurately capturing high-frequency nominal sales variations and adjusting for quality changes in consumer goods. The discrepancies and issues that we find between the POS data and the official statistics only get amplified during the pandemic episode when government agencies had to deal with declining response rates to surveys as well as a large shift away from in-person consumption to online sales.

Through a comprehensive examination of these interconnected themes, this dissertation endeavors to provide a nuanced understanding of the intricate interplay between government systems, economic measurement methodologies, and their tangible impact on individuals and businesses. By bridging theoretical insights with empirical evidence, I aim to measure the tangible macroeconomic consequences of these government systems. In regards to unemployment insurance, Chapters 2 and 3 unravel the macroeconomic consequences on aggregate consumption and business formation, shedding light on the profound implications for economic stability and growth. Regarding the measurement systems employed by government agencies in producing real sales, Chapter 4 highlights the weaknesses of the current measurement systems for real sales in Consumer Tech and Aggregated Food. We propose potential consequences of these shortcomings in official statistics, such as the slow response by the U.S. central bank in addressing the surge in inflation that started in 2021.

Chapter 2: COBOLing Together UI Benefits: How Delays in Fiscal Stabilizers Affect Aggregate Consumption

2.1 Introduction

The COVID-19 pandemic caused a severe contraction in US economic activity, and the fiscal policy response was unprecedented. The federal government spent over $5 trillion on subsidies, transfers, grants, and tax cuts. But a lack of administrative capacity hindered the policy response, with implementation issues affecting both the types of programs enacted and the effectiveness of those programs. The massive spike in unemployment insurance (UI) claims at the beginning of the pandemic, combined with the creation of new UI programs, led to long delays in the disbursement of benefits and even outright crashes of UI systems, particularly in states with antiquated UI benefit systems.[1]

In this chapter, I examine how problems administering unemployment insurance during the pandemic reduced the effectiveness of UI as a fiscal stabilizer. Specifically, I compare consumption changes during the pandemic in states that had not modernized their UI systems with those that had. I proxy for a lack of UI modernization with the use of COBOL (Common Business Oriented Language) in a state's UI benefit system. COBOL is an antiquated programming language

[1] For anecdotal evidence on the problems in COBOL states, see news articles about COBOL usage in New Jersey, Wisconsin, and Connecticut.

developed in 1959 that was once used by all state UI programs. As of 2020, COBOL had been abandoned in the UI benefit systems of 22 states through modernization of their UI system.

I find that while aggregate consumption (as measured by credit and debit card purchases) fell precipitously in all states at the start of the pandemic and remained below pre-pandemic levels for several months, it was slower to recover in states with antiquated UI systems. Using a two-way fixed-effects (TWFE) estimator, I find that the relative decline in consumption from March 13, 2020 to December 31, 2020 was 2.8-percentage-points larger in COBOL states than in non-COBOL states. Using this estimate in a back-of-the-envelope calculation, I find that the failure to invest in updating UI-benefit systems in COBOL states caused real GDP to be at least $105 billion (in 2019 dollars) lower during this period.

UI serves as not only a safety net for laid-off workers during recessions but also as a macroeconomic buffer. UI benefits increase income for households with unemployed workers, which in turn increases household consumption. Because the fiscal-multiplier effect of a dollar of UI benefits during a recession is likely greater than 1, UI has positive general equilibrium effects, including effects on household consumption.[2] My estimates, therefore, likely reflect a combination of the direct effect on UI-eligible households in COBOL states and indirect effects in the form of a dampened fiscal multiplier. In other words, the consumption effect that I estimate is driven by claimants in COBOL states experiencing a relatively higher administrative burden, which led to UI functioning as a less effective fiscal stabilizer in COBOL states.

The primary mechanism by which COBOL usage in UI benefits led to lower aggregate consumption is delays in the UI benefits disbursement. COBOL states could have experienced delays

[2] Kekre (2022) find that UI benefit extensions have a contemporaneous output multiplier of around 1 using data from 2008 to 2014. Di Maggio & Kermani (2016) estimate a local fiscal multiplier of unemployment insurance expenditures of around 1.9 using data from 1999 to 2013.

in UI disbursements, both because it took longer for claimants to successfully file a claim and because processing those claims took longer.[3] Delayed claims could affect aggregate consumption through a dampened UI fiscal multiplier. Ganong, Greig, Noel, Sullivan & Vavra (2022a) find that the one-month marginal propensity to consume (MPC) for claimants who received their benefits was highest in April 2020 compared to later periods when UI benefits changed, such as the expiration of the $600 supplement in July 2020. When households save a larger share of their UI benefits due to delayed disbursement, they will purchase fewer goods and services, directly lowering consumption. That household's consumption decisions will have spillovers, leading to a lower UI fiscal multiplier, since the UI fiscal multiplier is directly related to the MPC. Given that UI relative replacement rates were on average over 100% (Ganong, Noel & Vavra, 2020b) from April 2020 to the end of July 2020, delaying UI benefits by months could significantly alter consumption behavior.

2.2 Background

In this section, I describe the changes to UI benefits in 2020, provide more justification for the use of COBOL as a proxy for administrative capacity, and relate this chapter to the broader literature.

2.2.1 Changes to UI Benefits in 2020

To understand the effects of administrative capacity on UI disbursements in 2020, it is helpful to review what changed in UI benefits during the pandemic and the difficulties that potential

[3] Second, there may have been more discouraged filers in COBOL states—potential claimants who did not file a claim (or who did not complete the filing process) because they viewed applying for UI as too complicated or laborious. A survey conducted by the Economic Policy Institute shows that some claimants chose not to file for benefits because it was too difficult.

claimants faced. Before the pandemic, potential claimants filed an initial claim with their state's UI office online, over the phone, or (least commonly) in person. After the emergency declaration on March 13, 2020, states faced an unprecedented increase in claims. This led many UI websites to crash, which further overwhelmed call centers.[4] Before and after the CARES Act, claimants had to demonstrate eligibility by having worked in covered employment during their base period. Prior to the CARES Act, a large share of jobs counted as covered employment, but self-employment, gig work, and contract work did not. However, these workers were made eligible for benefits by the CARES Act through its Pandemic Unemployment Assistance provision.[5] COBOL states disproportionately struggled to process claims, both with the unprecedented increase in initial claims filed and with implementing changes to UI benefit systems such as eligibility rules.

During recessions, the federal government often extends the duration of eligibility for UI benefits and rarely increases the benefit amount.[6] However, the magnitude of the benefit enhancement in the CARES Act was unprecedented—an additional $600 per week from April 2020 until July 31, 2020. The CARES Act also increased the maximum duration of benefits by 13 weeks.

When designing the CARES Act specifics, many policymakers were aware that UI benefit systems were antiquated and slow. They would have liked to increase the UI replacement rate (the share of base period earnings replaced by UI benefits), but making that change to state UI systems would have been difficult and would have greatly delayed the disbursement of the enhanced

[4]Some states opened pop-up UI offices to process claims. Kentucky, a COBOL state, opened multiple pop-up offices.

[5]For a more complete discussion of the typical claim process during the pandemic, refer to Cajner, Figura, Price, Ratner & Weingarden (2020b).

[6]The Federal Additional Compensation (FAC) program, part of the American and Reinvestment Act of 2009, provided an additional $25 per week to UI recipients.

benefits.[7] Instead, policymakers increased benefits by $600 per week for every beneficiary,[8] even though such a change led to replacement rates of over 100% for the majority of UI recipients. In other words, this supplement led to the average UI recipient receiving more from UI benefits than from their previous employer (Ganong et al., 2020b).

Although initial claims spiked just after the emergency declaration and then declined, they remained elevated throughout 2020. Initial claims in every week in 2020 after the emergency declaration surpassed the previous recorded maximum initial claims between 1967 (when the data began to be collected) and the emergency declaration.[9] This persistent state of elevated initial claims meant that states had to process not only a large stock of existing claims from the first weeks after the emergency declaration but also a large flow of initial claims. (Initial claims may represent people who newly become unemployed but also new repeat claims by filers whose claims have not been processed and who may be unsure if they filed correctly.)

Across all states, the CARES Act provided the same additional benefits for claimants, in terms of both benefit amounts and maximum weeks to receive benefits. However, because states had different initial levels of maximum benefits and maximum duration, UI benefits during the pandemic were not uniform across the United States; these increases were different relative to baseline levels across states. For example, prior to the pandemic, Florida (a non-COBOL state) offered a maximum weekly benefit of $275 for a maximum of 12 weeks, while New Jersey (a COBOL state) offered a maximum weekly benefit of $713 for a maximum of 26 weeks.[10] How-

[7] Personal communication with Wendell Primus, senior policy advisor to House Speaker Nancy Pelosi.

[8] After $600-a-week UI supplements expired at the end of July 2020, claimants received $300-a-week UI supplements until summer 2021.

[9] To be precise, the week ending on March 21, 2020 was the first week to surpass the previous recorded peak of 695,000 claims.

[10] For a complete list of how states varied in maximum UI-benefit allocation prior to the pandemic, see this table from a research brief by the Brookings Institution.

ever, variation in which states used COBOL affected both when and whether claimants received UI benefits. For example, Wisconsin, a COBOL state, experienced delays so significant that in June 2020, the Wisconsin Department of Workforce Development was still processing initial claims filed in March.[11]

2.2.2 The Effect of COBOL on Administrative Capacity

COBOL can perform the same tasks as any modern programming language, but systems running on it were differentially overloaded by both the sudden influx of claims and the changes to UI benefits. COBOL states likely struggled more to implement the changes introduced by the CARES Act because of the difficulty of changing COBOL-based systems. Other UI-system failures were also related to COBOL. When the pandemic started, several COBOL states needed COBOL programmers. Because demand for COBOL programmers exceeded supply, some programmers came out of retirement to work in UI offices in COBOL states, but there were insufficient COBOL programmers.[12] Furthermore, COBOL states may have experienced issues not specific to COBOL but symptomatic of an antiquated system, including having a less user-friendly website, the absence of a mobile version of their website, or legacy platforms such as mainframes. These additional issues would have increased the administrative burden that potential claimants would have faced, which would affect delays in disbursement and increase the administrative burden for claimants.

Despite COBOL being capable of completing the same tasks as modern programming languages, implementing these changes were more difficult in COBOL states. One of the issues

[11] See Wisconsin news report.

[12] A group of retired COBOL programmers called the COBOL Cowboys exists solely to aid during crises. See NPR news article.

is that UI benefit systems using COBOL have spaghetti code; this programming code is complex, difficult to read, and highly complex. For example, code written using a fourth generation programming language (most modern programming languages) has half the number of lines of code as the same program written in COBOL (third generation programming language). For example, as of 2020 Wisconsin's (COBOL state) UI benefit system comprises of roughly 8.6 million lines of does, which has been updated and extended numerous times over the previous 50 years.[13] There is usually a lack of automation that requires frequent human intervention resulting in redundant and inefficient processing workflow. For example, Wisconsin experienced a surge of over 250,000 claims in 2020 rejected by the UI benefit system relative to 50,000 to 100,000 claims per year pre-pandemic. These claims that were rejected required human intervention, from the Adjustment and Special Programs (ASP) unit in Wisconsin's case. Wisconsin's ASP unit went from a staff of 16 to 140 in 2020. The staff increased disproportionately because the complexity of these claims increased by the introduction of new federal programs such as Pandemic Emergency Unemployment Compensation (PEUC) program and the PUA program. The increased complexity results in each rejected claim taking longer to process for adjudicators.[14]

All states once used COBOL in their UI-benefit system. Some states have modernized their systems, in part by switching to a more modern programming language, such as C# or Java, as shown in Figure 2.1. The decision to modernize could systematically differ between states. As noted by the Government Accountability Office (GAO) in a report prior to the pandemic, some states had trouble modernizing their UI benefit systems for reasons ranging from a lack

[13]For an in depth discussion on how an antiquated UI benefit system including COBOL usage affected the functioning of Wisconsin Department of Workforce Development please refer to their informational briefing on unemployment modernization that can be streamed here.

[14]For an in depth discussion on how costly the adjudication process (rejected claims) was for Wisconsin in 2020 please refer to their informational briefing on unemployment modernization that can be streamed here.

of funding to difficulties with operating legacy stems in tandem with new systems.[15] Perhaps surprisingly, the states with more generous UI benefits and a more generous social safety net such as California and New Jersey were not more likely to have modernized their UI systems. In fact, the states with less-generous UI benefits during nonrecessionary times, such as Florida and North Carolina, were more likely to have modernized. As I discuss in Section 2.4.2, the fact that COBOL is not randomly distributed across states is a potential source of bias, because the COBOL states are more likely to be Democratic states, and these states were more cautious about COVID-19. In order to account for these differences, I use the Republican vote share in 2016 as a proxy control variable in my empirical analyses.

These longer delays in UI disbursement in COBOL states could have affected aggregate consumption. Figure 2.2 shows the means in relative consumption without controlling for potential confounders. The figure plots the mean population-weighted values of total credit and debit card consumption for COBOL and non-COBOL states. After the emergency declaration, COBOL states recovered more slowly than non-COBOL states. Figure 2.2 provides suggestive evidence that delays in UI benefit disbursement affected aggregate consumption.

2.2.3 Related Literature

This chapter adds to the growing literature measuring the economic impacts of the pandemic recession (Faulkender, Jackman & Miran (2023); Cajner, Crane, Decker, Grigsby, Hamins-Puertolas, Hurst, Kurz & Yildirmaz (2020a); Chetty, Friedman, Hendren, Stepner & Team (2020); Coibion, Gorodnichenko & Weber (2020); Marinescu, Skandalis & Zhao (2021); Ganong et al. (2020b)). I exploit a novel source of heterogeneity across states, COBOL usage for UI benefit

[15] See GAO report from 2013.

systems, to measure the relative decline in aggregate consumption caused by the increased administrative burden faced by potential UI claimants in COBOL states. My work also contributes to the literature on fiscal stabilizers (Eilbott (1966); Dolls, Fuest & Peichl (2012); McKay & Reis (2016)) by being the first to directly look at the aggregate economic consequences of delaying a fiscal stabilizer: UI.

The paper in the pandemic-recession literature that most closely resembles mine is Ganong et al. (2022a). The authors exploit delays in UI benefit payments using micro data to calculate the consumption response to UI benefits at the individual level, and then they use those estimates to calculate the effect on aggregate consumption. They find that the UI-benefit enhancements of \$600 and \$300 led to a 2.7% and 1.5% increase in aggregate spending, respectively, and that the \$600 UI supplements increased aggregate consumption by \$430 billion (in 2019 dollars) nationally from April 2020 to July 2020. Because the variation they exploit is at the individual level, they cannot directly estimate multiplier effects.

I ask a different question: how did the higher administrative burden in COBOL states affect aggregate consumption in those states? I find that aggregate spending from March to December 2020 was 2.8 percentage points lower in COBOL states relative to non-COBOL states, leading aggregate consumption to be at least \$105 billion lower than it would have been if COBOL states had modernized their UI systems. My estimates are limited to the effects of using an antiquated UI-benefit system but do include multiplier effects (at least to the extent that these effects differentially affected the local economy).

This chapter also contributes to the literature on the effects of administrative burdens on program effectiveness. The delays in COBOL states are a form of administrative burden: people had to spend hours on the phone or online trying to file, or go in person to the UI office during a pandemic, and they experienced long and uncertain wait times for their claims to be processed. These administrative burdens could have discouraged potential claimants from receiving benefits.[16] My work is the first to look at the macroeconomic consequences of administrative burdens in the context of a fiscal stabilizer. Herd & Moynihan (2018) note that administrative burdens can be a form of policymaking known as targeting, whereby states deliberately increase the administrative burden in order to reduce a program's take-up rate. This could be true with respect to UI before the pandemic: states with less-generous UI benefits, which tended to be more Republican-leaning states, also had stricter eligibility rules (Skandalis, Marinescu & Massenkoff, 2022). An alternative explanation for the increase in administrative burden in UI is to reduce fraud. But unlike other administrative burdens where policymakers may be making deliberate choices to weaken a government program (Herd & Moynihan, 2018) or reduce fraud, COBOL is not chosen to increase the administrative burden on claimants. COBOL was not problematic in processing claims prior to the pandemic during nonrecessionary periods, as reflected by COBOL states having a lower share of topcoded claims prior to the pandemic recession in Figure 2.3. Topcoded claims are claims that experience a processing delay greater than 70 days. COBOL usage in UI benefit systems becomes a binding constraint when UI systems are overwhelmed with claims or when large changes are made to UI benefits, as was the case during the pandemic recession.

2.3 Data

I use two main sources of data for my outcomes variables: the Department of Labor Employment and Training Administration (DOLETA) and Affinity Solutions, which is part of the Economic Tracker (Chetty et al., 2020).

[16] Using the DOLETA 5159 report, I only find weak evidence of discouraged filers.

To analyze the impact of UI disbursement delays, I use the 9050 report from DOLETA. The 9050 report contains information on disbursement delays. To measure my main outcome variable—relative aggregate consumption—I use the Economic Tracker, which provides daily consumption data from a set of debit and credit cards by state.

2.3.1 Data on Delays in Processing UI Claims (9050 Report)

DOLETA's 9050 report contains monthly information on how long after receiving a claim each state takes to make the first regular UI benefit payment. These data are reported by states to DOLETA and are used for multiple purposes such as measuring state performance and allocating UI administrative funding. The report only imperfectly captures the difference in delays between COBOL and non-COBOL states. First, it only captures delays in processing time and not in filing. Delays in filing are the time between when a claimant starts filing a claim and when that claim is successfully filed, while delays in processing are the time between when a claimant successfully files a claim and when they receive their first UI payment. It is likely that it took longer for claimants in COBOL states to successfully file a claim because the UI systems were more overwhelmed in those states. Second, the report topcodes delays greater than 70 days. COBOL states had more topcoded claims, making it difficult to get an accurate comparison of delays in COBOL and non-COBOL states. Third, even for non-topcoded delays, the report does not include the number of days of delay but instead assigns delays to buckets of discrete weeks (e.g., delays between 1 and 10 weeks). Finally, the report only covers regular UI, with no information available on claims processing for the Pandemic Unemployment Assistance program (the program through which people ineligible for regular UI, like gig workers, received benefits during the pandemic). However, it seems likely that delays in regular UI would be a reasonable proxy for delays in other programs, and data from Ganong et al. (2022a) suggest that the majority of claims processed were regular UI claims. PUA claims could have caused larger issues for UI benefit systems given that this program added new complexity to claims. However, the PUA program could have also affected regular UI by drawing resources such as staff away from processing regular UI.

Whether a claim is topcoded is a lagging indicator of when a claim was originally filed, because claims are only reported as delayed once benefits have been paid out. For example, if a claimant files for UI benefits in March 2020 and gets benefits starting in June 2020, then the claim will be reported as topcoded in June 2020. However, if that same claimant starts to receive UI benefits in July 2020, then the claim would not be reported as topcoded until July 2020. Given that I cannot observe when claims are initially filed, I use as my measure of delay the number of people whose first benefit was more than 70 days late (i.e., the number topcoded) as a share of all the people receiving their first benefit in a given month. Ideally, I would be able to determine when a topcoded claim was initially filed and then calculate the share of topcoded claims with the month filed instead of the month paid. This would fix the lagging indicator issue in the numerator. As an alternative indicator of delays in UI benefit disbursement, I measure the share of claims that are delayed at least 5 weeks. This measure will suffer from being a lagging indicator similar to the topcoded claims, but the lag will be mechanically shorter for these non-topcoded claims.

The limitations of the data also give rise to nonclassical measurement error. Processing delays may be a relatively poorer measure of total delays in COBOL states than in non-COBOL states. Total delays are the sum of processing delays and delays in filing (the time between when the claimant starts filing a claim and when it is successfully submitted). Given the relatively larger administrative burden in COBOL states, delays in filing could be longer in COBOL states than non-COBOL states. Another concern within processing delay is that the accuracy of measuring processing delays could be related to COBOL usage particularly during recessionary periods. Specifically, non-COBOL states may do a better job of measuring in processing delays given their more modernized UI benefit system. As a result, the estimated effect of COBOL on the timeliness of claim processing should be viewed as a lower bound of the true difference in total delays between COBOL and non-COBOL states.

2.3.2 Data on Consumption

I use Opportunity Insights' Economic Tracker to track consumption patterns at the state level. The main advantage of using these data is that they are available at the daily frequency. The consumer-spending data are credit and debit card spending information provided by Affinity Solutions, which is then transformed and aggregated by Chetty et al. (2020). Seasonally-adjusted daily consumption data—measured relative to consumption in January 2020—are available from January 13, 2020, through the present, although I only use data through the end of 2020.

The data cover about 10% of all debit and credit card consumption in the US (Chetty et al., 2020). Chetty et al. (2020) find that the Affinity Solutions data has broad coverage across industries as shown in their comparisons to Quarterly Services Survey and Advance Monthly Retail Trade Survey, but over-represent categories in which credit and debit card transactions are used. The exclusion of cash consumption would only be problematic if different trends in cash usage emerged between COBOL and non-COBOL states after the emergency declaration.[17] Chetty et al. (2020) compare cash transactions captured in CoinOut grocery data with the Affinity So-

[17] In 2019, the San Francisco Fed found that consumers used cash in about 26% of transactions.

lutions data on total card consumption of groceries and find a signal correlation of 0.9 for the period between January 1, 2020, and June 1, 2020. This high correlation suggests that cash transactions are similar to card transactions. Furthermore, credit and debit card transactions accounted for roughly half of all PCE recorded in national accounts (Chetty et al., 2020). Throughout this chapter, I use the terms "consumption," "total card consumption," and "credit and debit card consumption" interchangeably.

Not having access to Chetty et al.'s (2020) raw individual data, I limit my analysis to changes in consumption relative to January 2020 by state. The data that I use are at the state level, but county-level data are also available. I focus on the state-level analysis because the variation I exploit, COBOL usage, is at the state level. The data are at the daily frequency with a seven-day moving average.

2.3.3 Other Control Variables

In addition to using DOLETA and the Economic Tracker as data sources, I also use data by state from a variety of sources for building a robust set of covariates. The covariates that I control for when estimating the impact of delays on aggregate consumption are new COVID-19 death rates, new COVID-19 case rates, and the interaction of 2016 presidential Republican vote share (vote share for candidate Donald Trump) and the period after the emergency declaration. The case and death rates are provided by the *New York Times*' COVID-19 repository. The COVID-19 data are also available at a daily frequency and are measured using a seven-day moving average. The 2016 election data are cross-sectional and come from the MIT Election Data and Science Lab. Finally, the 2019 population estimates come from the US Census Bureau and are used for weighting purposes. The means of these variables by COBOL usage, along with other covariates that I will discuss later, are shown in Table 2.1.

Additionally, I select five covariates that represent state characteristics from before the emergency declaration and interact each of them with the binary variable $Post_t$. These five covariates have statistically insignificant differences between COBOL and non-COBOL states, but are selected due to potential concerns one may have a priori. All five control variables were selected for concerns for the main outcome of interest: consumption. However, in order to be consistent, I apply these same controls and the Republican vote share interacted with $Post_t$ in the same order across regression tables even when these controls are less relevant (such as for processing delay outcomes). The five state characteristics are (1) income share in accommodation and food services (2019), (2) the percentage of the population living in urban areas (2010), (3) UI generosity (Jan. 2020), (4) the percentage of the population living in poverty (2019), and (5) the percentage of the population with at least a bachelor's degree (2019). Data on these covariates come from (1) BEA, (2) U.S. Census Bureau, (3) Brookings Institution and Department of Labor, (4) Small Area Income and Poverty Estimates program (U.S. Census Bureau), and (5) ACS (U.S. Census Bureau), respectively.

These five confounders do not meaningfully affect my consumption results, but I will briefly list what concerns one may have. Accommodation and food services were disproportionately affected by the pandemic recession, so one may be concerned that COBOL states had a higher income share in accommodation and food services. Another concern may be that COBOL states had less generous UI benefits, so they were slower to recover from the pandemic recession. Both of these concerns are assuaged given that COBOL states had lower income share in accommodation and food services and more generous UI benefits as shown in Table 2.1. In terms of poverty, there was expansion of the safety net during the pandemic recession, so poorer households may have received more transfers from other government transfer programs. COBOL states have a lower poverty rate, but I do not find statistically significant differences in transfers between COBOL and non-COBOL states across the Paycheck Protection Plan (PPP), the Economic Impact Payments (EIP), and the Supplemental Nutrition Assistance Program (SNAP). This analysis is reflected in Tables B.3, B.4, and B.5, respectively. Another concern could be that lower levels of education may lead to lower take up rates for the same administrative burden. I find that COBOL states had higher bachelor's degree completion rates as shown in Table 2.1. Finally, one may be concerned that rural areas may have recovered faster from the pandemic given their lower population density and potentially different response to COVID-19 relative to urban areas.

2.3.4 COBOL Status

I identify COBOL usage for all 50 state UI systems primarily through emails to state officials, news articles, and information from the National Association of State Workforce Agencies (NASWA) Information Technology Support Center (ITSC).[18] This definition of an antiquated UI benefit system closely follows NASWA's definition of an antiquated UI benefit system, but my definition of COBOL usage is more clear given that UI benefit systems are always undergoing a modernization effort.[19] Figure 2.1 is a map of the United States showing which states use COBOL as of 2020. COBOL is the most common language used by state workforce agencies, with 28 states categorized as COBOL and 22 as non-COBOL. There is a scattered distribution

[18]Seven states were identified via direct email, 26 via news articles, and 13 via the ITSC's definition of "modernized" to rule out COBOL states. Secondary sources were used for the remaining states: one state was identified via a Freedom of Information Act request, two states via UI-office reports, and one state through a UI-office job posting requesting COBOL skills.

[19]My results are robust to using NASWA's definition of antiquated. Only two states would change following NASWA's definition: Alabama and Nebraska.

of COBOL states throughout the United States, with no one region accounting for the majority of COBOL states. NASWA provides an overall description of state UI systems, describing them either as antiquated or modernized: typically, a COBOL state is also an antiquated state, and a non-COBOL state is a modernized state.

2.4 Empirical Strategy

I perform two empirical exercises. First, I estimate the relative decline in consumption in COBOL states versus non-COBOL states. Second, I determine whether COBOL states systematically experienced longer delays in UI disbursement.

2.4.1 TWFE Estimator

I use the same empirical strategy to address both questions. The main specification is a two-way fixed-effects (TWFE) estimator in which the treatment group is states that used COBOL in their UI-benefit systems in 2020. The specification is as follows:

$$Y_{it} = \alpha_0 + \beta_1 Post_t * Cobol_i + \beta_2 Post_t * X_i + \gamma Z_{it} + \phi_t + \psi_i + \varepsilon_{it}. \tag{2.1}$$

States are denoted by i and time by t. For analysis using DOLETA (Economic Tracker) data for the outcome variable, t corresponds to month (day). Y_{it} is the outcome variable, which differs for each exercise. For the relative consumption analysis, Y_{it} corresponds to the relative change in consumption in state i on day t. For the processing-delay analysis, Y_{it} corresponds to either the share of topcoded claims or the share of claims delayed at least 5 weeks in state i in month t. $Post_t$ is a binary variable taking the value of 1 for the post-period, which also differs across exercises. In the relative consumption exercise, $Post_t$ takes the value 1 starting on March 13, 2020, while $Post_t$ takes the value 1 starting March 2020 in the delay analysis. $Cobol_i$ is a binary variable taking the value of 1 for states that use COBOL in their UI-benefit system in 2020. X_i are state characteristics from before the emergency declaration, such as the 2016 Republican presidential vote share in state i. Z_{it} is a set of time-varying control variables such as the unemployment rate. ϕ_t is a month or day fixed effect, ψ_i is a state fixed effect, and ε_{it} is the error term.

2.4.2 Relative Consumption

My dependent variable for this exercise is Rel_Cons_{it}, a continuous variable that measures the relative change in relative total card spending for day t in state i. I define the start of the post-period as the date of the emergency declaration, March 13, 2020. Prior to the pandemic, the slower and less-efficient UI systems (in COBOL states) did not cause noticeable delays in benefit disbursement relative to systems in non-COBOL states. This constraint was not binding until the massive spike in UI claims after the emergency declaration, which overwhelmed UI systems. Therefore, my treatment is COBOL states after the emergency declaration.

The Rel_Cons_{it} measure of relative total card consumption incorporates seasonal adjustment and normalization components. The measure is given by

$$Rel\_Cons_{it} = \frac{\frac{C_{i,t2020}}{C_{i,t2019}}}{\frac{C_{i,index2020}}{C_{i,index2019}}} - 1. \tag{2.2}$$

Chetty et al. (2020) publicly share the consumption data expressed in these relative percentage-point changes instead of levels. For more details on the construction of Rel_Cons_{it}, see Chetty et al. (2020). In Equation 2.2, $C_{i,index2020}$ corresponds to the index period in 2020 (the first four complete weeks of January). $C_{i,t2019}$ and $C_{i,index2019}$ represent the same period in 2019. Dividing by the 2019 value seasonally adjusts the data, while dividing by the index period normalizes the data as changes relative to January 2020. In other words, the consumption measure can be thought of as the percentage-point change in seasonally-adjusted consumption relative to seasonally-adjusted mean consumption during the baseline period of January 2020. For example, a value of -0.419 on March 30, 2020, in Wisconsin represents a 41.9-percentage-point seasonally-adjusted decline in average total card consumption in the week ending on March 30, 2020, relative to average total card consumption in Wisconsin in January 2020.

To measure the dynamic effect of COBOL usage on aggregate consumption, I estimate a weekly event study:

$$Rel\_Cons_{ik} = \alpha_0 + \sum_{k=-5}^{41} \beta_k (COBOL_i \times I_k) + \beta_{42} Post_t * X_i + \gamma Z_{ik} + \phi_k + \psi_i + \epsilon_{ik}. \quad (2.3)$$

In Equation 2.3, k denotes the number of weeks since March 13, 2020. $Cobol_i$ is a binary variable that takes the value 1 if state i uses COBOL in its UI benefit system. I_k is binary variable that takes the value 1 if the week is week k. The event study allows me to track the evolution of the impact of higher administrative burdens in UI benefits on aggregate consumption. I define the relative weekly-consumption measure using total card spending from each Friday of the week; I choose Friday because March 13, 2020, fell on a Friday. Similarly, for the state controls, Z_{ik}, I use the Friday value of new COVID-19 death rates and new COVID-19 case rates. X_i denotes state characteristics from before the emergency declaration such as the 2016 Republican vote share. These state characteristics are then interacted with $Post_t$, which takes the value of one for weeks after the emergency declaration.[20] This step ensures that each week's value is not mechanically related to the previous week's value through the seven-day moving average.

To interpret the relative consumption effect from Equations 2.1 and 2.3 as causal, I need to satisfy the conditional parallel trends assumption. A potential violation of the (conditional) parallel trends assumption would be if COBOL states responded differently to COVID-19 than non-COBOL states. Despite experiencing similar COVID-19 case numbers and deaths as reflected in Table 2.1, individuals in COBOL states were more COVID-19 cautious than individuals in non-COBOL states. This difference in attitude would upward-bias my consumption estimates. This would violate the (conditional) parallel trends assumption because COBOL states would have had lower consumption after the emergency declaration relative to non-COBOL states for reasons unrelated to UI benefits. To address this concern, I include the interaction of the 2016 Republican presidential vote share and post-period (after the emergency declaration), which serves as a proxy for COVID-19 cautiousness. This variable also helps control for differential policies between Democratic and Republican states. Specifically, Republican states could have had more relaxed policies in general toward COVID-19 transmission such as fewer stay-at-home orders and fewer school closures.

2.4.3 Delays in UI-Benefit Disbursement

As described above, my primary measure of disbursement delay is the share of processed claims that were paid 70 or more days after filing, i.e., the share of claims that are topcoded. I focus on intrastate claims, which form the majority of claims for all states (Washington, DC, is omitted

[20] In figure B.5, I interact X_i, Republican vote share, with I_k in a similar fashion to COBOL. I find consistent results.

from the sample).[21] This level of analysis circumvents issues with workers living in one state but working in another, which is particularly pronounced in counties on state borders.

I use data starting in January 2019. I drop March, April, and May of 2020 from the topcoded analysis because the share of claims that are topcoded during this period is mechanically decreasing for both COBOL and non-COBOL states. The numerator is not increasing because topcoded claims are a lagging indicator, but the denominator, UI recipients, is contemporaneously increasing. The share of topcoded claims first increases for both COBOL and non-COBOL states in June 2020.

As a second outcome variable highlighting that COBOL states experienced longer delays in the disbursement of benefits, I also compare the share of claims delayed at least 5 weeks between COBOL and non-COBOL states. I only exclude March 2020 from this analysis, but this becomes a less clean exercise since the variation of lags from when a claimant files to when they receive their benefit is by construction larger.

2.5 Main Results

In this section, I document the effect of antiquated UI benefit systems on aggregate consumption. After presenting the main finding on consumption, I provide evidence that these antiquated UI benefits contributed to the increased processing delays in the disbursement of benefits. For my main outcome of interest, consumption, I estimate the effect of these antiquated UI benefit systems by comparing relative changes in consumption between COBOL and non-COBOL states in a static TWFE setting, an event study difference-in-differences setting, and through a back-of-the-envelope exercise. Finally, I compare processing delays between COBOL and non-COBOL

[21] The Economic Tracker does not provide consumption data for Washington, D.C.

states in a static TWFE setting.

2.5.1 Effects of Antiquated UI Systems on Consumption

For the TWFE estimator to be a valid approach for identifying the causal effect of COBOL usage on consumption, the conditional parallel-trends assumption must hold. This means that, absent the surge in claims and changes to UI rules during the pandemic recession, relative consumption trends between states with and without antiquated UI benefit systems would have been the same conditional on the covariates included in my regression equation.

As preliminary evidence of relative consumption differences between COBOL and non-COBOL states after the emergency declaration, I compare their mean daily consumption relative to January 2020 without any controls. Chetty et al. (2020) only provide consumption data starting in January 2020, and I start the analysis period in February 2020 to exclude the index period.[22] Both consumption series hover around zero prior to the emergency declaration on March 13, 2020, which is reflective of the lack of change in consumption patterns. This similarity in consumption patterns between COBOL and non-COBOL states suggests that the assumption of common pre-trends holds from February 1, 2020, to March 12, 2020. The relative consumption patterns of COBOL and non-COBOL states are similar prior to the emergency declaration.

The consumption patterns in COBOL and non-COBOL states begin to diverge after the emergency declaration. As seen in Figure 2.2, both sets of states reach a trough in relative consumption slightly below 30-percentage-points lower than their base period consumption on the same day: March 30, 2020. In addition, consumption did not fully recover by the end of the

[22] I cannot go further back in time due to data limitations. The data start in January 2020, and I exclude January 2020 given that it is the index period.

sample on December 31, 2020. The average relative consumption decline for all states from March 13, 2020 to December 31, 2020, is 7.3-percentage-points (not population weighted). The largest relative declines were immediately after the emergency declaration, but neither COBOL nor non-COBOL states had fully recovered as reflected in Figure 2.2. Figure 2.2 shows that after the emergency declaration, COBOL states consistently had lower relative consumption than non-COBOL states, as reflected in the gap between the two series that appears after March 13, 2020, marked by the red vertical dashed line. Specifically, the gap starts to form in early April, after the spike in initial UI claims that occurred at the end of March 2020. By the end of 2020, non-COBOL states' consumption is almost back to baseline period levels. COBOL states' consumption also recovered, but more slowly.

To formally test whether relative consumption was lower in COBOL states than in non-COBOL states after the emergency declaration, as suggested in Figure 2.2, I run a TWFE estimator. The dependent variable is relative consumption, as defined in Equation 2.2, with the sample period starting on February 1, 2020, and ending on December 31, 2020. The coefficient of interest is the interaction of COBOL and Post, where Post is a binary variable that takes the value 1 after March 13, 2020. In Table 2.2, I present results from Equation 2.1. In column 1, I only add state and day fixed effects when estimating the macroeconomic impact of increased administrative burdens to a fiscal stabilizer in COBOL states relative to non-COBOL states on aggregate consumption. In this naive specification, I find a 4.1-percentage-point larger decline in relative consumption in COBOL states relative to non-COBOL states. However, these consumption differences between COBOL and non-COBOL states could have emerged after the emergency declaration for reasons unrelated to an increased administrative burden in the UI benefit system. In column 2, I address this concern by using the interaction of the 2016 Republican vote share and Post, the coefficient of which is positive and significant. This result shows a positive association between consumption after the emergency declaration and the 2016 Republican vote share. One possible explanation for this result is that more-Republican states had policies that stimulated consumption more, such as shorter and fewer stay-at-home orders. Another possibility is that the more-Republican states were less cautious about COVID-19, with residents being more likely to go outside and consume goods and services that they might not have consumed if they had stayed at home. After controlling for the interaction of Republican and Post, the coefficient on the interaction of COBOL and Post decreases from a 4.1-percentage-point relative decline to a 2.8-percentage-point decline.

Even though column 2 in Table 2.2 is my preferred specification, the concern remains that other omitted variables could be driving the relative drop in consumption between COBOL and non-COBOL states. One potential concern is that COVID-19 affected COBOL and non-COBOL states differently.[23] In column 3, I add the daily new COVID-19 case rates and new death rates as additional controls. The coefficient on the interaction on COBOL and Post holds at a 2.8-percentage-point decline and significant at the 5% level. This result is unsurprising given that COBOL and non-COBOL states experienced similar COVID-19 new case rates and new death rates, as reflected by Table 2.1. There could still be other omitted variables that could be driving the differences in relative consumption between COBOL and non-COBOL states. In column 4, I add the five previously discussed confounders interacted with Post as additional controls. The coefficient on the interaction of COBOL and Post only marginally decreases, dropping to a 2.6-percentage-point decline. In column 5, I control for the monthly unemployment rate and find

[23]Table 2.1 suggests that COBOL and non-COBOL states had similar COVID-19 new case rates and new death rates.

an insignificant positive association between the unemployment rate and relative consumption. However, the monthly unemployment rate is problematic because one channel through which these antiquated UI benefit systems could affect consumption is through the unemployment rate. Specifically, UI is an automatic stabilizer and if COBOL states are relatively slower at disbursing funds, then UI will be a less effective fiscal tool at increasing aggregate demand. The post-treatment bias from using the unemployment rate as a control variable downward biases my coefficient on the interaction of COBOL and Post. This conservative estimate decreases the coefficient to a 2.4-percentage-point decline and remains significant at the 5% level.

2.5.2 Consumption Results by Week

The estimated 2.8-percentage-point decline in total card consumption in COBOL states relative to non-COBOL states represents the average effect in the post-period. There could be larger consumption differences earlier in the sample and potential convergence in consumption patterns between COBOL and non-COBOL states by the end of 2020. In Figure 2.5, I plot the results from Equation 2.3 to show how the effect varies by week. The x-axis denotes weeks relative to the emergency declaration, and the y-axis denotes the percentage-point decline in relative consumption in COBOL states. The dashed red line corresponds to the week starting on March 13, 2020, which marks the beginning of the post-period. Figure 2.5 highlights that the lower relative consumption in COBOL states was persistent given that relative consumption in COBOL states remained lower than relative consumption in non-COBOL states for every week in the post-period. Even though COBOL states were still experiencing a decline in relative consumption relative to non-COBOL states, both groups of states were recovering in the second half of 2020, as reflected in Figure 2.2.

Not only was relative consumption in COBOL states lower than in non-COBOL states every week in the post-period, but the recovery in COBOL states was tepid. Relative consumption in COBOL states fell from week 0 until it reached its trough 11 weeks after the emergency declaration: -4.4 percentage points. One might expect a similar speed of recovery for COBOL states, but COBOL states experienced low relative-consumption growth. At the end of my sample, 41 weeks after the emergency declaration, COBOL states still had 2.4-percentage-point lower consumption relative to non-COBOL states. This protracted recovery suggests UI state agencies struggled not only with the initial inflow of UI claims soon after the emergency declaration but with the continued flow of claims in the subsequent weeks. However, I cannot determine whether consumption fell because states continued to experience delays in processing older claims or because states were suffering dynamic general equilibrium effects through the multiplier effects of previous consumption delays. It could be that the recovery after the trough corresponds to the eventual disbursal of delayed UI claims. The persistence could be, in part, driven by the discouraged filers who never received UI benefits. My results are likely the sum of these effects.

2.5.3 Aggregate Effects of Lack of UI Modernization

To convert the 2.8-percentage-point relative decline in consumption from column 2 of Table 2.2 into a dollar amount, I perform the following back-of-the-envelope calculation:

$$Cost = \frac{PI_{2019}}{PI_{2012}} * Relative\_Decline * 0.8 * \sum_{i=1}^{28} PCE\_Cat_{i,2019}, \tag{2.4}$$

where $\sum_{i=1}^{28} PCE\_Cat_{i,2019}$ represents the total nominal personal consumption expenditures

(PCE) (denominated in 2012 dollars) in certain categories of all 28 COBOL states in 2019. I exclude five PCE categories because they may not be reflected in total card consumption: (1) motor vehicles and parts, (2) housing and utilities, (3) health care, (4) financial services, and (5) final consumption expenditures of nonprofit institutions serving households.[24] These PCE category exclusion results in credit and debit card consumption accounting for roughly 50% of PCE. I multiply by 0.8 because the post period, March 13, 2020 to December 31, 2020, corresponds to four-fifths of a year. To convert real GDP denominated from 2012 dollars into real GDP denominated in 2019 dollars, I divide the 2019 implicit price deflator for GDP by the 2012 implicit price deflator for GDP: $\frac{PI_{2019}}{PI_{2012}}$. The estimated 2.8-percentage-point relative decline, $Relative\_Decline$, in aggregate consumption then translates into a real-GDP decline of $105 billion in COBOL states relative to non-COBOL states; this figure corresponds to roughly 0.9% of real GDP in COBOL states. This estimate is conservative as I assume there was no consumption effect in the excluded PCE categories.[25] However, the $105 billion cost estimate likely underestimates the true overall economic costs of administrative burdens in UI-benefit systems given that non-COBOL states also experienced increased administrative burdens after the emergency declaration that resulted in delays. The estimate instead represents the GDP decline that could have been avoided if COBOL states had modernized their UI benefit systems to the same extent as non-COBOL states, which would have resulted in a lower administrative burden for claimants in COBOL states.

The cost that COBOL states would have incurred to modernize their systems is likely only a small fraction of the $105 billion relative decline in GDP that they incurred. The problems with the COBOL systems were apparent after the Great Recession, but issues that arose from the

[24]Table 2.1 suggests that COBOL and non-COBOL states had similar COVID-19 new case rates and new death rates.

[25]If I were to include all PCE categories, then the real-GDP decline would be closer to $210 billion.

pandemic recession provided renewed interest in COBOL states to modernize their UI benefit systems. For example, the Wisconsin Department Workforce Development, a COBOL UI state agency, signed a contract with Flexion Inc. in 2021 to modernize legacy IT systems that are largely written in COBOL.[26] The initial contract lasts one year, with three optional one-year renewals. According to the contract, the total proposed cost if the contract were renewed all three times is $16.5 million. If we assumed that Wisconsin's contract stays within budget and other states have similar modernization costs, then total modernization costs of all 28 COBOL states would be less than $500 million. This amount reflects the large discrepancy between the costs of modernization and the costs of having an antiquated UI benefit system to real economy.

2.5.4 COBOL-Induced Delays

One way that relatively larger administrative burdens in COBOL states could be affecting aggregate consumption is through longer processing delays in UI disbursement. One approach to determining whether COBOL states experienced longer processing delays in UI disbursement than non-COBOL states is to ascertain whether COBOL states had a relatively higher share of claims that were topcoded, meaning they experienced a processing delay of over 70 days. These delays are out of the norm; DOLETA does not keep track of delays beyond 70 days. Under normal circumstances, these topcoded claims only account for a small share of claims. In 2019, less than 5% of all intrastate regular UI claims were topcoded in each state in each month, as reflected in Figure 2.3. However, from July 2020 to October 2020, at least 20% of all intrastate claims were topcoded in both COBOL and non-COBOL states. Representing topcoding as a share might understate the severity of the issue given the drastic increase in the number of claimants receiving

[26] See the public contract.

UI after the emergency declaration in 2020.

Figure 2.3 shows both that topcoding was more common in COBOL states and how common it was for claimants to have to wait over 70 days in both COBOL and non-COBOL states. In both COBOL and non-COBOL states, fewer than 2.5% of intrastate claims were delayed over 70 days in every month from January 2019 to February 2020 for the aggregated COBOL and non-COBOL states. In each of the last six months of 2020, for both COBOL and non-COBOL states, over 15% of intrastate claims were delayed by more than 70 days. Given that topcoded claims are a lagging indicator, I would not expect to see a spike in topcoded claims until 70 days after the emergency declaration. COBOL states experienced a higher share of topcoded claims than non-COBOL states starting in June 2020, when the March 2020 claims would have first been topcoded. Importantly, weekly initial claims for every week in 2020 after the emergency declaration were higher than the maximum number of claims prior to the emergency declaration, so both the numerator and denominator of the fraction of intrastate regular claims that are topcoded drastically increased in June 2020. Also, March, April, and May 2020 mechanically saw the fraction of topcoded claims fall because of the surge of new claims being processed, which is reflected in the denominator; however, topcoded claims, the numerator, cannot appear until 70 days after the spike in claims.

Figure 2.4 focuses on a larger set of claims that were delayed: claims that experienced at least a 5 week processing delay. It is harder to see the pattern in Figure 2.4 relative to Figure 2.3, but COBOL states experience a larger share of claims that experience at least a 5 week delay. Given that these delays encompass a larger set of claims, it is unsurprising that the values are by construction are higher relative to the topcoded case. For example the maximum value in Figure 2.4 is around 40%, while the maximum value in Figure 2.3 is around 25%. One should note that when analyzing claims that are at least delayed by 5 weeks that the range in delays mechanically increases relative to claims that are topcoded. Delays now range from 5 weeks to over 10 weeks. Unlike the topcoding analysis, I only exclude March 2020 from the sample given that claims made early in the pandemic that are delayed by at least 5 weeks could appear in April 2020 (unlike topcoded claims). Given this heterogeneity in lags in conjunction with COBOL states experiencing more topcoded claims, non-COBOL states peak sooner than COBOL states (June 2020 peak for non-COBOL states and August 2020 peak for COBOL states).

In Table 2.3, I show results estimated using the TWFE model described in Equation 2.1. Column 1 presents results with only state and month fixed effects as controls. The parameter of interest is the coefficient on the interaction of COBOL and Post. The coefficient of 2.3 corresponds to COBOL states experiencing a 2.3-percentage-point increase in the share of topcoded claims after the emergency declaration relative to non-COBOL states. This is a large increase in the share of topcoded claims given that on average states only delayed claims by over 70 days for about 18.7 percent of claims from June 2020 to December 2020.[27]

As discussed above, one concern with comparing COBOL and non-COBOL states is that they differ politically, which could affect the results. In column 2, I add the interaction of Republican and Post, and the coefficient on the interaction of COBOL and Post does not meaningfully change, going from 2.3 to 2.1 and remaining significant at the 1% level. The coefficient on the interaction of Republican and Post is insignificant.

Despite topcoded claims only accounting for a small fraction of regular UI intrastate claims prior to 2020, there were differences between COBOL and non-COBOL states. Specifically, COBOL states had a lower share of topcoded claims prior to the emergency declaration, as re-

[27] Note that this 18.7 percent statistic is not population weighted.

flected in Figure 2.2. Not only are the average shares different between COBOL and non-COBOL states, but there appears to be convergence in the average shares right before the emergency declaration. To address the concern that confounders are driving the differences in topcoding, I add five previously discussed confounders interacted with Post in column 3 of Table 2.3.[28] In column 3, once I add those five confounders interacted with Post, the coefficient on the interaction of COBOL and Post increases to 3.8 ppt. These interaction terms meaningfully change the point estimate on the coefficient of interest. A potential reason for this change could be that laid-off workers in accommodation and food services have more complicated work histories that could lead to longer processing delays than a typical claim, and non-COBOL states have more workers in accommodation and food services, as reflected by Table 2.1. Another interpretation of this increase is that by adding these confounders, I can partially address the non-classical measurement error previously described with the introduction of these control variables.

Although COBOL and non-COBOL states had similar unemployment rates on average during the pandemic, as shown in Table 1, as a robustness check, I include the unemployment rate as an additional control in column 4 of Table 2.3. By including the unemployment rate, the coefficient on the interaction of COBOL and Post increases to 4.4 ppt and remains significant at the 1 percent level. I find that higher unemployment rates are associated with a higher share of topcoded claims, but the coefficient on the interaction term of COBOL and Post does not meaningfully change. Even though the unemployment rate is potentially endogenously affected by UI processing delays, longer processing delays in COBOL states could lead claimants to return to work sooner or they could increase the unemployment rate through the multiplier channel. However, it is reassuring to see that the larger share of topcoded claims in COBOL states is not

[28] I use these same five confounders throughout my analyses for both delay and consumption outcomes.

exclusively being driven by the unemployment rate.

I perform a similar analysis in Table 2.4 with the share of claims that were delayed at least 5 weeks as the dependent variable. I use the same confounders and I find that COBOL states experienced between a 3.1 ppt. increase and a 4.4 ppt. increase in the share of claims with at least a 5 week delay relative to non-COBOL states after the emergency declaration. These results are significant at the 5 percent level across all specifications.

In sum, I find that COBOL states experienced longer delays in the form of higher share of topcoded claims and higher share of claims delayed by at least 5 weeks relative to non-COBOL states after the emergency declaration. The covariates included in Table 2.3 do not meaningfully change the significance of the coefficient on the interaction of COBOL and Post, but they do meaningfully change the point estimates.

2.5.5 Robustness Check on Consumption Results

As a robustness check of the results in Table 2.2, I use the penalized synthetic control method developed by Abadie & L'Hour (2021) to measure the decline in relative consumption for COBOL states relative to non-COBOL states after the emergency declaration.

This method uses covariates in the pre-intervention period and the donor pool to create a synthetic control for each treated unit. In my setting, I have 28 treated units, COBOL states, and 22 control units in the donor pool, non-COBOL states. The innovation of the penalized synthetic control method over the traditional synthetic control method is that there is a tuning parameter, λ, that puts additional weight on pairwise comparisons instead of the aggregate comparison. The higher the value of this tuning parameter, the more sparse the synthetic controls will be, and fewer non-COBOL states will be selected from the donor pool. As the tuning parameter approaches 0, the penalized synthetic control method becomes the traditional synthetic control method that minimizes the sum of pairwise discrepancies. As the tuning parameter approaches ∞, the penalized synthetic control method becomes the nearest-neighbor matching with replacement estimator.

The penalized synthetic control method is similar to the traditional synthetic control method in that they are both heavily dependent on the controls selected from the pre-intervention period. These controls affect which non-COBOL states are selected in the synthetic control in addition to the weight assigned in the synthetic control. Typically, more controls are used in a synthetic control setting than in a TWFE setting given the lack of fixed effects. I select 15 covariates to match on: (1) Republican vote share (2016), (2) income share in accommodation and food services (2019), (3) the percentage of the population living in urban areas (2010), (4) UI generosity (Jan. 2020), (5) the percentage of the population living in poverty (2019), (6) the percentage of the population with at least a bachelor's degree (2019), (7) the employment-to-population ratio (2019), (8) the log of income per capita (2019), (9) median age (2019), (10) the African American population share (2019), (11) the relative replacement rate (2020), (12) teleworkable employment (2019), (13) a Republican governor indicator (2019), (14) labor force population, and (15) real GDP (2019).

Table 2.5 reports the relative consumption decline for COBOL states using the penalized synthetic control method. The last three columns report results using this method. The column labeled PSC fixed λ corresponds to a fixed value for the tuning parameter of 0.1. The other two penalized synthetic control estimator columns choose the tuning parameter in a data-driven manner. The column labeled PSC MSE λ uses a leave-one-out cross-validation procedure to select λ by minimizing the mean squared prediction error in the post-intervention period (after the emergency declaration). The column labeled PSC Bias λ uses validation over the outcomes (relative consumption) in the pre-intervention period (prior to the emergency declaration) to select the tuning parameter. The average treatment effects across these three specifications with different tuning parameters yield a relative decline in consumption for COBOL states of between 3.7 and 4.8 percentage points. The results from the penalized synthetic control method are not meaningfully different than the results from Table 2.2.

To conduct inference with the penalized synthetic control method, permutation tests are typically conducted. I randomly assign treatment across 28 of the 50 states 10,000 times and estimate a relative consumption decline using a tuning parameter identical to the one from the column labeled PSC MSE λ in Table 2.5 (0.01) in each iteration. To be consistent with the results from Table 2.5, I aggregate the 28 cohort treatment effects using population weights. Figure 2.6 shows the distribution of these 10,000 simulations. The red dashed line corresponds to the average treatment effect for the 28 COBOL states with a tuning parameter of 0.01. This permutation test yields an effect that is significant at the 10% level.

2.5.6 Heterogeneity Analysis

Throughout this analysis, I have focused on relative consumption for all consumption categories among all consumers. However, one of the benefits of using the Affinity Solutions data is that consumption at the state level can be decomposed by type of goods purchased, by type of services purchased, or by income quartile.[29] I focus on each income quartile and the four mutually exclusive aggregated consumption types defined by Chetty et al. (2020): durable goods, non-

[29]Due to data limitations, I cannot decompose results by goods or services within an income quartile at the state level using the Economic Tracker.

durable goods, remote services, and in-person services. If discouraged filers are playing a role in the relative consumption decline, then I expect durable-goods consumption to be negatively affected by the large UI benefit transfers during the pandemic recession. [30] Parker et al. (2013) found a shift towards durable goods under the Economic Stimulus Act of 2008 where the typical single household received $300 to $600. Unemployment Insurance benefits between April 2020 and July 2020 were much larger because of the provision granting unemployed workers an extra $600 per week in benefits. I also decompose consumption by income quartiles because income groups were differentially exposed to the COVID-19 shock. I expect consumption of in-person services to have decreased more in COBOL states than non-COBOL states because when households receive a negative income shock, they reduce their consumption at restaurants and on entertainment, which count as in-person services. In addition, studies such as Amburgey, Birinci et al. (2020) find that the top quintile had the smallest shift in its unemployment rate during the pandemic recession. If one was not unemployed, they would have been ineligible for UI benefits and could not have suffered from delayed UI benefits nor become a discouraged filer. I therefore expect the richest income quartile, quartile 4, to experience at most a small drop in relative consumption in COBOL states relative to non-COBOL states.

I use the Affinity Solutions data to see whether durable-goods consumption was affected by having an antiquated UI system. Durable-goods consumption is defined as consumption in the following merchant category codes: (1) building materials, gardening equipment, and supplies; (2) electronics and appliances; (3) furniture and home furnishings; (4) sporting goods, hobbies, musical instruments, and bookstores; (5) telecommunications; and (6) vehicles and parts.

[30] Studying the effects of the Economic Stimulus Act of 2008, Parker, Souleles, Johnson & McClelland (2013) find that about 50%–90% of household payments under the act went to durable goods and related services.

Nondurable-goods consumption is defined as consumption in the following codes: (1) clothing and clothing accessories; (2) food and beverage stores; (3) general merchandise; (4) health and personal care stores; and (5) wholesale trade. Remote-services consumption is defined as consumption in the following codes: (1) administrative and support and waste management and remediation services; (2) education; (3) finance and insurance; (4) information; (5) professional, scientific, and technical; (6) public administration; and (7) utilities, construction, and manufacturing. In-person-services consumption is defined as consumption in the following codes: (1) accommodation and food services; (2) healthcare and social assistance; (3) arts, entertainment, and recreation; (4) transportation and warehousing; (5) rental and leasing; (6) repair and maintenance; and (7) personal and laundry services.

To formally estimate the effect of increased administrative burden by type of good or service consumed, I use a TWFE estimator. Specifically, I use a TWFE estimator similar to the one in Equation 2.1 to estimate heterogeneous relative consumption differences between COBOL and non-COBOL states. The only difference is that the dependent variable changes from relative consumption in all categories to relative consumption in one of these four aggregated categories. In all four specifications, I match the controls used in column 2 of Table 2.2 where I include state and day fixed effects in addition to controlling for the interaction of Republican and Post. Column 1 of Table 2.6 provides suggestive evidence that increased administrative burden from antiquated systems reduced durable-goods consumption. The coefficient on the interaction of *COBOL* and *Post* is marginally insignificant at the 10% level (significant at 10.2%level). The coefficient corresponds to a 2.4-percentage-point decline in durable-goods consumption in COBOL states relative to non-COBOL states. The patterns in durable goods consumption differ from those of aggregate consumption in that durable goods consumption declined sharply at the start of the pandemic but quickly recovered, topping pre-pandemic levels soon afterward. By the end of May, durable-goods consumption was above baseline values and remained elevated for the remainder of the sample period for both COBOL and non-COBOL states. This suggestive finding of a relative decline in consumption of durable goods may suggest more discouraged filers in COBOL states. Durable goods are large purchases that arguably are less sensitive to delays in the disbursement of UI benefits.[31]

Unlike durable goods, nondurable goods should be affected by both discouraged filers and delayed payments. In column 2 of Table 2.6, I estimate the effect of the increased administrative burden in UI benefit systems on nondurable goods consumption. The table shows a relative decline of 3.5 percentage points more in COBOL states than non-COBOL states, which is significant at the 5% level. If households are not able to perfectly smooth consumption, then they will reduce their consumption of nondurable goods prior to the receipt of their delayed UI benefits. Column 3 estimates the impact on in-person services, showing a 2.8-percentage-point relative decline, which is significant at the 10% level. Consumers typically reduce their consumption of in-person services such as dining in restaurants when they receive a negative income shock. In column 4, I estimate the impact on remote services and find no effect.

Instead of looking at the goods or services purchased, as in Table 2.6, I next examine which individuals experienced the largest declines in relative consumption. Specifically, I sort the individuals into income quartiles.[32] I rerun the analysis from Table 2.6 using income quartiles as the dependent variable. I report the results in Table 2.7. Column 1 corresponds to consumption in

[31] Note that delays may also partially encourage durable goods consumption because the first payment of delayed UI payments will be larger than UI payments that are not delayed. For example, if an individual is entitled to 8 weeks of UI, but their claim is delayed over 10 weeks (topcoded), then the recipient will receive all their UI benefits in one lump sum transfer.

[32] Two states, Alaska and Hawaii, are omitted from the sample because consumption data for the bottom quartile are unavailable.

the bottom quartile, column 2 corresponds to consumption in the second income quartile, and so forth. All the specifications include state and day fixed effects as well as the interaction of Republican and Post as an additional control. The results show that as the income quartiles increase, the standard errors shrink. Even though the top income quartile has the smallest standard error, I find an insignificant result, as expected. The richest income quartile was the least likely to become unemployed during the pandemic recession and thus the least likely to receive UI benefits, independent of administrative burdens. The strongest effects are for the second- and third-income quartiles. There are two possible reasons for not finding a strong result for the bottom quartile: (1) larger standard errors or (2) fewer discouraged filers. It could be that unemployed individuals in the bottom quartile were less susceptible to an increase in administrative burden because they were more likely to depend on the benefits to cover necessities such as rent payments and food expenses. Administrative burdens in programs like UI may be a form of targeting (Nichols & Zeckhauser, 1982) in which only the most motivated claimants overcome all the hurdles. In columns 2 and 3, I find a 3.2- and 2.7-percentage-point decline for the second and third income quartiles, respectively, both of which are significant at the 10% level.

2.6 Conclusion

Using a TWFE estimator, I find that problems with UI-benefit systems in states whose systems ran on COBOL resulted in a decline in relative consumption after the emergency declaration that was 2.8 percentage points larger than in non-COBOL states. I cannot definitively attribute all the difference to the use of COBOL, since COBOL states' systems may have also been antiquated in other ways; however, my results do clearly show that not having a well-functioning UI benefit system during the pandemic meaningfully harmed Americans. My results illustrate the economic consequences of only the increased administrative burden on potential claimants, but my results do not capture the nonpecuniary costs that potential claimants faced in COBOL states by repeatedly being disconnected, losing time in filing a claim, and experiencing added uncertainty regarding whether and when they would receive their benefits.[33] The large negative effect of UI system delays in COBOL states during the pandemic strongly suggests that the effectiveness of UI benefit systems as a countercyclical tool.[34]

Given data limitations, I cannot decompose how much processing delays contributed to the relative decline in consumption in COBOL states (relative to non-COBOL states) during the pandemic. COBOL states experienced longer delays in disbursing benefits and could have also had relatively more discouraged filers due to increased administrative difficulties. In particular, relative consumption declines among UI-eligible claimants in COBOL states may have led to a lower UI fiscal multiplier relative to non-COBOL states. Such an effect would dampen consumption even among households that remained employed and households that received benefits promptly. Future work could potentially decompose the overall decline in consumption into these components.

One potential policy approach is to create federal incentives for states to update their UI systems. With the American Recovery and Reinvestment Act of 2009, the federal government made $7 billion available for states who chose to modernize.[35] Thirty-nine states chose to imple-

[33]NJ Labor Comissioner Asaro-Angelo discussed how his team received death threats from claimants frustrated with issues in filing their claims during the pandemic recession in a panel discussion held by the Heldrich Center for Workforce Development on UI systems in New Jersey.

[34]How a UI benefit system should modernize lies outside the scope of this chapter. Other work has shown that states should allocate enough time for modernization and incorporate extensive user testing throughout the process (Simon-Mishel, Emsellem, Evermore, Leclere & Coven, 2020).

[35]States had until August 22, 2011, to submit their applications to the Department of Labor to receive funding. See the briefing paper from the National Employment Law Project on how federal incentives after the Great Recession

ment at least some of the changes required to receive funding. These changes included expanding the definition of eligible workers to include part-time workers. Only $4.4 billion was allocated, but the program shows that the financial cost of enticing states to change their UI-benefit system is only a fraction of the cost incurred during the pandemic recession by COBOL states. Instead of focusing first on transitioning away from COBOL, states could focus on other less costly issues that likely contribute to delays, such as not having all UI information available in commonly spoken languages, not making it easy to reset passwords, and not making it possible to complete the entire online filing process on mobile devices. These other issues are less costly to fix and can cause large problems such as locking out a claimant from the state UI website, which leads to delays in filing.

Another approach is for states to join a consortium in which they share the same UI system and split the maintenance costs. A few states have taken this approach. For example, Mississippi, Rhode Island, Maine, Connecticut, and Oklahoma have joined together in the ReEmploy USA Consortium.[36] A more radical approach is to form a federal UI system, which might—or might not—improve efficiency. States might oppose this idea, especially if they are trying to decrease their UI uptake rate by discouraging eligible claimants.

Although lack of UI system modernization is a central problem, Lachowska, Mas & Woodbury (2022) show that modernization by itself is not necessarily sufficient to fix UI administrative issues. It is important that federal funding to state workforce agencies be tied to a pay-for-performance scheme to achieve the outcomes desired, such as shorter processing delays or reduced call center volumes. Some of these approaches are already being implemented through

helped states modernize their UI benefit systems.

[36]Though five states joined the consortium, only Maine and Mississippi had fully implemented the software program to modernize their UI benefit system away from COBOL by the start of the pandemic.

a grant of up to $600 million to support state UI information technology modernization under the American Rescue Plan Act (Parton, 2023). However, it is important that policymakers understand the importance of modernizing these antiquated UI systems.[37] Regardless of which approach policymakers choose to take, antiquated UI systems hamper the effectiveness of UI as a fiscal stabilizer.

[37]For example, the initial support of $600 million available in grant opportunities was reduced to $200 million through the Fiscal Responsibility Act of 2023.

Table 2.1: Unweighted Summary Statistics by COBOL Usage

	Non-COBOL	COBOL
Relative Consumption	-5.44	-7.51
	(9.82)	(10.08)
Fraction Topcoded	10.74	13.23
	(12.93)	(16.27)
Relative First Payments (Ratio)	7.79	6.74
	(11.95)	(10.48)
New COVID-19 Death Rate	0.29	0.29
	(0.33)	(0.46)
New COVID-19 Case Rate	18.81	18.10
	(23.75)	(25.40)
Unemployment Rate	7.76	7.73
	(4.04)	(3.69)
Population (Thous.)	5,800.01	7,143.62
	(4,650.14)	(8,809.51)
Republican (2016)	50.59	48.11
	(9.06)	(10.89)
Urban (2010)	72.56	74.38
	(13.72)	(14.90)
UI Generosity (Jan. 2020)	10154.82	12470.57
	(4,710.13)	(3,378.89)
Acc. and Food Services Inc. Share (2019)	4.14	3.70
	(2.40)	(1.46)
Bachelor's Degree (2019)	31.23	32.90
	(5.09)	(5.32)
Poverty (2019)	12.43	11.88
	(3.14)	(2.11)

Note: The table provides summary statistics for the variables used in my main specification and covariates used as controls. Relative consumption, the new COVID-19 death rate, and the new COVID-19 case rate come from the Economic Tracker (Chetty et al., 2020). Fraction of intrastate claims that are topcoded and relative first payments come from DOLETA. Relative consumption is measured in percentage point changes from the index period of January 2020. State population estimates for 2019 come from the US Census Bureau. The relative consumption variable is identical to the all-spending variable in (Chetty et al., 2020). The monthly unemployment rate estimates come from the Bureau of Labor Statistics. The remaining covariates are cross-sectional data from a point in time prior to the emergency declaration. The reported statistics are the means of the corresponding group, with their standard deviations in parentheses. These summary statistics cover the sample period of February 1, 2020, to December 31, 2020.

Table 2.2: TWFE COBOL Usage on All Card Consumption

	(1) Rel Cons	(2) Rel Cons	(3) Rel Cons	(4) Rel Cons	(5) Rel Cons
COBOL × Post	-0.041**	-0.028**	-0.028**	-0.026**	-0.024**
	[0.020]	[0.013]	[0.013]	[0.011]	[0.011]
Republican × Post		0.003***	0.003***	0.003***	0.004***
		[0.001]	[0.001]	[0.001]	[0.001]
UR					0.002
					[0.002]
State FE	Yes	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes	Yes
COVID-19 Controls	No	No	Yes	Yes	Yes
State Char. × Post	No	No	No	Yes	Yes
Days	335	335	335	335	335
States	50	50	50	50	50
Observations	16,750	16,750	16,750	16,750	16,750

Note: The table presents results from a TWFE estimator with day and state fixed effects. The dependent variable is the percentage-point change relative to the base period in credit and debit card consumption measured at the daily frequency. *Post* is a binary variable that takes the value 1 if the date is on or after March 13, 2020. *COBOL* is a binary variable that takes the value 1 if a state uses COBOL in its UI benefits system. The interaction term of interest is the product of *COBOL* and *Post*. *Republican* is the Republican vote share in the 2016 presidential election. COVID-19 controls include new COVID-19 death rates and new COVID-19 case rates. Column 1 only includes state and day fixed effects. Column 2 adds the interaction of *Republican* and *Post*. Column 3 adds the COVID-19 controls. Column 4 adds five terms of *Post* interacted with another confounder: (1) income share in accommodation and food services (2019), (2) mask mandates in July 2020 (2020), (3) the percentage of the population living in poverty (2019), (4) the percentage of the population with at least a bachelor's degree (2019), and (5) UI generosity (Jan. 2020). Column 5 adds the monthly unemployment rate. These estimates cover the sample period of February 1, 2020, to December 31, 2020. State populations in 2019 are applied as analytic weights. Standard errors are clustered at the state level.

Standard errors: *** p<0.01, ** p<0.05, * p<0.1

Table 2.3: TWFE Fraction of Claims Topcoded

	(1) Frac Intra Top	(2) Frac Intra Top	(3) Frac Intra Top	(4) Frac Intra Top
COBOL × Post	2.3***	2.1***	3.8***	4.0***
	(0.53)	(0.61)	(0.72)	(0.62)
Republican × Post		-0.0	-0.4***	-0.3***
		(0.06)	(0.09)	(0.08)
UR				0.8***
				(0.22)
State FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
State Char. x Post	No	No	Yes	Yes
Obs.	1050	1050	1050	1050
Depvar	0.88	0.88	0.88	0.88

Note: This table relies on first-payment time-lapse data from the Department of Labor Employment and Training Administration's 9050 reports. The dependent variable is the fraction of intrastate claims that are topcoded, that is, delayed by over 70 days. All specifications correspond to a TWFE estimator with state and month fixed effects. Column 1 does not include any additional controls. Column 2 includes the interaction of 2016 presidential Republican vote share and Post. Column 3 adds multiple interaction terms of post and another confounder: (1) income share in accommodation and food services (2019), (2) the percentage of the population living in urban areas (2010), (3) UI generosity (Jan. 2020), (4) the percentage of the population living in poverty (2019), and (5) the percentage of the population with at least a bachelor's degree (2019). Column 4 adds the unemployment rate. The sample starts in January 2019 and ends in December 2020, with the post-period starting in June 2020. Given the spurious nature of topcoding being a lagging indicator, for 2020, I drop March, April, and May from the sample. Depvar corresponds to the average value of the fraction of topcoded claims from January 2019 to February 2020 across all 50 states (unweighted). The standard errors are clustered at the month level.

Standard errors: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.4: TWFE Fraction of Claims Delayed at Least 5 Weeks

	(1) Frac Intra Top	(2) Frac Intra Top	(3) Frac Intra Top	(4) Frac Intra Top
COBOL $\times$ Post	3.1**	3.3**	4.3***	4.4***
	(1.42)	(1.54)	(1.13)	(1.08)
Republican $\times$ Post		0.0	-0.3**	-0.3**
		(0.07)	(0.11)	(0.12)
UR				0.2
				(0.65)
State FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
State Char. x Post	No	No	Yes	Yes
Obs.	1150	1150	1150	1150
Depvar	4.83	4.83	4.83	4.83

Note: This table relies on first-payment time-lapse data from the Department of Labor Employment and Training Administration's 9050 reports. The dependent variable is the fraction of intrastate claims that are delayed at least 5 weeks. All specifications correspond to a TWFE estimator with state and month fixed effects. Column 1 does not include any additional controls. Column 2 includes the interaction of 2016 presidential Republican vote share and Post. Column 3 adds multiple interaction terms of post and another confounder: (1) income share in accommodation and food services (2019), (2) the percentage of the population living in urban areas (2010), (3) UI generosity (Jan. 2020), (4) the percentage of the population living in poverty (2019), and (5) the percentage of the population with at least a bachelor's degree (2019). Column 4 adds the unemployment rate. The sample starts in January 2019 and ends in December 2020, with the post-period starting in April 2020. Given the spurious nature of topcoding being a lagging indicator, I drop March 2020 from the sample. Depvar corresponds to the average value of the fraction of claims delayed at least 5 weeks from January 2019 to February 2020 across all 50 states (unweighted). The standard errors are clustered at the month level.

Standard errors: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.5: Penalized Synthetic Control Method, Results

	Treated	Control	PSC fixed λ	PSC MSE λ	PSC Bias λ
Sample Size	28	22	18	22	20
Republican	44.43	48.80	48.88	49	48.42
Urban	82.95	77.45	81.61	81.46	82.31
UI Generosity	12.3	9.11	7.02	7.55	7.31
ACF Incshare	3.37	3.930	4.290	4.360	4.220
Poverty	12.20	12.540	12.650	12.470	12.400
Education	33.81	31.810	31.330	31.380	31.620
EPOP	46.01	45.540	44.340	44.230	44.670
Income Per Cap.	10.96	10.87	10.89	10.9	10.89
Median Age	37.72	38.92	39.73	40.14	39.72
AA Pop. Share	0.11	0.140	0.150	0.140	0.150
Rel. Rep. Rate	1.01	1.09	1.13	1.12	1.12
Teleworkable Emp.	0.37	0.350	0.350	0.350	0.350
Republican Governor	0.46	0.780	0.750	0.900	0.800
Labor Force Pop.	8,903,412	4,695,989	6,589,242	6,579,186	6,578,733
Real GDP	1,179,523	484,750	652,204	647,904	650,749
Treatment Effect	NA	-0.041	-0.039	-0.048	-0.037
λ	NA	NA	0.100	0.000	0.010
Min. Density	NA	NA	1	1	1
Median Density	NA	NA	2	22	2
Max. Density	NA	NA	3	22	4

Note: The table presents results from the penalized synthetic control method (Abadie & L'Hour, 2021) in comparison and the traditional TWFE. The sample size corresponds to how many states are used to create synthetic control groups. In the TWFE setting, the control group is all 22 non-COBOL states. With the penalized synthetic control method, not all states from the donor pool may get selected. For both the traditional TWFE estimator and the penalized synthetic control method, the 28 treated states are the COBOL states. Fifteen covariates are used for creating synthetic controls that are measured prior to the emergency declaration. The one parameter changing across the three penalized synthetic control methods is the penalization parameter: λ. This parameter makes a trade-off between the component-wise fit (to a COBOL state) and aggregate fit (to all COBOL states). The column labeled "PSC fixed λ" corresponds to a fixed value for λ of 0.1. The other penalized synthetic control estimator columns choose lambda in a data-driven manner. One uses a leave-one-out cross-validation procedure to select λ by minimizing the mean squared prediction error in the post-intervention period (after the emergency declaration). The other method chooses λ on validation over the outcomes (relative consumption) in the pre-intervention period (prior to the emergency declaration). All five confounders as well as the 2016 Republican vote share are included in this analysis. The density refers to the number of non-COBOL states used for creating the synthetic control of the COBOL states. For example, a maximum density of 22 refers to at least one COBOL state using all 22 non-COBOL states in its synthetic control. All results use 2019 population weights.

Table 2.6: TWFE COBOL Usage by Consumption Type

	(1) Durables	(2) Nondurables	(3) Inperson_serv	(4) Remote_serv
COBOL × Post	-0.024	-0.034**	-0.028*	-0.014
	[0.014]	[0.014]	[0.015]	[0.016]
Republican × Post	0.002***	0.001	0.006***	0.001
	[0.001]	[0.001]	[0.001]	[0.001]
State FE	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes
Days	335	335	335	335
States	50	50	50	50
Observations	16,750	16,750	16,750	16,750

Note: The table provides results from a two-way fixed-effects (TWFE) estimator with day and state fixed effects, where consumption is broken down by consumption type. The dependent variable is the percentage-point change in a type of credit and debit card consumption (measured at a daily frequency) relative to the base period (January 2020). Column 1 corresponds to durable-goods consumption, column 2 to nondurable-goods consumption, column 3 to in-person services consumption, and column 4 to remote-services consumption. *Post* is a binary variable that takes the value 1 if the date is on or after March 13, 2020. *COBOL* is a binary variable that takes the value 1 if a state uses COBOL in its UI benefits system. The main interaction term is the product of *COBOL* and *Post*. As an additional control in all specifications, I interact *Post* and the 2016 Republican presidential election vote share. These estimates cover the sample period of February 1, 2020, to December 31, 2020. State populations in 2019 are applied as analytic weights. Standard errors are clustered at the state level.

Standard errors: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

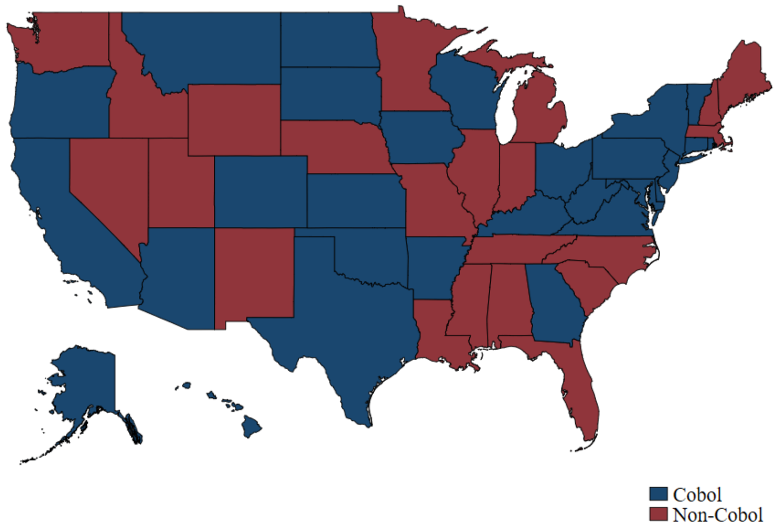
Table 2.7: TWFE COBOL Usage on All Card Consumption by Income Quartiles

	(1) Rel Cons (Q1)	(2) Rel Cons (Q2)	(3) Rel Cons (Q3)	(4) Rel Cons (Q4)
COBOL × Post	-0.014	-0.032*	-0.027*	-0.011
	[0.020]	[0.019]	[0.015]	[0.010]
Republican × Post	0.004***	0.002	0.002	0.002**
	[0.001]	[0.001]	[0.001]	[0.001]
State FE	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes
Days	335	335	335	335
States	48	50	50	50
Observations	16,080	16,750	16,750	16,750

Note: The table provides results from a two-way fixed-effects (TWFE) estimator with day and state fixed effects, where consumption is broken down by income quartiles. Column 1 corresponds to the bottom quartile, column 2 to the second quartile, column 3 to the third quartile, and column 4 to the top quartile. The dependent variable is the percentage-point change in credit and debit card consumption (measured at a daily frequency) for the relevant income quartile relative to the base period. *Post* is a binary variable that takes the value 1 if the date is on or after March 13, 2020. *COBOL* is a binary variable that takes the value 1 if a state uses COBOL in its UI benefits system. The main interaction variable is the product of *COBOL* and *Post*. As an additional control in all specifications, I interact *Post* with the 2016 Republican presidential election vote share. Alaska and Hawaii are omitted in column 1 because their consumption data are missing. These estimates cover the sample period of February 1, 2020, to December 31, 2020. State populations in 2019 are applied as analytic weights. Standard errors are clustered at the state level.

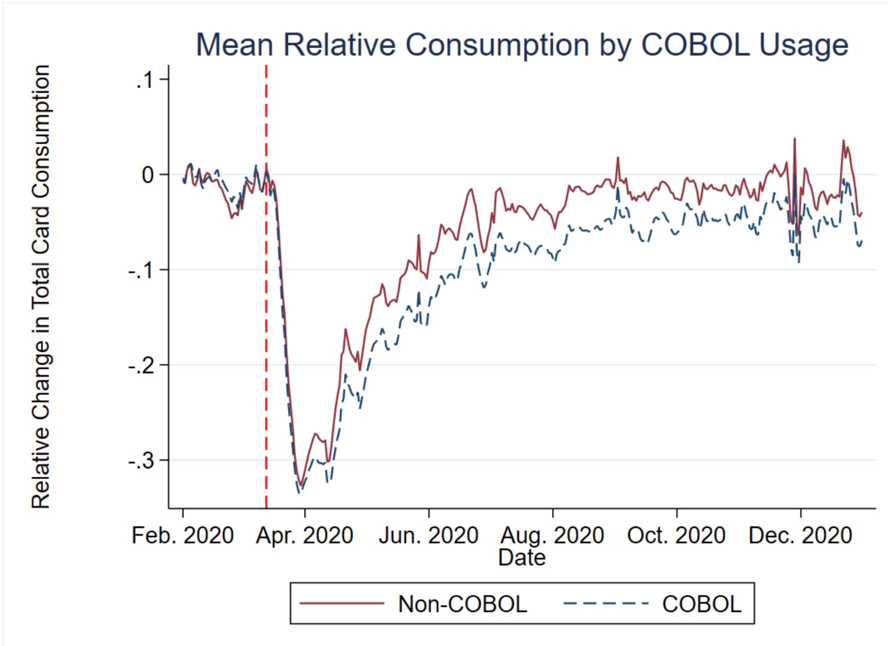
Standard errors: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 2.1: Map of COBOL Status



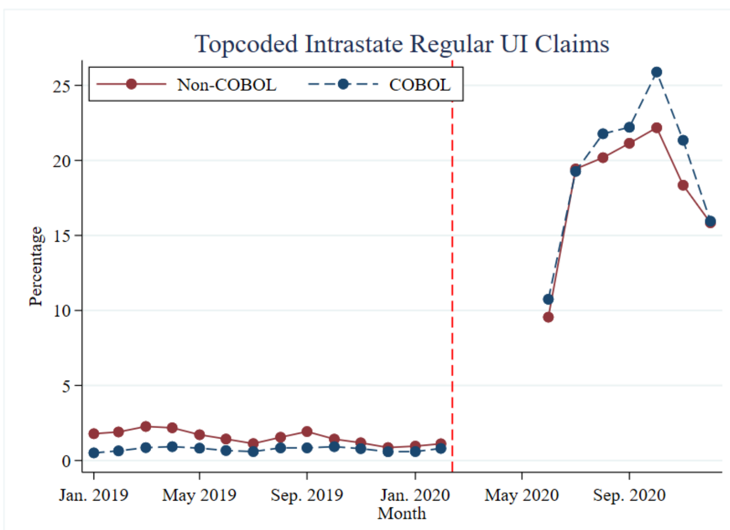
Note: The data on COBOL usage were collected by the author primarily from emails, news articles, and information from the UI Information Technology Support Center. Washington, DC, uses COBOL, but it is excluded from the analysis because of lack of consumption data.

Figure 2.2: Relative Credit and Debit Card Consumption for All Consumers



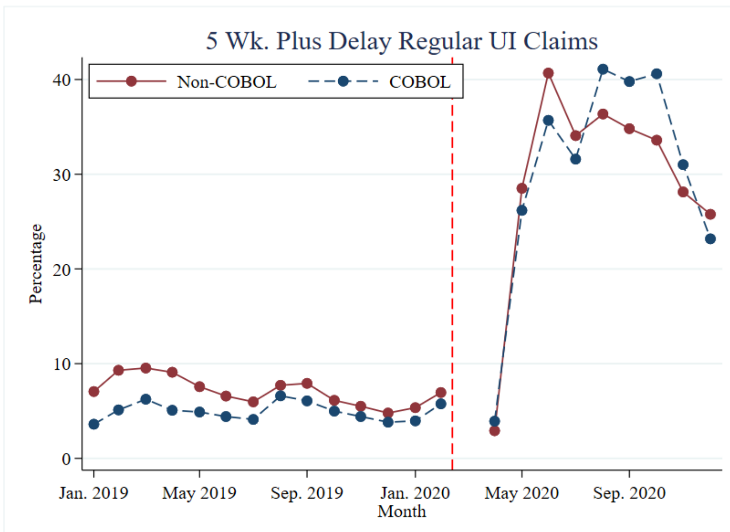
Note: This figure relies on data of all credit and debit card spending at a daily frequency for each state. A population-weighted average across states is computed when aggregating to COBOL and non-COBOL states. The dashed maroon line denotes states expected to experience longer delays and higher shares of discouraged filers: COBOL states. This figure covers the sample period from February 1, 2020, to December 31, 2020.

Figure 2.3: Percentage of Topcoded Claims (Processing Delays)



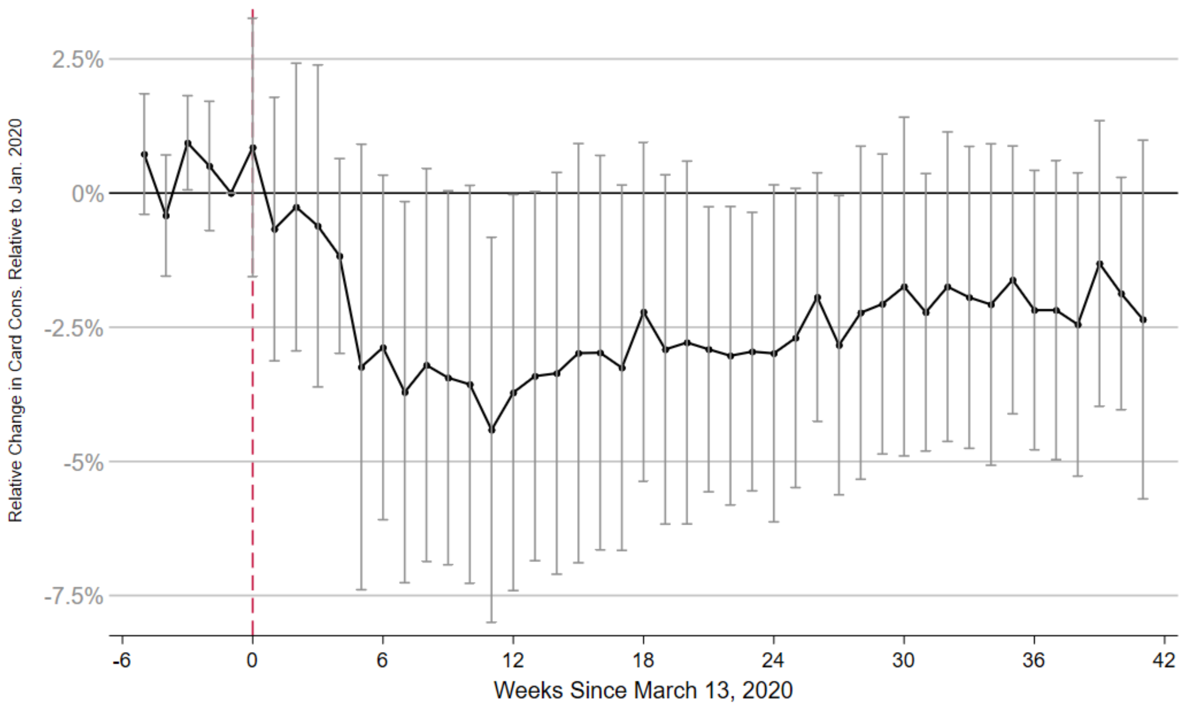
Note: This figure is based on first-payment time-lapse data from the Department of Labor Employment and Training Administration's 9050 reports. The groups are population weighted using 2019 Census estimates. The figure depicts the percentage of intrastate regular UI claims reported as having over a 70-day delay between January 2019 and December 2020 for COBOL and non-COBOL states. The vertical red dashed line corresponds to March 13, 2020. Because topcoding is a lagging indicator, for 2020, I drop March, April, and May from the sample.

Figure 2.4: Percentage of Claims Delayed at Least 5 Weeks (Processing Delays)



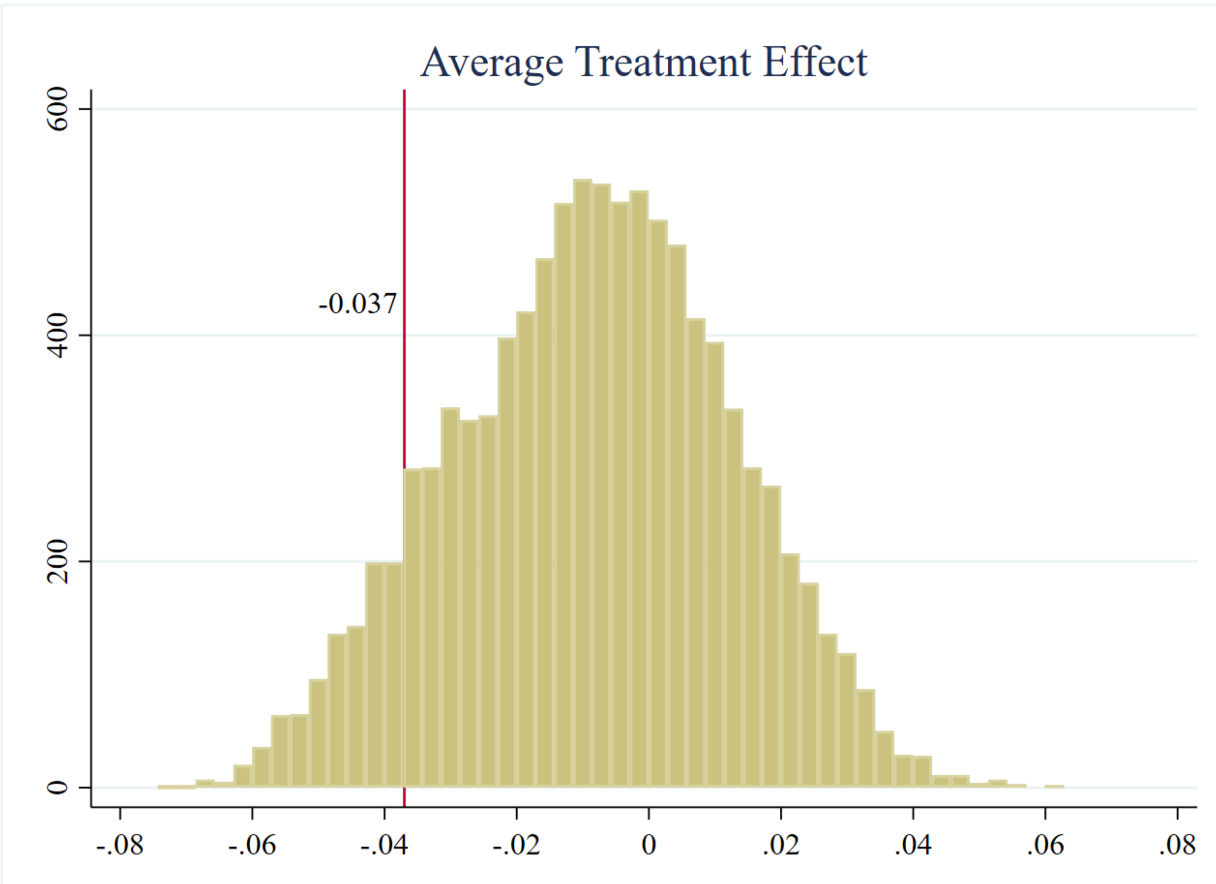
Note: This figure is based on first-payment time-lapse data from the Department of Labor Employment and Training Administration's 9050 reports. The groups are population weighted using 2019 Census estimates. The figure depicts the percentage of intrastate regular UI claims reported as having at least a 5 week processing delay between January 2019 and December 2020 for COBOL and non-COBOL states. The vertical red dashed line corresponds to March 13, 2020. Because topcoding is a lagging indicator, I drop March 2020 from the sample.

Figure 2.5: Relative-Consumption Weekly Event Study:
Relative Difference between COBOL and Non-COBOL States



Note: The graph is a coefficient plot showing the coefficient on β_k from Equation 2.3. State and week fixed effects are used in conjunction with an interaction term of *Post* and Republican presidential election vote share in 2016. The red dashed line that goes through week zero corresponds to March 13, 2020. This figure shows that in each week after the week of the emergency declaration, COBOL states saw a larger decline in relative consumption than non-COBOL states. The bounds on each point estimate correspond to a 95% percent confidence interval. State populations from 2019 are applied as analytic weights. Standard errors are clustered at the state level.

Figure 2.6: Permutation Test for Penalized Synthetic Control Method (10,000 Simulations)



Note: The histogram shows the distribution of average treatment effects when treatment is randomly assigned across 28 of the 50 states 10,000 times using the penalized synthetic control method. The tuning parameter is identical to the one from the column labeled PSC MSE λ in Table 2.5 (0.01) in each iteration. To be consistent with the results from Table 2.5, I aggregate the 22 cohort treatment effects using population weights. The red dashed line corresponds to actual treatment effect with the 28 COBOL states: a 3.7-percentage-point decline. This permutation test yields an effect that is significant at the 10% level.

Chapter 3: Unemployment Benefits Expansion and Business Formation

with Joonkyu Choi, Samuel Messer, and Veronika Penciakova

3.1 Introduction

Business startups play a critical role in job creation, innovation, and productivity growth in the United States. As such, the secular decline in the firm entry rate over the past several decades has received widespread attention.[1] Unexpectedly, the slowdown in startup activity has reversed substantially since the onset of the COVID-19 recession. Figure 3.1 shows that the number of new business applications surged starting in April 2020, and has remained elevated since. While this phenomenon has received much attention given its unprecedented scale and its potential implications for economic growth, its causes are not well understood.[2]

In this chapter, we use the U.S. Census Bureau's Business Formation Statistics (BFS) data on applications for Employer Identification Numbers (EINs), coupled with the Department of Labor Employment Training Administration (DOLETA) data on unemployment insurance (UI). We estimate the effect of antiquated UI benefit systems on business formation. We find a decline in business formation during the pandemic recession in states that used an antiquated UI benefit

[1] For example, see Decker, Haltiwanger, Jarmin & Miranda (2014), Alon, Berger, Dent & Pugsley (2018), Pugsley & Şahin (2019), Hopenhayn, Neira & Singhania (2022), and Akcigit & Ates (2023).

[2] See Decker & Haltiwanger (2023) for an in-depth and extensive documentation of this phenomenon.

system relative to states with a more modernized UI benefit system. The CARES Act, enacted in March 2020, contained several provisions that increased UI generosity. We hypothesize that, when taken together, these provisions incentivized some unemployed individuals to pursue their entrepreneurial ideas.

There are three key aspects of the UI expansion under the CARES Act that are relevant for this chapter. First, the CARES Act increased UI generosity. Specifically, UI benefit amounts were increased by an additional $600 per week until the end of July 2020 (Federal Pandemic Unemployment Compensation) and the maximum duration was extended by an additional 13 weeks (Pandemic Emergency Unemployment Compensation). Another important expansion of UI through the CARES Act was the creation of the Pandemic Unemployment Assistance (PUA) program, which expanded eligibility of UI to include gig workers, self-employed workers and contract workers.[3] These provisions enabled unemployed individuals to accumulate excess savings.[4] It is well documented in the entrepreneurship literature that in the presence of financial constraints, an increase in savings facilitates business entry (Evans & Jovanovic, 1989; Buera, 2009). Second, work search during this period was lower for several reasons such as fewer job vacancies, a pandemic, and UI recipients no longer having to search for work while receiving benefits. The work search requirement (i.e. that UI recipients provide proof of active job search) was waived in virtually all states in 2020 and in much of 2021 by the Families First Coronavirus Response Act (FFCRA).[5] Lastly, the UI expansion had significant stimulative effects on aggre-

[3]The risk of fraud was especially high for PUA because applicants could self-certify their eligibility and were not required to provide any documentation of self-employment or income as documented in the U.S. Government Accountability report GAO-23-106586.

[4]See Ganong et al. (2020a), Ganong, Greig, Noel, Sullivan & Vavra (2022b), and Aladangady, Cho, Feiveson & Pinto (2022) on the effect of the UI expansion on excess savings.

[5]It is important to note that the increase in the UI benefit amount did not necessarily raise the opportunity cost of entering entrepreneurship for unemployed workers, because they could still receive partial UI benefits particularly when the work search requirements were suspended.

gate demand (Ganong et al. (2022b) and Chapter 2), which raised expected profits for potential business entrants.

We estimate the effect of antiquated UI benefit systems on business formation. One limitation of using COBOL variation in the business formation setting is that we cannot decompose the effect. Potential discouraged filers in COBOL states could be driving these effects where not receiving any UI payments discourages business formation, smaller relative increases in UI payments in COBOL states could lead to lower business formation, or the delayed payments could lead to lower business formation. Our effect is most likely a combination of these effects. This affects how we interpret the findings of which component of antiquated UI benefit systems is driving the differences in business formation between COBOL and non-COBOL states. However, this limitation does not pose a threat to identification.

Using a two-way fixed-effects estimator (TWFE), we show that COBOL states experienced relatively lower business formation than non-COBOL states from March 2020 to December 2020. In our preferred specification, we estimate that the increase in business formation in COBOL states was 6.6 percentage points lower than in non-COBOL states from March 2020 to July 2020.[6] For the TWFE estimator to be a valid approach for identifying the causal effect of COBOL usage on business formation, the conditional parallel-trends assumption must hold. This means that, absent the surge in claims and changes to UI generosity (FPUC program) during the pandemic recession, business formation trends between states with and without antiquated UI benefit systems would have been the same conditional on the covariates included in our regression equation.[7]

[6]Note that all our results are robust to ending the sample in December 2020.

[7]In our preferred specification we control for state fixed effects and Census division by month fixed effects, but we find similar results when we only control for state fixed effects and month fixed effects.

A natural question is whether UI benefits expansion affects not just the quantity, but also the *quality* of business formation. Because low-skilled workers were more likely to become unemployed during the COVID-19 recession, we investigate whether higher UI generosity is associated with lower quality business applications. Specifically, we consider signals about the quality of new businesses available in the BFS data, such as incorporation status, hiring plans, and high likelihood of job creation (estimated by the U.S. Census Bureau). We find suggestive evidence that higher UI generosity led to a deterioration in the quality of new business ideas. This deterioration is evidenced by the relatively higher share of high business applications (HBAs) in COBOL states after March 2020 relative to non-COBOL states. COBOL states were the states that had relatively less UI generosity. HBAs are business applications that are more that are more likely to transition to businesses.

We document supporting evidence that we satisfy the conditional parallel trends assumption. We document that COBOL states and non-COBOL states are statistically indistinguishable along many dimensions that are potentially related to business formation as shown in Table 3.1. There are two characteristics where we find notable differences, which are homeownership rate and Republican governor. COBOL states were less likely to have a Republican governor and less likely to own their homes. To address these concerns we add Census division by month fixed effects. As shown in Table 3.2, once we control for Census division fixed effects, COBOL and non-COBOL states become indistinguishable along many dimensions including homeownership and Republican governor. We also show that COBOL and non-COBOL states experienced similar pandemic-related restrictions, such as stay-at-home requirements and school closures as shown in Table A.1.

Taken together, our analysis suggests that the expansion of UI benefits as an important driver behind the surge in business formation after the pandemic recession. Our results highlight the potential role of UI policy in enhancing the pace of recovery from recessions by fostering entrepreneurship.

One may want to address directly whether UI generosity caused the surge in business formation through an instrumental variable approach. However, this lies outside the scope of this chapter as does directly measuring UI generosity differences between COBOL and non-COBOL states.

Related Literature This chapter lies at the intersection of two strands of literature: the literature on the impact of unemployment benefits on labor market outcomes and the literature on the cyclicality of business formation.

Recently, the literature studying the labor market impact of unemployment insurance has begun to explore the impact of UI on self-employment and entrepreneurship.[8] Several papers in the literature study labor market reforms enacted across Europe. Focusing on a reform in Spain that reduced the replacement rate of long term UI, Camarero Garcia & Hansch (2021) find that the cut in UI reduces the probability, but not quality, of self-employment. Hombert, Schoar, Sraer & Thesmar (2020) show that firm entry rose significantly after France extended UI to unemployed individuals that start a business and that the quality of those new firms did not deteriorate relative to those that entered before the reform. Meanwhile, Gaillard & Kankanamge (2023) focus on the U.S. context, where self-employment is generally not covered by UI, and finds that higher UI generosity lowers the probability of unemployed individuals entering self-employment. Through

[8]Traditionally, the literature on the impact of unemployment benefits on labor market outcomes focuses on re-entry into paid employment. See, for example, Atkinson & Micklewright (1991), Boone, Dube, Goodman & Kaplan (2021), Card, Johnston, Leung, Mas & Pei (2015), Chodorow-Reich, Coglianese & Karabarbounis (2019), Farber, Rothstein & Valletta (2015), Katz & Meyer (1990), Lalive, Ours & Zweimuller (2006), Meyer (1990), Nekoei & Weber (2017), Schmieder, von Wachter & Bender (2012).

the lens of a structural model, calibrated to the U.S. economy, Gaillard & Kankanamge (2023) show that a UI system that covers self-employed individuals fosters business creation. In this chapter, we build on this literature by establishing empirical evidence using a U.S. episode in which much of the increase in UI generosity was in the form of larger benefit amounts rather than extended benefit duration, accompanied by the relaxation of the FFCRA.

There is a growing interest in better understanding drivers of nascent entrepreneurship during normal times and economic downturns.[9] Using the U.S. Census Bureau's Business Formation Statistics (BFS), Dinlersoz, Dunne, Haltiwanger & Penciakova (2021) compare the evolution of business formation during and after the Great Recession versus COVID-19. Fazio, Guzman, Liu & Stern (2021) use data from the Startup Cartography Project to explore the local correlates of state business registrations during the COVID-19 pandemic. Decker & Haltiwanger (2023) use Business Formation Statistics (BFS) and the Bureau of Labor Statistics' (BLS) Business Employment Dynamics (BED) to show that the rise in business formation is observed consistently across various measures of new business entry, and their cross-industry and cross-regional patterns are consistent with pandemic-induced structural shift toward remote work and online shopping. While these papers discuss the surge in business formation during COVID-19, they do not establish causal evidence for its underlying drivers. This chapter advances the literature by identifying UI generosity as a potential contributor to the rise in business formation during the 2020 phase of COVID-19.

[9] For example, see Andrews, Fazio, Guzman, Liu & Stern (2022), Bayard, Dinlersoz, Dunne, Haltiwanger, Miranda & Stevens (2018), Dinlersoz, Dunne, Haltiwanger & Penciakova (2023), and Guzman & Stern (2020).

3.2 Background and Hypothesis

In this section, we document how much UI benefits increased during this period. We hypothesize how this increase in UI benefits could have affected business formation. Finally, we note that this was a period of low work search, which could amplify the effects that we are finding.

3.2.1 Increase in UI Generosity

Two provisions of the CARES Act increased UI generosity. First, the Act established the Federal Pandemic Unemployment Compensation (FPUC) program, which added an extra $600 in weekly federal benefits through the end of July 2020.[10] This increase in UI benefits was substantial. To illustrate, Figure 3.2a shows that the median replacement rate of UI benefits for each state before the pandemic varied between 34 percent and 65 percent. With the FPUC, the median replacement rate ranged from 115 percent to 166 percent as shown in Figure 3.2b. Consistent with the large increases, Ganong et al. (2020a) estimate that 76% of unemployed individuals received more than 100% of their pre-unemployment income as UI benefits due to the FPUC program. The duration of UI benefits was also extended through the Pandemic Emergency Unemployment Compensation (PEUC) program, which provided an additional 13 weeks of UI compensation. Another component of the CARES Act was the Pandemic Unemployment Assistance (PUA) that expanded who was eligible to receive unemployment insurance benefits. Note that these three programs FPUC, PEUC, and PUA contributed $439 billion, $84 billion, and $130 billion respectively worth of UI benefits to be disbursed.[11]

[10]Upon expiration, the FPUC was immediately followed by the Lost Wage Assistance (LWA) program, which provided $300 extra weekly benefits until September 5, 2020.

[11]Note that these expenditures include UI benefits after 2020 given that these programs ended in September 2021. See information from the Pandemic Response Accountability Committee.

The large increase in UI generosity enabled unemployed individuals to increase not only their consumption, but also their savings. For example, Ganong et al. (2022b) use bank account data to show that while unemployed individuals who received UI benefits increased their spending, they also increased their savings relative to their employed counterparts. These increases in both spending and savings were possible due to replacement rates of over 100 percent for the majority of claimants.[12] Consistent with this micro-level evidence, Aladangady et al. (2022) estimate that UI contributed $836 billion in aggregate excess savings through the middle of 2022. Similarly, Cox, Ganong, Noel, Vavra, Wong, Farrell, Greig & Deadman (2020) find that while the massive increase in unemployment was especially concentrated in low-income households, these households contributed disproportionately to the aggregate increase in liquid bank account balances, relative to their pre-pandemic shares.

We hypothesize that the large increase in UI generosity contributed to the rise in business formation through at least two channels. First, it enabled unemployed individuals with potential business ideas to accumulate savings to fund their startups. Indeed, a wide body of research indicates that an increase in savings leads to a higher propensity to enter entrepreneurship, as startup businesses often face financial constraints.[13] The rise in business formation during this period is most pronounced in industries with relatively low fixed costs, such as Nonstore Retailers (NAICS 454) and Professional, Scientific, & Technical Services (NAICS 541), making savings from UI expansion adequate for startup capital.[14] Second, the increase in UI had a significant stimulative effect on consumer spending (Ganong et al. (2022b) and Chapter 2), raising expected

[12] The main estimates in Ganong et al. (2020a) do not include PUA, but Ganong et al. (2020a) show that most PUA recipients had over 100% replacement rates as well. Note that FPUC was an additional $600 per week and this amount applied to regular UI claims and PUA claims.

[13] For example, see Evans & Jovanovic (1989), Buera (2009), Kerr & Nanda (2011), Corradin & Popov (2015), and Schmalz, Sraer & Thesmar (2017) among many others.

[14] Table A.5 shows the top 10 industries with the largest increases in business formation in 2020.

profits of businesses and thus incentivizing new business entry.

3.2.2 Decline in Work Search

Work search from March 2020 to December 2020 was reduced (Faberman, Mueller & Şahin, 2022). There were a myriad of factors that could have led to this reduction such as decreased job vacancies, a pandemic, and the removal of work search requirements. Job vacancies declined as was typical during recessions.[15] This was compounded by a pandemic where some businesses could not open or at least open to the same capacity as before the pandemic. The pandemic led to individuals to potentially not want to search for work if that job required them to interact face-to-face with other people, which would have increased the risk of being infected with COVID-19. Finally, The FFCRA allowed states the flexibility to modify or waive work search requirements.

Even though work search was reduced, an increase in the UI benefit amount could have raised the opportunity cost of starting a business if individuals were required to relinquish their UI benefits entirely upon generating business income. That was not the case, however, as self-employed individuals could continue receiving partial UI benefits. State governments require unemployed workers to report any income earned, including from their businesses or self-employment activities. While specific provisions vary by jurisdiction, workers are allowed to earn a certain amount of money while still receiving their full UI Weekly Benefit Amount (WBA). This is known as "disregarded earnings." State governments subtract the disregarded earnings from weekly income, and then subtract this adjusted earnings from the worker's WBA to calculate their

[15] See the April 2020 Job Openings and Labor Turnover (https://www.bls.gov/news.release/archives/jolts$_0$6092020.pdf)news report.

partial benefit amount.[16] Furthermore, as documented by Bayard et al. (2018) and Dinlersoz et al. (2023) there are long lags from a business application to a business that generates income. Given UI benefit enhancements during this period, partial UI was more generous than in the past.

3.3 Data

3.3.1 Business Formation per capita

Our main outcome of interest is business formation per capita. For the denominator, we use state population annual estimates from the IPUMS National Historical Geographic Information Systems (NHGIS). For the numerator, business formation, we use the U.S. Census Bureau's Business Formation Statistics (BFS). Throughout our analysis, we focus on the monthly estimates at the state level. The BFS program receives data on the universe of new applications for Employer Identification Numbers (EINs) from the IRS on a weekly basis. In the United States, all employer businesses are required to have an EIN for payroll purposes, but individuals may hold an EIN for other reasons, including applications for trusts or estates.

Consequently from the universe of filings, the BFS program constructs a measure of Business Applications (BA) that consists of applications that are likely to be associated with new business formation. This is done by excluding applications for tax liens, estates, trusts, financial filings, as well as applications within certain industries—NAICS 11 (agriculture, forestry, fishing, and hunting), NAICS 92 (public administration), and NAICS 814110 (private households).[17]

While our baseline analysis focuses on the BA series, we also look at three additional

[16] For a full list of these provisions across states, see Comparison of State Unemployment Insurance Laws issued by the Department of Labor.

[17] For full details on the micro data and series constructions, see Bayard et al. (2018).

measures published by the BFS program that allow us to look at business application quality. The baseline BA series includes applications for businesses that are both intended to be employer and non-employer. In fact, only about 10 percent of all BAs transition into employer businesses within two years (Dinlersoz et al. (2023)).

Using information available in the EIN application, the BFS program constructs three additional series on applications that have a higher probability of transitioning into an employer business: business applications with planned wages (WBAs), business applications from corporations (CBAs), and high-propensity business applications (HBAs). WBAs and CBAs are subsets of HBAs, but they are not mutually exclusive. WBAs are defined as BAs that indicate an intent to hire employees and/or pay wages when applying for an EIN. CBAs are defined as BAs that are submitted by a corporation or a personal service corporation. HBAs are defined as BAs that include at least one of the following characteristics: application for a corporate entity, an application indicating that they are hiring employees, provide a first wages-paid date (planned-wages), or have a NAICS industry code in accommodation and food services (72) or in portions of construction (237, 238), manufacturing (312, 321, 322, 332), retail (44, 452), professional, scientific, and technical services (5411, 5413), educational services (6111), and health care care (621, 623).

In addition to data on business applications, the BFS program also publishes data on the transition of applications to employer businesses. It does so by linking EIN applications to employer business births via the Business Register (BR) and Longitudinal Business Database (LBD). Specifically, BFS publishes tabulations of business formations within four and eight quarters of the application date. Because of the time lag associated with the release of LBD, while application (BA, WBA, CBA, and HBA) data are available through March 2024, formation data is only available through the end of 2019 for the eight quarter business formation series and the end of

2020 for the four quarter business formation series.

3.3.2 COBOL Data

Our main source of variation, COBOL, is a binary variable that indicates whether a state used COBOL in its UI benefits system at the onset of the COVID-19 pandemic. To construct this measure, we directly follow Chapter 2. Specifically, the data on COBOL usage were collected primarily from emails, news articles, and information from the UI Information Technology Support Center. There are 28 COBOL states and 22 non-COBOL as of June 2020. COBOL states are slow to modernize, so given that our analysis is restricted to 2019 to 2020, COBOL usage is time invariant. We also rely on the Department of Labor Employment Training Administration (DOLETA) 9050 report similar to Chapter 2 to measure official delays in benefit disbursement between COBOL and non-COBOL states.

Figure A.1 shows that COBOL usage is dispersed throughout the country. In our empirical analysis, we also consider specifications that control for Census division fixed effects to deal with any potential geographic clustering of COBOL states. The reason for doing so is the potential concern that the differences in the growth rate of business applications per capita is being driven by the region, rather than COBOL usage in UI benefit systems.

3.3.3 Additional Variables and Data Sources

As part of our analysis, we evaluate whether COBOL and non-COBOL states are different along a number of demographic, labor, income, political, and other characteristics. From the 2019 American Community Survey (ACS), we obtain data on population, median age, education, race, ethnicity, poverty, employment to population ratio, labor force participation, self employment, homeownership, home value, and employment shares by occupation. We also obtain income per capita in 2019 from the Bureau of Economic Analysis (BEA), urban population share in 2010 from the U.S. Census Bureau, Republican governorship in 2018 from the National Conference for State Legislatures (NCSL), Republican vote share in the 2016 elections from MIT Election Data and Science Lab, union membership are in 2018 from the Bureau of Labor Statistics (BLS), and employment share by firm age, size, and sector from the U.S. Census Bureau's Business Dynamic Statistics (BDS). We calculate unemployment risk exposure using characteristics of those who became unemployed during April-July 2020 obtained following Ganong et al. (2020a) based on the 2020 CPS Merged Outgoing Rotation Group (MORG) Earnings data and local demographic characteristics in 2019 from the 2019 ACS. We also consider pandemic-related restrictions in place between March 13 and December 31,2020, measured by the Oxford COVID-19 Government Response Tracker (OxCGRT).

In our regression analysis, we also include Census division fixed effects. The U.S. Census Bureau defines nine divisions in the United States with divisions including between three and eight states.[18]

3.4 Identification Strategy

We rely on the conditional parallel trends assumption for identification. The threat to identification is whether there are systematic differences between COBOL and non-COBOL states and those differences affect business formation from March 2020 to December 2020 even after

[18]The South Atlantic division would include nine states if we were to treat Washington D.C. as a state. Note that our analysis similar to Chapter 2 in that we also exclude Washington D.C.

conditioning on our control variables. We take a conservative approach and look at a broad set of confounders and only find two that are statistically significant differences between COBOL and non-COBOL states: Republican Governor and homeownership rate. Once we control for Census division fixed effects, we find no confounders that are statistically significant.

Prior to discussing confounders, we will discuss COBOL usage in UI benefit systems and how its usage could have affected business formation. The relatively longer delays in the disbursement of UI benefits in COBOL states could have hampered business formation. During this period, states struggled, to varying degrees, to manage the simultaneous surge in UI claims and changes to benefits amounts and eligibility criteria. As a result, the timeliness of UI processing significantly declined. Figure 3.3a illustrates that, on average, the share of first UI payments delayed by 70 days or more increased from close to zero percent in March 2020 to 25 percent by October 2020.[19]

COBOL was first introduced in 1959. It was once the standard programming language for all UI systems. By 2020, nearly half of states (22) had modernized their UI systems and phased out COBOL. In the remaining 28 states that continued using COBOL, UI systems became particularly overwhelmed at the start of the COVID-19 pandemic.[20] As illustrated in Figure 3.3b, the share of first UI payments delayed by 70 or more days was close to zero percent in both COBOL and non-COBOL states prior to COVID-19, but by October 2020 the two groups of states diverged significantly, with the share hovering around 19 percent for non-COBOL states and 27 percent for COBOL states.

Several factors contributed to the overloading of COBOL-based UI systems. First, although

[19]Note that the values for October 2020 are associated with claims filed in August 2020, July 2020, or earlier (for claims delayed well over 70 days).

[20]Figure A.1 shows the map of states by COBOL usage status.

COBOL is capable of performing the same tasks as more modern programming languages, adapting COBOL systems to accommodate changes in eligibility and benefits amounts is more challenging. Second, implementing the necessary changes required some states to hire additional COBOL programmers, which was difficult because many COBOL programmers had retired. Third, and more generally relevant for any antiquated UI system, states using COBOL also had less user-friendly platforms through which unemployed individuals could file UI claims.[21] As shown in Chapter 2, these factors contributed to COBOL states facing relatively more severe UI processing and benefits disbursement delays. As an anecdotal example, the situation was so dire in New Jersey (a COBOL state) that in April 2020, Governor Phil Murphy pleaded for assistance from any programmers who knew how to program in COBOL.[22]

One concern is that COBOL usage in UI benefit systems could be correlated with a confounder where this confounder is driving the business formation differences between COBOL and non-COBOL states. Even though states have direct control over COBOL-usage, modernizing away from COBOL is a slow and expensive process. The decision to modernize could systematically differ between states, which would be a concern if these systematic differences were also correlated with factors associated with the rise in business formation during the pandemic. One specific concern, is whether states who modernized COBOL were the states with more generous UI systems. However, that is not the case. [23]

To address the concern that a confounder is driving our results, we compare COBOL and non-COBOL states along several dimensions, including demographics, household characteris-

[21]See Chapter 2 for a thorough discussion of the challenges faced by COBOL states.

[22]See Feldman, B. (2020, April 6). NJ Governor Requests Expertise of 6 People Who Still Know COBOL. *New York Magazine*.

[23]As documented in Chapter 2, the decision does not go in the expected manner. The states with more repressive UI regimes were more slightly more likely to modernize away from COBOL.

tics, political environment, and labor market characteristics. Table 3.1 displays the results from regressing various state-level demographic, household, and political characteristics on a COBOL indicator.[24] We find that COBOL and non-COBOL states are statistically indistinguishable along a variety of dimensions. Two exceptions are whether the state's governor is a Republican (measured in 2020) and the homeownership rate. Differences in the party affiliation of the state governor could potentially matter because people's political leanings were correlated with the degrees of voluntary and involuntary social distancing, which in turn may have affected economic activity and the pace of business formation.

We view these two dimensions as potential threats to the conditional parallel trends assumption. The political differences could violate our conditional parallel post-trends assumption. Even though COBOL is not directly affecting political affiliations nor COVID-19 cautiousness, it would be correlated with these factors that could be affecting business formation. Potentially business formation was higher in non-COBOL states from March 2020 to July 2020 because Republican governors had policies more friendly to starting a business. Another potential concern is that increased fear of COVID-19 in COBOL states hampered the idea generation process of creating a business. There were other policies that happened at the same time such as mortgage payment forbearance, which was part of the CARES Act.[25] However, there is no systematic differences in percent of households with mortgages as shown in Table 3.1, but there is a statistically significant difference in homeownership rate. Differences in the homeownership rate may matter because of the documented importance of housing as collateral for young and small businesses (Kerr, Kerr & Nanda, 2022; Lastrapes, Schmutte & Watson, 2022; Davis & Haltiwanger,

[24] The characteristics chosen here are not identical to Chapter 2 because confounding threats to business formation are different than confounding threats to consumption.

[25] In Chapter 2, I discuss why some other pandemic policies are not problematic such as the Supplemental Nutrition Assistance Program, Paycheck Protection Plan, and Economic Impact Payments.

forthcoming). Furthermore, there was an appreciation of home prices during this period, so the differences seen in Table 3.1 are potentially problematic.

Motivated by the fact that there is strong regional clustering in people's political leanings and in homeownership rates, we control for Census division fixed effects in Table 3.2, and find that the differences in state governor's party affiliation and homeownership rates become statistically insignificant.[26] Moreover, all other demographic, household, and political differences also remain insignificant. As an additional test, in Table A.1 we use the Oxford COVID-19 Government Response Tracker (Hale, Angrist, Goldszmidt, Kira, Petherick, Phillips, Webster, Cameron-Blake, Hallas, Majumdar et al., 2021) to directly compare the pandemic-related restrictions imposed by state governments across COBOL and non-COBOL states, and find no significant differences. We view the usage of Census division fixed effects as a cleaner method to isolate COBOL variation given that any differences between COBOL and non-COBOL states disappear.

Because there were strong systematic differences in the adverse impact from the COVID-19 pandemic across different industries and occupations (Adams-Prassl, Boneva, Golin & Rauh, 2020)[27], in Table A.2 and Table A.3 we compare firm and labor market characteristics of COBOL and non-COBOL states. We find no statistically significant differences in employment shares by firm age, firm size, sectors, or occupations. Collectively, these results support that COBOL and non-COBOL states were not observably different along many dimensions that might have affected business formation after taking Census division by month fixed effects, which assuages concerns of violating the conditional parallel trends assumption.

[26] Controlling for Census division fixed effects also controls for any region-specific factors that could affect business formation.

[27] In Chapter 2, I find differences in relative time spent at home using Google Mobility measures, but these differences disappear once I take Census division fixed effects into account.

3.5 Main Results

We first document the relationship between COBOL usage and business formation in the aftermath of the emergency declaration in March 2020. Then as a secondary exercise to try to answer the question of whether the initial quality of businesses was affected by the lower increase in UI generosity in COBOL states, we look at the share of higher quality type businesses in COBOL states relative to non-COBOL states.

3.5.1 Quantity of Business Applications

We estimate the effect of COBOL-usage on business formation with a monthly two-way fixed effects (TWFE) event study:

$$ln(BApc)_{s,t} = \sum_{\tau=-T_0}^{T_1} \beta_\tau (COBOL_s \times I_\tau) + \delta_s + \lambda_t + \epsilon_{s,t} \tag{3.1}$$

where t denotes the number of months since March 2020. $ln(BApc)_{s,t}$ is the log of BFS business applications (BA) per capita (pc) in state s in month t. $COBOL_s$ indicates whether state s uses COBOL, and I_τ indicates whether the corresponding month is month τ. δ_s and λ_t are the state and time fixed effects, respectively. $\epsilon_{s,t}$ is the error term. This TWFE event study design allows us to track the monthly evolution of business formation in COBOL states relative to non-COBOL states.

Figure 3.4a shows evidence of parallel pre-trends and that COBOL states experienced a slower pace of business formation following the pandemic recession. The figure presents the estimated β_τ with 95 percent confidence intervals from the TWFE model. First, we see parallel pre-trends in business formation between COBOL and non-COBOL states in 2019, lending support to the assumption required for a difference-in-differences design. Second, from March 2020 to December 2020, COBOL states experienced a slower growth rate in per capita business applications compared to non-COBOL states. Notably, we observe a V-shaped trend: the difference in business formation between the two groups of states gradually widens from March 2020 to June 2020. In July 2020, coinciding with the expiration of the FPUC, the gap reaches its peak at 15 percent, and subsequently, from August to November 2020, this difference progressively narrows.

The V-shaped pattern of the estimated effects is consistent with the our hypothesis in several ways. First, as shown in Figure 3.3b, the difference in the share of initial claims delayed by 70 days or more between COBOL and non-COBOL states gradually widens until October 2020, indicating that the difference in the proportion of eligible unemployed workers who did not receive their UI payment increases over this period.[28] In turn, the adverse impact on business formation was likely to have increased over this period as well. Second, we claim that a channel through which the UI expansion affects business formation is the increase in savings. Given that UI payments occur on a weekly or biweekly basis, it is likely to take time for the unemployed workers to accumulate enough savings for their business projects, and hence, the impact on business formation is also likely to increase gradually.

Given the documented differences in Table 3.1 between COBOL and non-COBOL states in the party affiliation of the governors and homeownership rates, there could be concerns that these omitted variables have differential, time-varying effects on business formation in the two groups of states in 2020. To address this concern, we augment our TWFEs model with Census

[28] Note that topcoded claims from October 2020 are most likely from July 2020 or earlier.

division by month fixed effects.[29]

Figure 3.4b confirms parallel pre-trends, the slower pace of business formation in COBOL states, and the V-shaped pattern of the estimated effects. Further, in Figure A.2, we adopt a doubly robust estimator developed by Callaway & Sant'Anna (2021) which allows us to directly control for time varying effects of ex-ante observables in an event study setting. Specifically, we control for both party affiliation of governor and homeownership rate. Similar to Figure 3.4a and Figure 3.4b, we find parallel pre trends as well as the V-shaped pattern with the trough in July 2020.

The standard errors are large in Figure 3.4 in part because we are looking at dynamic effects. However, if we estimate a static TWFE as described in equation (3.2), then the standard errors are reduced in this pooled sample.

$$Y_{s,t} = \beta_0 + \beta_1 Post_t * COBOL_s + \delta_s + \lambda_t + \varepsilon_{s,t} \tag{3.2}$$

States are denoted by s and month by t. The dependent variable $Y_{s,t}$ is the log of BFS business applications (BA) per capita (pc) in state s in month t. $COBOL_s$ indicates whether state s uses COBOL, and $Post_t$ indicates whether the corresponding month is after March 2020. δ_s and λ_t are the state and time fixed effects, respectively. $\varepsilon_{s,t}$ is the error term.

Table 3.3 corresponds to our results from equation (3.2). In columns 1 and 3, we end our sample in December 2020 to be consistent with Figure 3.4. In columns 2 and 4, we end our sample in July 2020 given that July 2020 coincides with the end of the FPUC program and we will using the growth rate UI payments per unemployed in our instrumented difference-in-differences

[29]Recall that after controlling for Census division fixed effects, the differences between COBOL and non-COBOL states in the party affiliation of the governor and homeownership rates become insignificant.

from March 2019-July 2019 to March 2020-July 2020. In column 1 when we only include state and month fixed effects, we estimate that COBOL usage in UI benefit systems led to a 7.1 percent lower increase in business formation per capita in COBOL states relative to non-COBOL states. By historical standards, a differential increase of 7 percent in business applications would be a large amount. However, one should note that in July 2020 business applications per capita were over 80 percent above pre-pandemic trends and after 2021 remained over 40 percent above pre-pandemic trends.

Our results in Table 3.3 are consistent across specifications. The odd numbered columns of the Table report results through July 2020; the even numbered columns report results through December 2020. In column 2, we estimate that the increase in business formation in COBOL states was 7.7 percent lower than in non-COBOL states after March 2020. The point estimates in ending in July 2020 versus December 2020 are not statistically different. In column 3, we use Census division by month fixed effects instead of month fixed effects. In column 3, we estimate that the increase in business formation in COBOL states was 6.6 percentage points lower than in non-COBOL states. In column 4, we estimate that the increase in business formation in COBOL states was 6.6 percent lower than in COBOL states relative to non-COBOL states. Our results from columns 1 and 2 are significant at the 10 percent level, while our results from columns 3 and 4 are significant at the 5 percent level.[30]

[30] July 2020 is an outlier (expiration of FPUC) as was seen from the event study difference-in-differences. However, we estimate that the increase in business formation in COBOL states was 5.6 percentage points lower than in non-COBOL states from March 2020 to December 2020 when we exclude July 2020. This effect is significant at the 5 percent level.

3.5.2 Quality of Business Applications

All our results thus far have focused on business applications, which is a necessary first step for entrepreneurs to start a business. However, one concern is that these business applications may be a weak signal of business formation particularly when it became relatively cheaper for potential entrepreneurs to file a business application. The opportunity cost to file a business application could be lower for a host of reasons during this period. Potential entrepreneurs might have faced a lower trade-off during this period between idea generation and leisure because of the increased social isolation that COVID-19 brought, particularly prior to vaccines becoming available. Stay at home orders forced individuals to stay at home, which could have led to potential entrepreneurs to allocate more time to idea generation relative to leisure. Households might be more likely to start a business if they are prevented from partaking in outdoor activities and socializing in-person with others; leisure potentially became less enjoyable for potential entrepreneurs. Finally, as discussed the relaxation of work search requirements could have also lowered the cost to apply to start a business, since potential entrepreneurs could allocate more time on starting a business rather than looking for a new employer.

In order for the quality of business applications to have deteriorated we look at whether the share of business applications from a corporation (CBA), business applications with planned wages (WBA), and high propensity business applications (HBA). All of these types of businesses have higher transition rates from business application to business. For example, the transition rate from BA to business is 10%, while the transition rate from WBA to business is 40% (Dinlersoz et al., 2023).

We use data from the BFS on the number of business applications that (i) are incorporated

(CBA), (ii) submitted actual hiring plans (WBA), and (iii) have high likelihood of becoming actual employer startups (HBA), as estimated by the Census Bureau, to proxy for the quality of business ideas. Specifically, we calculate the share of business applications with these characteristics, and use them as the outcome variable to estimate equation (3.2) with state and division by month fixed effects. The share of HBAs, CBAs, and WBAs declined after March 2020 in both COBOL and non-COBOL states. This is consistent with our hypothesis of additional UI funds leading to potentially lower quality entrants.

In Table 3.4, we estimate the effect of antiquated UI benefit systems on the initial quality of business formation. In all three specifications we use state fixed effects, Census division by month fixed effects, and end our sample in July 2020.[31] We find no difference in the share of CBA (column 1) and WBA (column 2) after March 2020 between COBOL and non-COBOL states. Given this lack of significance, we focus our business quality analysis on the share of HBAs. In column 3, when we look at the share of HBAs, we estimate that the decrease in HBAs in COBOL states was 0.8 percentage points lower than in non-COBOL states from March 2020 to July 2020. This implies there was a relative increase in business quality as measured by HBAs in COBOL states. Alternatively, we estimate that the decrease in HBAs in non-COBOL states was 0.8 percentage points higher than in COBOL states from March 2020 to July 2020. This alternative interpretation is equivalent to the previous one. However, this alternative interpretation emphasizes that the states, non-COBOL states, that experienced a larger increase in UI payments also experienced experienced a relative decline in the share of BAs that were HBAs; business quality deteriorated in the states (non-COBOL states) that received more UI generosity even

[31] We end in July 2020 given that our preferred specification for business formation ends in July 2020. Our business quality results are robust to ending in December 2020.

though this increased UI generosity led to relatively more business applications.

Taken together these results imply that the business applications formed during this period were less likely to transition into a business generating income. This could have been in part due to lower quality ideas being generated or the composition of businesses applications shifting to industries that typically lead to lower transition rates.

3.6 Conclusion

A number of studies have documented and discussed the surge in business formation during the COVID-19 pandemic (Dinlersoz et al., 2021; Fazio et al., 2021; Decker & Haltiwanger, 2023). In this chapter, we establish a causal link between antiquated UI benefit systems during the 2020-phase of COVID-19 and the rise in business applications. Using a TWFE design, we estimate that the increase in business formation in COBOL states was 6.6 percentage points lower than in non-COBOL states between March 2020 and July 2020. We also estimate a deterioration of initial business quality in COBOL states relative to non-COBOL states.

As already documented by Ganong et al. (2022b) and Chapter 2, unemployment insurance played an important role in the sharp economic recovery from the pandemic recession by spurring consumption. In this chapter, we document a new channel through which unemployment insurance further helped stabilize local labor markets. Unlike its more immediate effect on consumption, the impact of UI expansion on local labor markets through business formation may take some time to fully materialize. This is due to the fact that it takes several quarters for new businesses to hire employees and grow large enough to have a quantitatively meaningful impact on the labor market.[32]

[32]See Bayard et al. (2018) and Dinlersoz et al. (2023) for documentation of the lag between the filing of business

Our analysis has focused on the necessary first step of creating a business: filing a business application. However, we are interested if these business applications transitioned into businesses generating income. Given the long lags between filing a business application and being an employer based business, we would not expect these business applications to transition until 2022. In September 2024, the BDS will release data on firm entry in 2022. Once these data are released, we will incorporate analysis on the transition rate from business applications filed during 2020. Furthermore, we will add a simple two-period model to illustrate how changes in employment benefits can affect the probability of entering entrepreneurship.

It is important to note that the UI expansion in this study was implemented in a period when pandemic-induced structural transformation, such as increases in remote work and online shopping, potentially created new business opportunities (Decker & Haltiwanger, 2023). Therefore, it is possible that the effects of UI expansion are amplified during this period, and caution is needed in extrapolating our results to other settings. However, to the best of our knowledge, this study is the first to document that UI expansion fosters business formation (particularly in the benefit amount dimension), and it opens the door for more discussion on the potential role of UI policy in promoting entrepreneurship during economic recoveries.

Furthermore, understanding what was one of the key drivers of the surge in business formation in 2020 is relevant for understanding economic conditions well after 2020. Specifically, the U.S. economy has exhibited strong consumer spending in 2022 and 2023, despite the conclusion of several pandemic transfer programs such as UI. Initially, excess savings was used as an explanation for the robust consumer spending. However, as this trend persisted, the credibility of excess savings as an explanation diminished. It remains a puzzle how the U.S. economy has

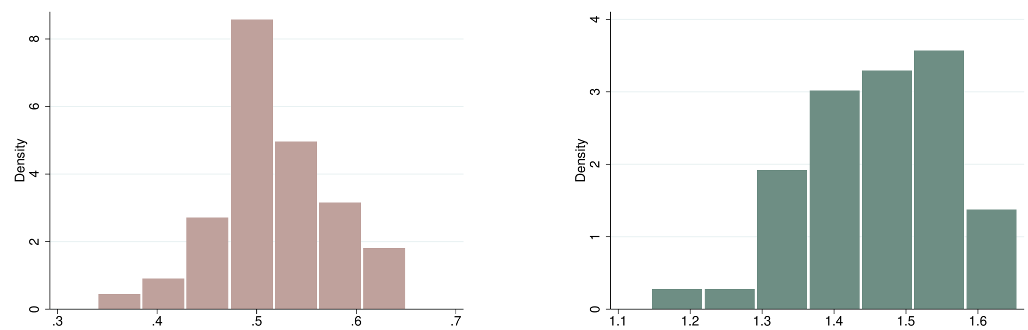
applications and their transition to employer businesses.

sustained strong consumer demand in the face of prolonged monetary policy tightening. One potential explanation is the increase in business dynamism that had been recently lacking from the U.S. economy. Specifically, given the long lags from a business application to a business generating income, we would expect a lag from when the surge in business applications translates into additional jobs that generate income.

Figure 3.1: Business Formation per Capita

Notes: Seasonally-adjusted monthly business applications per 1,000 people relative to its 2019 average (0.89 business applications per 1,000 people). Business application data is obtained from the U.S. Census Bureau's Business Formation Statistics.

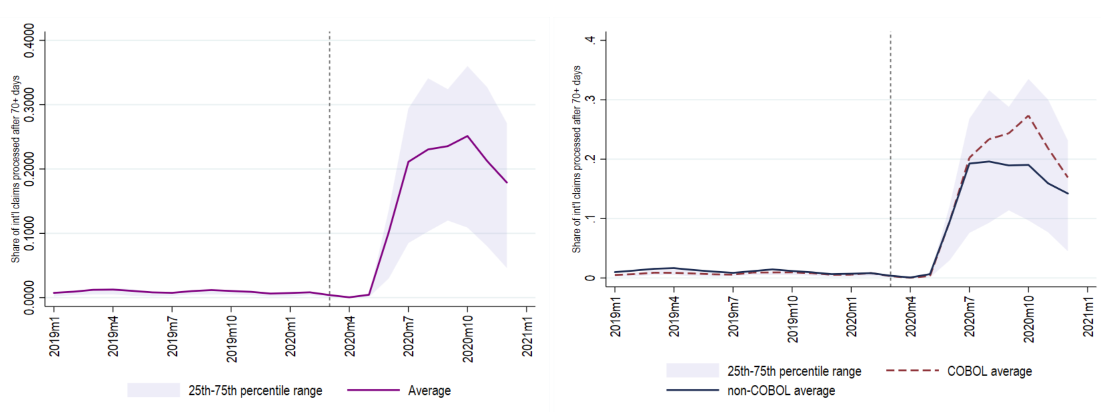
Figure 3.2: Distribution of State-Level Median Replacement Rates



(a) Replacement Rate without FPUC (b) Replacement Rate with FPUC

Notes: These figures depict the distribution of state-level median replacement rates before the passage of the CARES Act (without FPUC) and after (with FPUC). Replacement rates are calculated using the methodology of Ganong et al. (2020a) using 2020 ASEC data on earnings.

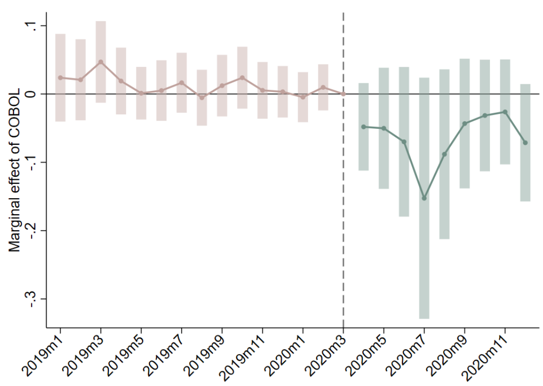
Figure 3.3: Share of First UI Payments Made After 70+ Days



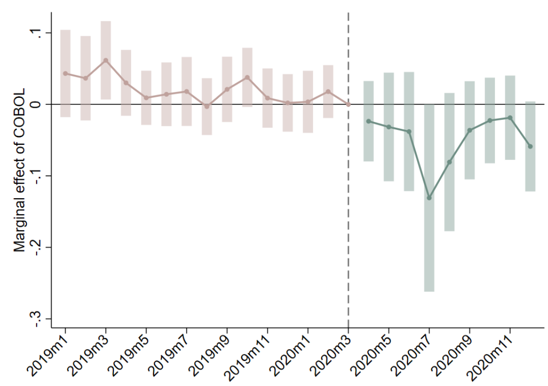
(a) Average (b) COBOL vs. non-COBOL

Notes: These figures depict the cross-state average (solid line) and 25th-75th percentile range of (a) the share of initial claims processed after 70+ days at the national level and (b) the share of initial claims processed after 70+ days in COBOL and non-COBOL states. Data are obtained from the "Benefits: Timeliness and Quality Reports" released by the BLS.

Figure 3.4: Marginal Effect of COBOL: TWFE



(a) TWFE



(b) TWFE with Division × Month FE

Notes: These figures depict the marginal effect of COBOL on log(business applications per capita). Figure (a) shows the two-way fixed effects results where month and state fixed effects are included. Figure (b) shows the two-way fixed effects where state fixed effects and month by division fixed effects. Standard errors are clustered at the state level in both figures.

Table 3.1: Balance of Characteristics

Dependent Variable	Coefficient on COBOL indicator: Est.	Std Err.	P-Value	Mean of Dep Var.
Demographics				
Log population	-0.096	(0.295)	0.75	15.21
Median age	-0.123	(0.706)	0.86	38.34
High school or lower	-0.001	(0.013)	0.91	0.41
Some college	-0.013	(0.009)	0.15	0.30
Bachelor's degree or higher	0.015	(0.014)	0.30	0.28
% White	-0.023	(0.036)	0.54	0.77
% Black	-0.024	(0.027)	0.37	0.10
% Hispanic	0.007	(0.028)	0.82	0.11
% Foreign born	0.023	(0.022)	0.30	0.12
Labor and Income				
Income per capita ($1,000)	2.812	(2.213)	0.21	53.45
% Below poverty	-0.009	(0.008)	0.29	0.13
Employment to population	0.010	(0.012)	0.40	0.60
Labor force participation rate	0.009	(0.011)	0.41	0.64
Self employment rate	0.002	(0.004)	0.64	0.07
Unemployment risk exposure	-0.002	(0.009)	0.82	0.08
Residential				
% Urban population	0.015	(0.042)	0.73	73.4
Homeownership rate	-0.029**	(0.011)	0.01	0.62
% Households w/ mortgage	-0.015	(0.012)	0.19	0.38
Median home value ($1,000)	454.56	(394.332)	0.26	617.96
Political Environment				
Republican governor	-0.283**	(0.132)	0.04	0.66
Republican vote share (2016)	-0.025	(0.029)	0.40	0.49
Union membership rate (2018)	0.019	(0.014)	0.21	0.10

Notes: This table reports results from regressions where each one of the state-level characteristics in Column (1) are dependent variables and the COBOL indicator is the independent variable. Variables under Demographics, Labor and Income, and Residential categories, except for income per capita and unemployment risk exposure, are obtained from the 2019 American Community Survey. We obtain income per capita in 2019 for each state from the Bureau of Economic Analysis. We calculate unemployment risk exposure using characteristics of those who became unemployed during April-July 2020 and local demographic characteristics in 2019. Percent urban is measured as of 2010 and is obtained from the U.S. Census Bureau, Republican governor share is measured as of 2018 and is obtained from the National Conference for State Legislatures, union membership is measured as of 2018 and is obtained from the Bureau of Labor Statistics. ***, **, and * indicate significance at the 1%, 5%, and 10% significance levels.

Table 3.2: Balance of Characteristics w/ Division Fixed Effects

	Coefficient on COBOL Indicator			
Dependent Variable	**Est.**	**Std Err.**	**P-Value**	**Mean of Dep Var.**
Demographics				
Log population	-0.339	(0.280)	(0.23)	15.21
Median age	-0.163	(0.587)	(0.78)	38.34
High school or lower	0.001	(0.011)	(0.93)	0.41
Some college	-0.006	(0.007)	(0.33)	0.30
Bachelor's degree or higher	0.005	(0.012)	(0.65)	0.28
% White	-0.003	(0.028)	(0.92)	0.77
% Black	0.014	(0.034)	(0.68)	0.10
% Hispanic	-0.032	(0.019)	(0.10)	0.11
% Foreign born	-0.005	(0.021)	(0.82)	0.12
Labor and Income				
Income per capita ($1,000)	0.561	(1.927)	(0.77)	53.45
% Below poverty	-0.005	(0.007)	(0.43)	0.13
Employment to population	0.009	(0.010)	(0.36)	0.60
Labor force participation rate	0.008	(0.009)	(0.42)	0.64
Self employment rate	0.003	(0.004)	(0.47)	0.07
Unemployment risk exposure	-0.007	(0.009)	(0.43)	0.08
Residential				
% Urban population	-0.019	(0.043)	(0.67)	73.4
Homeownership rate	-0.018	(0.012)	(0.13)	0.62
% Households w/ mortgage	-0.010	(0.011)	(0.38)	0.41
Median home value ($1,000)	45.012	(388.741)	(0.91)	617.96
Political Environment				
Republican governor	-0.101	(0.122)	(0.41)	0.66
Republican vote share (2016)	-0.003	(0.024)	(0.90)	0.10
Union membership rate (2018)	0.006	(0.010)	(0.55)	0.49

Notes: This table reports results from regressions where each one of the state-level characteristics in Column (1) are dependent variables and the COBOL indicator is the independent variable. Variables under Demographics, Labor and Income, and Residential categories, except for income per capita and unemployment risk exposure, are obtained from the 2019 American Community Survey. We obtain income per capita in 2019 for each state from the Bureau of Economic Analysis. We calculate unemployment risk exposure using characteristics of those who became unemployed during April-July 2020 and local demographic characteristics in 2019. Percent urban is measured as of 2010 and is obtained from the U.S. Census Bureau, Republican governor share is measured as of 2018 and is obtained from the National Conference for State Legislatures, union membership is measured as of 2018 and is obtained from the Bureau of Labor Statistics. ***, **, and * indicate significance at the 1%, 5%, and 10% significance levels.

Table 3.3: Effect of Antiquated UI Benefit System on Business Formation

	(1) ln(BA/Pop)	(2) ln(BA/Pop)	(3) ln(BA/Pop)	(4) ln(BA/Pop)
$COBOL_s \times POST_t$	-0.071*	-0.077*	-0.066**	-0.066**
	(0.040)	(0.040)	(0.029)	(0.030)
State FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	No	No
Division x Month FE	No	No	Yes	Yes
Last Month	Dec. 2020	Jul. 2020	Dec. 2020	Jul. 2020
Obs.	1200	950	1200	950
R-sq	0.94	0.94	0.97	0.97

Notes:
This table depicts the marginal effect of COBOL on log(business applications per capita) from a static TWFE model as described in Equation (3.2). The post period starts in March 2020 across all specifications. In columns 1 and 3, we end our sample in December 2020. In columns 2 and 4, we end our sample in July 2020 when the FPUC program ends and we would expect our effects to start to dissipate after July 2020. Standard errors are clustered at the state level. ***, **, and * indicate sig. at the 1%, 5%, and 10% sig. levels. Standard errors are clustered at the state level.

Table 3.4: Initial Quality of New Businesses Formation

	(1) % CBA	(2) % WBA	(3) % HBA
$COBOL_s \times POST_t$	0.005	0.002	0.008**
	(0.004)	(0.002)	(0.003)
State FE	Yes	Yes	Yes
Division x Month FE	Yes	Yes	Yes
Mean Dep. Var	0.11	0.14	0.33
Obs.	950	950	950
R-sq	0.95	0.88	0.87

Notes: Unit of analysis is state by month between January 2019 and July 2020, with the post period beginning in March 2020. We use three additional series on applications that have a higher probability of transitioning into an employer business: business applications with planned wages (WBAs), business applications from corporations (CBAs), and high-propensity business applications (HBAs). The dependent variable in column 1 is % CBA, column 2 is % WBA, and column 3 is % HBA. All three dependent variables are shares of the respective type of business application over the core business application series from the BFS. ***, **, and * indicate sig. at the 1%, 5%, and 10% sig. levels. Standard errors are clustered at the state level.

Chapter 4: Measuring Real Output and Inflation: Official Statistics vs Economics Transactions Data

with Gabriel Ehrlich, John Haltiwanger, David Johnson, Ron Jarmin, Seula Kim, Jake Kramer, Edward Olivares, R. Benjamin Rodriguez, and Matthew D. Shapiro

4.1 Introduction

Current official measures of real output in the United States rely heavily on the integration of survey data across multiple government agencies including the Bureau of Economic Analysis (BEA), the U.S. Census Bureau, and the Bureau of Labor Statistics (BLS). In this chapter, we compare the official statistics on inflation, nominal sales, and real output from the Personal Consumption Expenditures (PCE) with counterparts constructed using item-level transactions data (referred to as Point of Sales or POS data hereafter) and examine the disparities between the resulting measures.

To start with, we document the current source information and methodology used for key components of PCE, specifically focusing on Food and Beverages, as well as Consumer Technology goods such as televisions and computers. First, the source of official statistics originates from various surveys conducted by multiple agencies, involving a complex integration process. Second, some of the source data required for the official statistics (e.g., expenditure weights) are only available with a substantial lag or rely on survey data that may have declining response rates.

For instance, the PCE data incorporate price data from the BLS and nominal sales data from the Census Bureau. At a disaggregated level, the PCE price indices are drawn directly from the

BLS CPI, which is based upon a Laspeyres price index drawn from monthly collection of price relatives of selected items in each product group (i.e., entry-level item level or ELI). Price indices at the ELI and at more aggregated levels are based on expenditure weights from the Consumer Expenditure Survey (CEX), with weights that may reflect substantially dated expenditure patterns in real time. At a more aggregate level (e.g., Food and Beverages), the PCE price index is a Fisher index using expenditure weights from the PCE expenditure data. The PCE expenditure data in turn largely stems from collection of nominal sales from the Monthly Retail Trade Survey (MRTS) by the Census Bureau with the challenge of converting industry statistics to product group statistics.[1] Also, the Economic Census is conducted only every five years, so it takes a number of years to be incorporated into the PCE.[2] This summary of the integration process of survey data into official statistics highlights the inherent challenges of measuring inflation and real output within the current system. Many of these challenges arise from the disparate and complex measurement of the different components involved.

In contrast, the POS data offer internally consistent prices and nominal sales data from an integrated underlying source in a timely manner. They therefore allow for a closer examination and potential rectification of several challenges present in the current measurement system for key national indicators. The POS data not only permit measurement internally consistent prices, nominal sales and expenditure shares but also permit real-time accounting for substitution bias (by using superlative rather than Laspeyres-style price indices) and quality change from product turnover (by using hedonic adjustment). Hedonic approaches to adjusting for quality change make use of item-level product attribute data that are increasingly available for POS data (Ehrlich et al., 2023).

Next, we demonstrate substantial differences between the patterns in the official statistics and those that emerge from the POS data. We find that these differences vary across product groups and time periods. Comparing price indices for disaggregated goods in the Food and Bev-

[1]The MRTS collects data at the industry level (e.g. grocery stores, general merchandise stores) and these sales are allocated to product groups (e.g., milk, coffee, cereal) through an intricate process that originates with product by industry information from the Economic Census.

[2]For example, the 2017 Economic Census was not adopted as the benchmark for the PCE until September 2023.

erages category (consumed at home), we find that the PCE- and POS-based Laspeyres indices track each other reasonably well. However, we uncover substantial substitution bias for overall Food and Beverages, evidenced by notable differences between inflation measured by Laspeyres indices and superlative indices such as the Tornqvist. Furthermore, we identify non-trivial evidence of quality change within the overall food and beverages category.[3]

Examining nominal and real sales, we observe significant disparities in patterns between the PCE and POS data at the disaggregated level. Our analysis reveals that these disparities arise predominantly from the official statistics allocating nominal sales at the industry level to product groups using outdated shares obtained from the Economic Census. In addition to tracking nominal shares more accurately, the adjustments for substitution bias as well as quality change lead to substantial differences in real output patterns at the disaggregated level for PCE categories between PCE and POS data. However, some of these differences are mitigated at the more aggregate level such as overall Food and Beverages consumed at home, for which the PCE and POS nominal sales patterns track each other reasonably well.[4]

Furthermore, we look into these patterns for consumer electronics and find that the disparities between the patterns observed in the PCE and POS data are even more pronounced. The substitution bias and quality change adjustments are substantially larger for consumer electronic items such as notebook computers. Moreover, the limitation of the PCE in capturing high frequency nominal sales patterns is, if anything, more problematic for consumer technology goods than Food and Beverage goods. The larger discrepancies seen in Consumer Tech relative to Food and Beverages could be due to survey methodology being better equipped to measure nondurable goods that are frequently bought than to measure durable goods that are infrequently bought and highly sensitive to temporary discounts.

[3] Our analysis of quality change for food draws heavily on Ehrlich et al. (2023). The value added of using the quality-adjusted prices from this earlier work in the current analysis is that we explore the implications for real output patterns by integrating with the nominal sales data.

[4] We also find that there are notable differences between the PCE price index at aggregate levels and the C-CPI-U, a chained Tornqvist version of the CPI, which is released with a significant lag. These differences reflect the distinct underlying source data for the expenditure weights used in the two series: the PCE price index uses PCE-based expenditure weights, while the C-CPI-U uses CEX based weights).

Lastly, we document striking differences between the official statistics and POS data during the COVID-19 pandemic and its aftermath.[5] The POS data displays lower inflation in 2020 than PCE and CPI. Both the POS data displays sharp increases in inflation in 2021 and 2022 as do the PCE and CPI. Yet the magnitude of the increases in 2021 and 2022 is notably larger in the POS data compared to the PCE for almost all of the selected product groups that we analyze. In addition, we find evidence that the slowdown in inflation in 2023 is more pronounced in the POS data than in the PCE. The PCE's difficulty in tracking high-frequency nominal sales variation also remains evident over this time period. Consequently, real output patterns exhibit notable differences between the PCE and POS for these detailed product groups in the aftermath of the pandemic.

The chapter proceeds as follows. Section 4.2 describes the source data used in the analysis, Section 4.3 delineates the measurement methodology employed by the PCE and for the POS data, Section 4.4 presents our main results comparing the PCE and POS indices, Section 4.5 discusses remaining limitations of the analysis, and Section 4.6 concludes.

4.2 Data

4.2.1 NielsenIQ Data

Our analysis for Food and Beverages is based on the NielsenIQ Retail Scanner dataset (referred to as RMS) provided by the Kilts Center at the University of Chicago Booth School of Business. The data consists of weekly pricing, volume, and store merchandising conditions generated by more than 100 retail chains across all U.S. markets, which includes over 40,000 individual stores. Total sales in the NielsenIQ RMS are worth over \$200 billion per year and represent 50% of total sales in grocery stores, 55% in drug stores, 32% in mass merchandisers, and 2% in convenience stores.

A key advantage of this data is that it contains detailed information at the finest product

[5] Our POS data during this period is limited to a relatively small number of detailed food and beverage product groups (milk, eggs, cereal, sodas, bakery goods and coffee).

level, 12-digit universal product codes (UPCs), that uniquely identifies specific goods. The data consists of over 2.6 million UPCs including food and nonfood grocery items, health and beauty aids, and general merchandise. NielsenIQ classifies UPC-level goods by 10 departments, 110 product groups, and 1,100 product modules. All products include UPC code and description, brand, multipack, and size, as well as NielsenIQ codes for department, product group, and product module. Some products contain additional characteristics (e.g., flavor). For each UPC code, participating stores report units, price, price multiplier, baseline units, baseline price, feature indicator, and display indicator. NielsenIQ also reports store-level information about store chain code, channel type, and area location. Retailer names are masked to protect identity.

We use a concordance provided by the Bureau of Labor Statistics (BLS) between NielsenIQ product modules and ELIs. These ELIs then map to PCE categories. This concordance allows us to make direct comparisons between NielsenIQ and PCE. The food sector is identified as the aggregation of 21 PCE food categories. The 21 disaggregated PCE food categories are Bakery, Beef and Veal, Beer, Cereal, Coffee and Tea, Dairy, Eggs, Fats and Oils, Fish and Seafood, Fruit, Milk, Other Foods, Other Meats, Pork, Poultry, Processed Fruit and Vegetables, Soda and Juice, Spirits, Sugar and Sweets, Vegetables, and Wine.

Furthermore, we aggregate the categories up to aggregate food. There are three aggregated food categories, which are Aggregate Food, Aggregate Food and Nonalcoholic Beverages, and Aggregate Food and Beverages. The Aggregate Food category aggregates the 16 food categories, excluding the following 5 beverage categories: Beer; Coffee and Tea; Soda and Juice; Spirits; and Wine. The Aggregate Food and Nonalcoholic Beverages excludes only the 3 alcoholic beverages: Beer, Spirits, and Wine. The Aggregate Food and Beverages category aggregates all 21 food categories. Our main focus is the broadest category, Food and Beverages.

Our sample using NielsenIQ data covers the period from 2006Q1 to 2019Q3, and our main analysis is at the PCE category-quarter level. For each PCE category, we generate quarterly nominal sales, prices, and real sales indices. The real sales indices are the nominal sales divided by the relevant price index. To calculate sales for aggregate food categories such as Aggregate

Food and Beverages, we sum the disaggregated category-level nominal sales. For prices for those aggregate items, we use Divisia-style market share weights at the disaggregated PCE category level and aggregate prices across the categories.

To construct our main dataset from NielsenIQ, we start with the raw weekly-store-UPC level data and winsorize the data following Hottman, Redding & Weinstein (2016) and Redding & Weinstein (2018). In particular, for each UPC, we drop outliers having prices above triple or below one-third of the UPC-level median price in a given month or having quantities sold more than 24 times of the median quantity sold per month. Also, we winsorize quarterly price and sales growth of UPCs at the top and bottom 1% over a pooled sample that includes all UPCs and quarters. And then, we map this data to the PCE using the BLS concordance, which allows us to create series for disaggregated Food and Beverages categories in NielsenIQ.

4.2.2 Circana data

As a secondary source of POS data we use proprietary data from Circana, formerly the NPD Group, provided the U.S. Census Bureau. The data consists of monthly sales and units from 2017Q1 to 2020Q4. Circana has more than 65,000 retail stores, and includes online retailers unlike the NielsenIQ Retail Scanner dataset. We focus on products in Circana that most closely resembles the aggregated PCE category video, audio, photographic, and information processing equipment and media.[6] We often refer to this aggregate category as aggregate tech. We create a concordance mapping subclasses to PCE categories to help us determine the coverage of Circana and look at patterns of disaggregated PCE categories. There is high turnover for these products and rich attribute detail as explained in Ehrlich et al. (2023).

[6]There are two PCE disaggregated categories that we exclude due to coverage issues: recording media and computer software.

4.2.3 Official PCE Data

We use the official personal consumption expenditures (PCE) and price index series provided by the U.S. Bureau of Economic Analysis (BEA) as a comparison to NielsenIQ sales and price index series. The BEA releases official statistics of real personal consumption expenditures and price indices by detailed types of products (e.g. Bakery products, Cereals, Beef and Veal, etc.) at the quarterly frequency. The product types are consistent with the PCE food categories that we referred to in the previous section. For each category, we use the seasonally adjusted quarterly data for both expenditure and price indices, and compute the real sales as before. Also, for the three aggregate food categories, we continue to use the concordance provided by the BLS.

4.2.4 Supplemental POS Data

We use supplemental POS data for disaggregated food categories provided by NielsenIQ and Circana (formerly IRI). The main benefit of this data is that it covers a later time period, which allows us to compare inflation patterns that occurred during and after the pandemic recession. The supplemental data from NielsenIQ covers 4-week interval data of sales and units by UPC from 2018q1 to 2023q3. The supplemental Circana data consists of monthly sales and units from 2018q1 to 2022q3. These data are pre-grouped by the relevant data provider into approximate PCE disaggregated categories. The supplemental NielsenIQ data roughly corresponds to four PCE disaggregated categories: (1) cereal, (2) milk, (3) eggs, and (4) soda.[7] The supplemental Circana data roughly corresponds to roughly three PCE disaggregated categories: (1) bakery goods, (2) coffee, and (3) soda.[8] Both supplemental datasets are better at covering online sales given than the NielsenIQ Retail Scanner dataset, which has no online coverage. The supplemental NielsenIQ data is much richer than our baseline NielsenIQ dataset given that it covers a larger share of the universe of sales. Specifically, the supplemental NielsenIQ data also has online

[7]NielsenIQ uses the following labels for these categories: (1) ready to eat cereal, (2) cow's milk, (3) chicken eggs, and (4) soft drinks.

[8]Circana uses the following labels for these categories: (1) bakery snacks, (2) coffee, and (3) carbonated soft drinks.

sales that go through the store. The supplemental Circana POS data also has even richer online coverage of first party sales.

4.3 Methodology

4.3.1 Concordance between NielsenIQ and PCE

In order to ensure that we compare the same set of goods between NielsenIQ and PCE across all categories, we use a concordance provided to us by the BLS, which has a mapping between NielsenIQ product modules and entry-level items (ELIs). ELIs are the smallest sampling unit used by the BLS in constructing Consumer Price Index (CPI) series. All 21 food PCE categories are aggregations of multiple ELIs, except for eggs. For example, the ELI that maps into the PCE category eggs is FH011.[9] For almost all of the disaggregated price indices that we analyze, the Consumer Price Index for All Urban Consumers (CPI-U) is identical to the PCE price index that we use. The potential differences arise once we aggregate given that different weights and different formulas are used by the BEA and the BLS.

There is a duplicates issue with using the concordance provided by the BLS to map NielsenIQ product modules to PCE categories. Given that some product modules could be mapped to more than one ELI category, some of these product modules are mapped to multiple PCE disaggregated categories. These duplicate product modules have some UPCs in one PCE category and some UPCs in another PCE category. This issue will be revisited in Section 4.5. In order to only include UPCs that are properly mapped to the PCE categories that we analyze, we exclude duplicate product modules. This is a conservative estimate where we assume that all of the sales that we exclude from these product modules belong entirely to the corresponding PCE food category.

[9] For a full mapping of ELIs to PCE categories, refer to the following BLS documentation.

4.3.2 Correcting Store Turnover in NielsenIQ

Using the winsorized dataset at the week-store-UPC-PCE category level, we generate sales indices in NielsenIQ. Here, we apply our own correction method to avoid potential issues attributed to store coverage variations that occurred across time in NielsenIQ. There is natural store turnover where stores enter and exit the sample over time. This would not be problematic if entry and exit reflected true entry and exit of stores in the population. The retail scanner dataset contains information from approximately 30,000-50,000 individual stores from approximately 90 retail chains, but these numbers vary by year with significant changes in certain years.[10] Furthermore, there exists a general trend of spurious store turnover where stores enter in the first two years (2006-2008) because store coverage is increasing in NielsenIQ during this period. Thus, using the raw sales data without a correction method would lead to spurious growth in nominal sales.

To resolve this concern, we develop and apply a correction method in the following way. For each PCE category in a given week, we sum the UPC-level sales up to the store level and get total sales per store. Next, for each PCE category, we identify stores that are continuously present for two consecutive quarters. To be specific, we use stores having non-negative total sales for every single week for two consecutive quarters (i.e. 26 consecutive weeks). For expositional convenience, we refer to these stores as "continuing stores."

We also apply a similar correction method at the retail chain level. We use the weekly-store-UPC-PCE category data, and link it to store chain codes provided by NielsenIQ to identify retail chains that each store belongs to. We apply the chain correction method in the same way as the store correction method, using the total sales per chain in a given week for each PCE category and identifying retail chains that have non-negative sales for 26 consecutive weeks. As before, we refer to these chains as "continuing chains". This is our preferred correction method given that spurious entry and exit as the chain level. If a store that belongs to a chain is in the NielsenIQ

[10] The data version released on May 31, 2020 had a significant drop in store coverage in 2017. Compared to 2011 when NielsenIQ covered 53% of the full universe food chain stores, it dropped to 26% in 2017. This was later amended by NielsenIQ updating the data.

sample, then typically all the stores of that chain are also in the NielsenIQ sample.

Next, for each PCE category, we restrict our sample to either the continuing stores or the continuing chains, and aggregate the weekly store or chain sales to the quarterly frequency using the National Retail Federation (NRF) calendar. Using the quarterly total sales, we calculate the growth rates (log first difference) to the continuous store or chain sales series and build a quarter-level index for the pairwise continuers. Note that for the initial quarter, we use total sales of all stores or chains in the sample, and accumulate the imputed growth rates to build a quarterly sales index. We call the final output, which is an imputed quarterly sales index, "store-corrected" or "chain-corrected" sales index. We then seasonalize the series. Our main analysis is based on the chain corrected sales series, and the store corrected series is used as robustness check.

4.3.3 Price Indices in NielsenIQ

To build quarterly price indices, we average prices across weekly transactions for each UPC. To do so, we aggregate the weekly UPC sales and units sold to the quarterly frequency from the raw data. Here, we acknowledge that the UPC level information is based on different size or package units. Thus, when we compute prices at the quarterly level, we unify the package units across all UPCs and normalize the prices to the same unit (e.g. ounces), following Hottman et al. (2016). Note that this was done both with and without the store/chain correction method, and the correction did not noticeably affect price indices. Our main analysis for price indices uses the whole sample without applying either correction method.[11] The price series is at the PCE category-quarterly level, and seasonally adjusted. We display more details on price indices in the following subsection.

4.3.3.1 Traditional Indices

For each PCE category, we construct price indices from the UPC-level quarterly price data. As a starting point, we have used traditional price indexes, such as Laspeyres, Paasche, and

[11] We checked price series with the correction method as a robustness test, which is available on request.

Tornqvist indexes. Our main focus is on the log geometric price index given by:

$$\ln \Phi_t^G = \sum_{k \in \mathbb{C}_{t-1,t}} w_{kt} \ln \frac{p_{kt}}{p_{kt-1}},$$

where w_{kt} is a weight assigned to product k (typically based on the product's market share). The set $\mathbb{C}_{t-1,t}$ is the set of all "continuing" goods that are sold both in period t and in period $t-1$. The expenditure share for product k is:

$$s_{kt} \equiv \frac{p_{kt} c_{kt}}{\sum_l p_{lt} c_{lt}} \tag{4.1}$$

where c_{kt} is the quantity of good k purchased in period t. The Laspeyres index uses lagged expenditure shares as weights ($w_{kt} = s_{kt-1}$), the Paasche index uses current expenditure shares ($w_{kt} = s_{kt}$), and the Tornqvist index uses average expenditure shares ($w_{kt} = \frac{s_{kt-1}+s_{kt}}{2}$).

4.3.4 Methodology used in PCE Estimates

The PCE estimates are only complete in benchmark years known as the Economic Census (EC) years. A benchmark year is a year in which the benchmark Input-Output (I-O) accounts are used to establish the level of PCE and its components during a comprehensive update. The primary sources for PCE estimates are BEA's Benchmark I-O Accounts, which are in turn based on Census Bureau's Economic Census (EC); BEA's International Transaction Accounts; the Census Bureau's Annual Retail Trade Survey (ARTS), Service Annual Surveys, Quarterly Service Reports, and Monthly Retail Trade Surveys (MRTS); and the BLS Consumer Price Indexes.

The PCE Handbook Chapter 5 summarizes the methodology used to prepare the estimate for PCE goods, in particular for the following items of our interest: Food and nonalcoholic beverages purchased for off premises consumption.

In the benchmark year estimates are prepared using the commodity-flow method, starting with manufacturers' shipments from the Economic Census (EC). The commodity flow method is generally used to derive estimates in EC years for various components of consumer spending, equipment and software, and the commodity detail for state and local government consumption expenditures and gross investment. Generally, the method begins with an estimate of the total supply of a commodity available for domestic uses.

In nonbenchmark years excluding the most recent one the retail control method is applied using retail sales from ARTS. Composition of goods sold largely based on scanner data from Information Resources Inc. and from Fresh Look Marketing Group (acquired by IRI). This method uses retail sales data (MRTS and ARTS), compiled by the Census Bureau, to estimate annual, quarterly, and monthly consumer spending on most consumer goods. In these nonbenchmark years, PCE is derived by extrapolation from the benchmark year using a retail control total of sales by kind of business from the monthly and annual surveys. Product ratios are calculated in the benchmark year and then these weights are extrapolated to the retail sale data by kind of business. However, the only store type with updated ratios in nonbenchmark years within Aggregated Food and Beverages is grocery stores. Similarly, the only store type with updated ratios in nonbenchmark years in Aggregated Tech is consumer electronic stores.

This is the general process for the retail control method. However, when the BEA uses scanner data such as for goods bought at grocery stores, the EC allocations are updated annually using retail point of sale scanner data from Circana, formerly Information Resources, Inc. This specifically means that instead of using product ratios in column 3 from the benchmark year, extrapolated ratios are calculated from the relevant scanner data. For example, if beef comprises 10% in grocery stores in Year 1 (Economic Census), 11% in grocery stores in scanner data in year 1, and 12% in grocery stores in the scanner data in year 2, then the extrapolated ratio would be 10.9% for beef within grocery stores. This is calculated by multiplying economic census ratio by the growth (10*12/11).

4.4 Results

In this section, we document how real sales indices using POS data compares with official PCE measures. We first focus on the aggregated results using NielsenIQ data for Aggregate Food and Beverages and Circana data for Aggregate Tech. We then focus on several disaggregated categories within each of those two broad aggregates. Given differences in real sales indices between the POS data and PCE data, we also provide external sources that validate the POS data. Next, to better understand why PCE measures could be deviating from the POS data, we perform a counterfactual exercise where we replicate BEA's methodology in creating sales indices. Finally, we discuss the patterns and disparities observed between POS data and official statistics during the pandemic period and beyond.

4.4.1 Aggregated Results

In order to compare aggregated real sales between POS and official PCE data, we focus on the geometric Laspeyres deflator to see if there are variations in underlying patterns beyond differences in price indices. In Figure 4.1, we find similar patterns in real sales for Aggregate Food between the NielsenIQ POS and PCE data based on the geometric Laspeyres price index. However, if we had used a different deflator such as the Tornqvist, we would have seen a 20% growth in real sales from 2012Q1 to 2019Q3 instead of 10%. The choice of deflator has significant implications for real sales growth. The official PCE measure understates real sales growth by using the geometric Laspeyres index.

Even though we observe similar patterns in real sales for Aggregate Food between the POS and PCE data, there could be differences at the price index or nominal sales level. In Figure 4.2 we notice some disparities between NielsenIQ POS and PCE data emerging. Specifically, the geometric Laspeyres price index relative to the PCE price index starts to diverge in 2012, with PCE exhibiting higher inflation than NielsenIQ until 2018. However, starting around 2015, nominal sales indices between PCE and NielsenIQ begin to diverge, with PCE reporting higher growth. PCE has higher sales growth and inflation than NielsenIQ for Aggregate Food, which results in offsetting outcomes where real sales appear close between the POS and PCE data.

Additionally, we compare NielsenIQ with PCE in terms of their trends, high-frequency fluctuations, and correlations for nominal sales, prices, and real sales indices. The results are presented in the last row of Tables 4.2, 4.3, and 4.4, respectively. In Table 4.2, we observe that the trend and overall growth pattern of sales appear similar between NielsenIQ and PCE, as indicated by the first three columns. However, based on the detrended series in the fourth and last columns, PCE exhibits smoother fluctuations than NielsenIQ. Tables 4.3 and 4.4 exhibit consistent patterns for price and real sales indices, respectively, with real sales showing more pronounced gap in the detrended series between NielsenIQ and PCE. These patterns are visualized in Appendix Figure C.1. Overall, both datasets closely mirror each other at the aggregate level.

For Aggregate Tech, we have a shorter time horizon to compare POS data to PCE data: 2017Q1 to 2020Q4. In Figure 4.3, we observe significant growth in real sales of Aggregate Tech for both PCE and Circana, with 80% and 90% growth, respectively. Prior to 2020, PCE generally reported slightly higher growth in real sales than Circana. However, relative to Circana, PCE fails to fully capture the surge in real sales during the pandemic. After the comprehensive update to PCE in September 2023, nominal sales in PCE for Aggregate Tech is downwardly revised. This revision is particularly pronounced in 2020 where the increase in real sales to the revised PCE series is closer to Circana. Despite nominal sales between PCE and the POS data getting closer after the revision, a larger gap emerges in real sales because the POS data exhibits more deflation than PCE.

Similar to the decomposition of real sales for Aggregate Food, there are larger discrepancies between POS and PCE in price indices and nominal sales for Aggregate Tech than in real sales. However, the differences between the POS and PCE data are especially pronounced for price indices in Aggregate Tech. In Figure 4.4, PCE price index reports a 20% deflation from 2017Q1 to 2020Q4, while Circana POS data reports almost 30% deflation during the same time period. Figure 4.4 illustrates that before the update in September 2023, PCE was reporting almost 50% growth in nominal sales, while Circana POS data reported close to 35% growth. After the comprehensive update, PCE shows similar growth in sales as Circana. However, the similar growth between the revised PCE nominal sales and Circana nominal sales between 2017Q1 and 2020Q4 understates the differences between both series. Specifically, Circana was reporting less growth than PCE for nominal sales from 2017 to 2019, but this was mostly offset by Circana reporting more growth than PCE for nominal sales in 2020.

4.4.2 Disaggregated Results

Once we switch to disaggregated PCE categories, we observe larger gaps in real sales between POS and PCE data. Generally, PCE and NielsenIQ tend to track each other closely for price indices in disaggregated food categories. Figure 4.6 illustrates that NielsenIQ closely aligns with PCE for price indices of eggs and milk for the majority of the period from 2006Q2 to 2019Q3. However, differences begin to emerge in 2015 for both eggs and milk, where PCE experiences more deflation than NielsenIQ.

There are more significant differences in nominal sales at the disaggregated PCE category within Aggregate Food. Figure 4.7 highlights the discrepancies in nominal sales for eggs and milk. For eggs, the PCE nominal sales appear too smooth relative to the NielsenIQ POS data. For example, PCE fails to capture the substantial increase in nominal sales in 2015 during the bird flu outbreak. The disparity between the POS and PCE data is even more pronounced for milk. Instead of merely missing high-frequency fluctuations, PCE and NielsenIQ have different trends starting around 2015. The PCE data shows a modest increase in sales, while NielsenIQ scanner data indicates a decline in sales. Even after the update in September 2023, neither of these issues is resolved.

Here again, we compare NielsenIQ with PCE in terms of their trends, high-frequency fluctuations, and correlations for nominal sales, prices, and real sales indices at the disaggregated level. We examine seven items, Cereal, Coffee and Tea, Dairy, Eggs, Milk, Other Foods, and Soda and Juices, and report the results in Tables 4.2, 4.3, and 4.4, for nominal sales, prices, and real sales, respectively (see the first seven rows). In Table 4.2, we find that the similarity between NielsenIQ and PCE (in terms of trends and general growth pattern) is lower for disaggregated items than for aggregate food. For milk, the correlation between NielsenIQ and PCE is even negative. Furthermore, it is generally observed that PCE is smoother than NielsenIQ across all items. Table 4.3 shows the results for price index, where NielsenIQ in general exhibits more growth than PCE between 2008Q1 and 2019Q3, as indicated by the first column. Table 4.4 demonstrates that very distinct patterns between PCE and NielsenIQ are yielded for real sales. There are even negative correlations for cereal, milk, and eggs for both the raw series and trend. In addition, the detrended series and fluctuations of PCE are significantly different from NielsenIQ. These differences are also evident in Appendix Figures C.2-C.8. Overall, unlike aggregate food, PCE struggles to match NielsenIQ at the disaggregated level.

There are only six disaggregated PCE categories for Aggregate Tech as opposed to the 21 disaggregated PCE categories for Aggregate Food. One of the six PCE categories where there are large differences between PCE and Circana POS data is photographic equipment. Figure 4.8 shows the differences in trends for price and sales indices within photographic equipment between PCE and Circana. In Figure 4.8 the POS Circana data shows slightly over 10% deflation from 2017Q1 to 2020Q4, while the PCE data shows slightly under 5% inflation over the same time period for photographic equipment. In Figure 4.8, Circana shows slightly over a 20% decline in nominal sales from 2017Q1 to 2020Q4, while the PCE data shows slightly over a 20% increase in sales over the same time period for photographic equipment.

4.4.3 External Validity

In order to provide evidence that POS data does a better job of capturing real sales, particularly for nominal sales, we provide external sources that validate the patterns described by the POS data. For example, Figure 4.9 shows external sources from the USDA that closely track the NielsenIQ nominal sales indices. Since the USDA data only provides units, we use the PCE price index to create a sales index. Figure 4.9 displays that PCE nominal sales missed the 2015 bird flu episode.[12] Similarly, Figure 4.9 depicts that the external source captures the secular decline of milk sales. NielsenIQ nominal sales also reflect this secular decline, while PCE sales show growth.

This external validation of nominal sales between the POS data and PCE can be applied to cereal as well as beer. In Figure 4.10, we observe stagnant growth in NielsenIQ nominal sales, while there are large increases for both beer and cereal PCE nominal sales. For cereal and beer, we have two external sources. For cereal, one external source lies above NielsenIQ nominal sales, and the other lies below NielsenIQ nominal sales. For beer we are only able to get external sales data starting from 2014. Both of these external beer nominal sales series are closer to the NielsenIQ series and depict lower growth than NielsenIQ. According to news articles, beer and cereal sales experienced a slump due to changes in consumer preferences away from these products.[13]

This external validation also applies to the disaggregated PCE categories within Aggregate Tech. Specifically, the external source for nominal sales of photographic equipment closely lines up with the Circana nominal sales as seen in Figure 4.11. The external source uses domestic and foreign shipments data and, similarly to Circana, shows a decline in sales from 2017Q1 to 2020Q4. In contrast, PCE sales show growth during this time period of slightly over 20%. Despite the September 2023 comprehensive revision, nominal sales growth for photographic equipment was not significantly reduced.

4.4.4 Counterfactual Exercise

In several cases at the disaggregated PCE category level, PCE fails to accurately capture real sales patterns. This is particularly evident for nominal sales, given the larger differences between the POS and PCE data. In this section, we document how the BEA creates sales indices and highlights limitations with their approach.

[12] According to a GAO report, the 2015 bird flu episode led to 43 million layers, birds that lay eggs, to be culled.

[13] A Bloomberg article cites Americans substituting away from cereal to other breakfast items as the reason for the slump in cereal sales. Similarly, a CNN Business article cites Americans substituting away from beer to other products such as spiked seltzer.

The PCE has benchmark years, which correspond to Economic Census (EC) years. An EC occurs every five years. However, there can be a substantial lag between when an EC occurs and when its results are incorporated in the PCE nominal sales. For example, the 2017 EC was not incorporated into PCE estimates until September 2023. As can be seen in Figure 4.2, there were meaningful upward revisions to Aggregate Food starting in 2013, the year after the previous benchmark year. After September 2023, 2017 becomes the current benchmark year, and all subsequent years are nonbenchmark years until the 2022 EC gets incorporated into PCE estimates.

Not only are there issues with lags in the incorporation of EC data, but there is also a concordance issue between PCE categories and EC categories. Table 4.1 shows that EC categories tend to be more aggregated than PCE categories, particularly for food items. For example, an EC category is "eggs and dairy", which comprises three PCE categories: eggs, dairy, and milk. This makes it more challenging to capture economic trends at the PCE category level, since the EC directly shows sales growth at the aggregated level. Furthermore, even if the EC were able capture egg sales in 2017, it would miss the surge in egg sales in 2015 as a result of the 2015 bird flu. The BEA's approach of using the EC leads to nominal sales that are too smooth, particularly in nonbenchmark years.

In nonbenchmark years, the BEA employs a share approach, which is based on store type. For example, within food categories, store types could include grocery stores, convenience stores, and warehouse clubs. However, only the share for grocery stores is updated within food categories. All other shares, such as convenience stores, remain fixed to the most recent benchmark year. That means that PCE nominal sales estimates prior to September 2023 for 2023Q2 were using weights from 2012, except for grocery store types. For grocery store type, the BEA uses scanner data from Circana (formerly IRI) to update nominal sales in nonbenchmark years.

Given the information on how the BEA updates its nominal sales data, we replicate their methodology for a few disaggregated PCE categories. In Figure 4.12, we find that the replication of the BEA methodology leads to much closer estimates to PCE than NielsenIQ for both eggs and milk. In order to replicate the BEA methodology, we separate sales by store type (i.e. grocery stores or mass merchandisers). Then, we measure the annual share of the aggregated EC category, eggs and dairy, in every year starting from 2012 (average from 4 quarters) from Food and Beverages for grocery stores. We also estimate the share of each EC category in 2012 (average from 4 quarters) from Food and Beverages for all other store types.[14] Finally, we sum the sales across all store types.

We are able to create counterfactual exercises using NielsenIQ. However, we cannot precisely replicate the BEA approach with Circana because we lack store type information. The BEA employs a similar approach for PCE categories in Aggregate Tech, updating only consumer electronic store types (instead of updating grocery store types only). They continue to use POS data to update consumer electronic store types, sourced from Circana (formerly NPD). Following this process, Figure 4.13 shows what happens with two different counterfactual exercises for photographic equipment. One uses the annual share of photographic equipment and updates annually, while the other series locks shares to 2017 levels. The counterfactual fixing 2017 shares gets closer to PCE estimates. However, this counterfactual exercise is not an ideal replication due to the lack of Circana data prior to 2017Q1 and the absence of store type information.

4.4.5 Disaggregated Sales Affect Aggregate Prices

In this section, we document how disaggregated nominal sales affect aggregate price indices. Specifically, we focus on PCE Food and Nonalcoholic Beverages, which is composed of 18 disaggregated series. Even if price indices are being measured correctly at the disaggregated level, price indices at the aggregate level could be unrepresentative if the expenditure weights employed are wrong.

In Figure 4.14, we plot eight price indices for Food and Nonalcoholic Beverages from 2006q2 to 2019q3. Four price indices are official price indices: (1) PCE Price Index (August 2023 revision), (2) PCE Price Index (September 2023 revision), (3) BLS Consumer Price Index

[14] These 2012 shares for non-grocery stores are not updated in order to be consistent with the BEA procedure.

(CPI), and (4) BLS Chained Consumer Price Index. We use the geometric Laspeyres price index constructed from POS data as a fifth series. The remaining three price indices are counterfactual PCE price indices where we use alternative weights or aggregation methods: (1) PCE with PCE Divisia weights, (2) PCE with NielsenIQ Divisia weights, and PCE with PCE weights and Fisher aggregation. One should note that the PCE and CPI use different aggregation methods. Specifically, the CPI is based on a Laspeyres formula, while PCE is based on a Fisher formula.[15]

Five of the eight series in Figure 4.14 are indistinguishable from each other: (1) PCE Price Index (August 2023 revision), (2) PCE Price Index (September 2023 revision), (3) BLS Consumer Price Index (CPI), (4) PCE with PCE Divisia weights, and (5) PCE with PCE weights and Fisher aggregation. There are three series that noticeably diverge from these five series: (1) BLS Chained Consumer Price Index, (2) PCE with NielsenIQ Divisia weights, and (3) NielsenIQ geometric Laspeyres. The Chained CPI reports the lowest cumulative inflation by the end of the sample followed by the PCE with NielsenIQ Divisia weights.

The Chained CPI uses the Tornqvist index formula for upper-level aggregation, while the CPI employs a Laspeyres index formula. From a theoretical perspective, Tornqvist weights are preferred to Laspeyres weights because Tornqvist weights allows more flexibility in adapting to consumption basket changes. Consequently, Tornqvist weights exhibit less substitution bias than Laspeyres weights. According to the BLS, they view the Chained CPI as a superlative index that better approximates a cost of living index than the CPI.[16] Given that our PCE with NielsenIQ Divisia weights price index better approximates the Chained CPI, we can see that the weights employed from the disaggregated series affect aggregate inflation at the Food and Nonalcoholic Beverages level.[17]

[15]For more details between PCE and CPI, see this BLS report.

[16]See the BLS response comparing the CPI and C-CPI on their website.

[17]We use Food and Nonalcoholic Beverages in this exercise to have similar consumption baskets between the CPI and BLS that excludes food away from home.

4.4.6 Patterns in the Pandemic and Its Aftermath

In this section, we document additional patterns observed over the pandemic period and beyond, using the two types of POS data and PCE series. The POS datasets consist of NielsenIQ and Circana IRI as before, but these are different datasets from what we have used in the main analysis. These data are from the same providers as the baseline datasets but span a more recent yet more limited time period.

Using the datasets, we compute seasonally-adjusted series of nominal sales, traditional prices, and real sales indices for each item at the quarterly level. Figures 4.15 and 4.16 illustrate the annual inflation rates and price indices, respectively, for cereal, milk, eggs, and soda in NielsenIQ and PCE. Generally we find that the PCE data reported higher inflation starting in the second quarter of 2020 and underreported the surge in inflation that started in 2020. This period is when we see consistently the largest differences in inflation between official statistics and POS data within Food and Beverages. Potentially, the official statistics were not well equipped to handle the pandemic shock where field collection was no longer possible. [18] Furthermore, there was a large shift in this period in the share of products purchased online, which would be difficult for official statistics to measure given the lagged expenditure weights. Another issue during this period is the lagging store type weights, which are only updated through the Economic Census.

In Figure 4.15, we see inflation surge during 2021 and 2022 for both NielsenIQ and PCE data. However, in 2021 inflation surges more in NielsenIQ POS data than in PCE, except for cereal. For instance, based on the geometric Laspeyres index, NielsenIQ inflation for eggs was 51.2% while PCE shows 36.8% on average in 2022. More details can be found in Table 4.5. Interestingly, for cereal, NielsenIQ exhibits less inflation than PCE prior to the surge, and price levels also remain higher for PCE than NielsenIQ.

Figure 4.17 presents the nominal sales indices for the same set of items in both NielsenIQ and PCE. In general, PCE displays less volatility and does not capture high-frequency variation well. For instance, PCE sales for eggs entirely missed the bird flu case in late 2022, unlike

[18] See this BLS documenation on the COVID-19 pandemic on CPI data.

NielsenIQ. Furthermore, there was a benchmark revision in PCE in September 2023, where PCE was benchmarked to 2012 Economic Census prior to September 2023, and the updated version was benchmarked to 2017 Economic Census from September 2023. This primarily impacts nominal sales and is reflected in the gap between the two PCE series in this figure.

Figure 4.18 displays the counterparts for real sales. There exists a notable gap between NielsenIQ and PCE real indices across all item. In particular, PCE shows discrepancies in real growth, which are mainly attributed to missing inflation, as well as the high frequency volatility and trends of nominal sales.

We do the same exercises for other food items, such as bakery goods and coffee, comparing Circana and PCE. The results are consistent, as presented in Figure 4.19. Inflation shows a greater surge in 2022 in NielsenIQ related to PCE for both items. For instance, bakery goods exhibit 15.1% inflation relative to 12.4% in PCE. Furthermore, the similar pattern holds for nominal sales, where PCE does not track well high-frequency variations as observed in Circana. This impacts the growth path of real sales in the similar fashion, driven by PCE not able to capture inflation, volatility, and trends of nominal sales as shown in Circana.

4.5 Limitations of POS data

While the POS data offer numerous advantages that we leverage to improve official statistics, there are remaining limitations of using it.

First, the coverage of the data is not perfect and varies across datasets. For instance, the NielsenIQ Retail Scanner dataset provides comprehensive coverage for food items at grocery stores or general merchandise stores but lacks coverage of online retail. This discrepancy may be particularly pronounced during the period from 2012 to 2017, when e-commerce experienced rapid emergence and growth. Conversely, Circana provides extensive coverage of online retail and is not subject to this issue. The supplemental data is also not subject to this critique.

Second, while the scanner data in NielsenIQ includes products labeled with a UPC, it lacks non-UPC products like meat, fruits, and vegetables that are weighed at the counter (referred to as "magnet data").[19] In Figure 4.20, we present the sales ratio of NielsenIQ to PCE for each food item, where the sales index is normalized at 2012Q1, and the ratio is taken at 2019Q3. NielsenIQ exhibits similar coverage to PCE for most items, except for Fresh Fruit, Poultry, and Beef and Veal, which may be attributed to the magnet item coverage. However, it's noteworthy that other meat categories, such as Pork or Other Meats, are less affected by this issue, indicating heterogeneity in sensitivity to the magnet issue across different food items. To mitigate potential bias stemming from this issue, we conduct robustness checks by excluding relevant food items related to meat, fruits, and vegetables from analyses for aggregate foods. The results remain largely unchanged.

Third, NielsenIQ has time-varying store coverage depending on its contract with each store, which can lead to store turnover attributed to each contract renewal. In essence, some stores enter and exit the sample due to these renewals, potentially not accurately reflecting the true entry and exit of stores in the population. One feature of these contract renewals is that stores get added or removed at the chain level, which is why our preferred correction method is at the chain level. This chain correction method resolves the spurious entry and exit of stores from the sample. As a robustness check, we apply this correction method at the store level as well and find consistent results.

Fourth, the current concordance between NielsenIQ and PCE is at the disaggregated level, using NielsenIQ product modules and entry-level items (ELIs) in BLS, to try to compare the same set of goods as much as possible. However, this concordance may not be perfect as mentioned before, there are duplicate product modules in NielsenIQ that are mapped into more than one PCE category. For the baseline analysis, we drop these duplicate product modules. These duplicate product modules account for less than 10% in each PCE category with the exception of eggs and milk. We have conducted robustness checks whre we include these duplicate product modules and find similar results.,[20]

[19]The magnet data began to be incorporated into the scanner data in 2020 or 2021, so we expect an improvement in coverage for recent years. Additionally, unlike the scanner data, the consumer panel in NielsenIQ includes the magnet data, which can be an avenue for checking robustness.

[20]For instance, milk comprises two duplicate modules out of five product modules. One of these duplicates

Lastly, unlike NielsenIQ, Circana data have a limited sample duration of four years and do not track store types. This short sampling period is not optimal for seasonalization using methods like X-13 ARIMA SEATS, which is necessary for comparison with PCE that is always seasonalized. Furthermore, Circana masks certain products, by only providing their sales (instead of prices, quanities, and attributes). These masked items do not affect sales since they are being included, but might affect our price indices. This limitation could pose a larger challenge for hedonic price indices, particularly as we rely on the attributes of item-level products.

4.6 Concluding Remarks

In this chapter, we conduct a comparative analysis of prices, nominal sales, and real sales between official statistics and measures derived from item-level transactions data and explore potential improvements in official statistical methodology. Our study highlights significant disparities between these two sets of measures, with the discrepancies largely stemming from the methodology used by official statistics. These official measures rely on multiple surveys and statistical sources that often suffer from long lags, resulting in outdated information regarding output composition. Our findings indicate that the PCE data is smoother than the item-level transaction data and often fails to capture high-frequency variations, such as the 2015 bird flu episode. These discrepancies are mainly observed at the disaggregated level for both trends and fluctuations and are mitigated at the aggregate level. Furthermore, we provide external validation for the transactions data by confirming their consistency with external sources when their nominal sales trends diverge significantly from official statistics. Our counterfactual exercises, which replicate the methodology used in official statistics, reveal similar patterns between official nominal sales and item-level transactions data, even in cases where their trends exhibit notable differences.

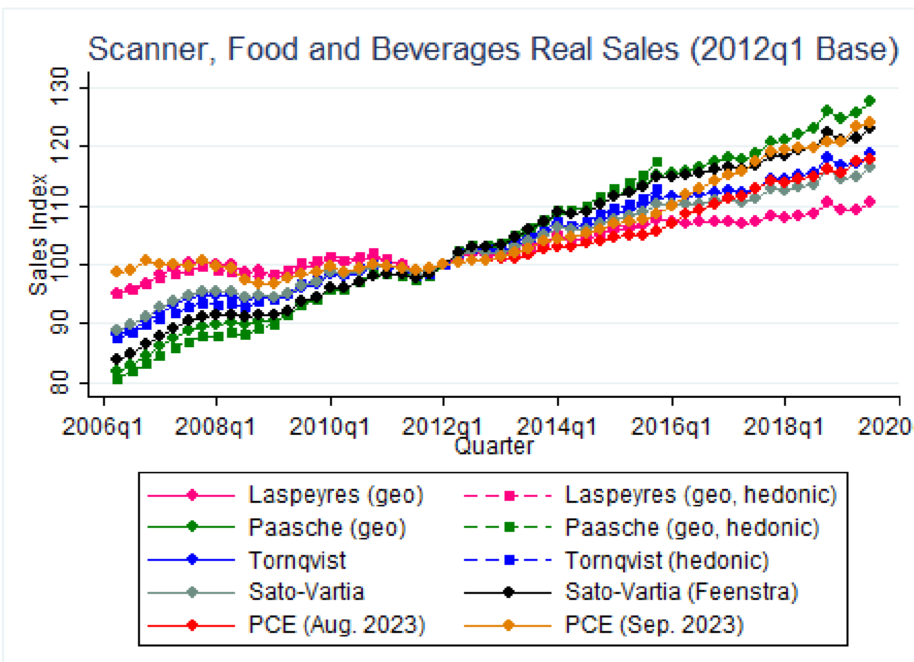
Furthermore, we document that although inflation patterns within Food and Beverages are

shows a sales share that increases over time, surpassing 10% after around 2012. Upon closer examination of the composition of the duplicate modules, they likely represent plant-based milk and its increasing demand over time. This trend is absent in the current concordance, unlike PCE, which includes plant-based milk items. Nevertheless, we observe that the overall pattern of sales index for milk remains largely similar even after incorporating the two duplicate modules.

similar between the POS data and PCE data, noticeable differences emerge during the pandemic. The infrastructure of official statistics was poorly equipped to handle a pandemic when personal visits by statistical agencies ceased in conjunction with large changes in expenditure weights across store types. Monetary policy in the United States was slow to respond to inflation by waiting until 2022 to increase the federal funds rate. Monetary policy may have responded more quickly during the recession had they known the true extent of inflation. Despite some remaining limitations in the use of item-level transactions data, this chapter could serve as a catalyst for discussions among researchers, policymakers, and statisticians on how to enhance the accuracy and timeliness of economic statistics.

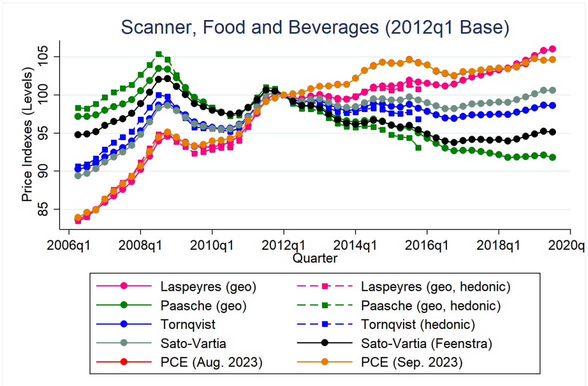
A Figures

Figure 4.1: Real Sales Index for Aggregated Food

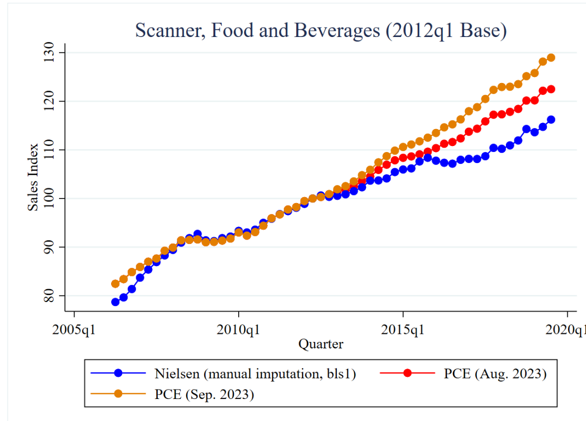


Note: The figure represents the real sales indices from 2006q2 to 2019q3 using NielsenIQ data in comparison to official PCE measures for aggregated Food and Beverages. For the NielsenIQ data we create eight different real sales indices by deflating nominal sales by eight different price indices. We use three traditional price indices: geometric Laspeyres, geometric Paasche, and Tornqvist. We use two demand-based price indices: Sato-Vartia and Sato-Vartia Feenstra-adjusted. We use four hedonic price indices based on the other price indices except the Sato-Vartia price index.

Figure 4.2: Real Sales Decomposition for Aggregated Food



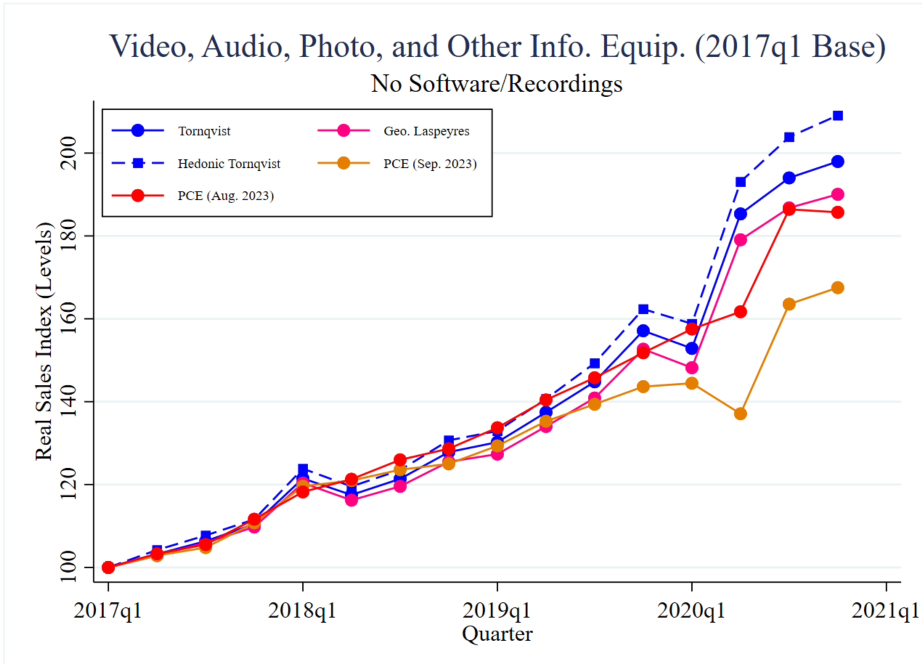
Price Index



Sales Index

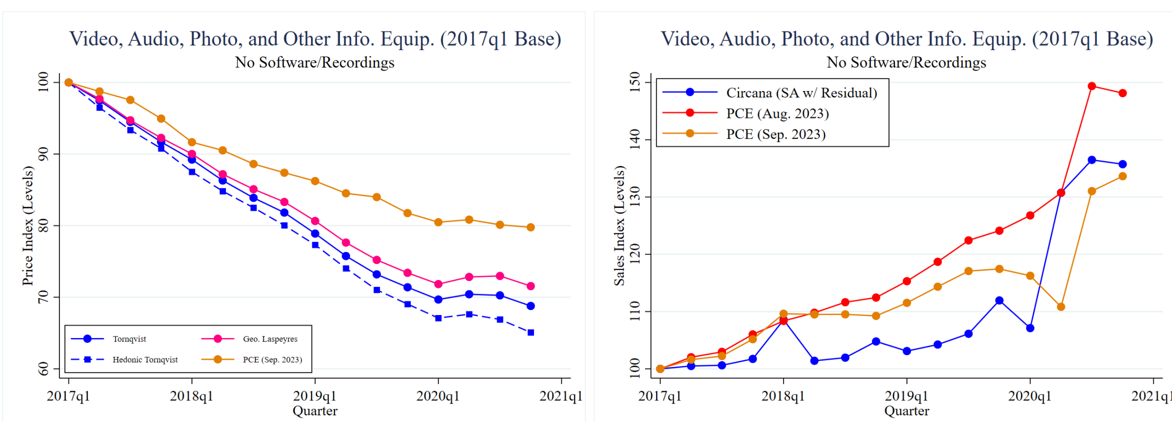
Note: These figures represent the price indices and nominal sales indices from 2006q2 to 2019q3 using NielsenIQ data in comparison to official PCE measures for aggregated Food and Beverages. The panel on the left has PCE official price index in addition to eight different price indices computed using NielsenIQ data. The panel on the right has nominal sales from PCE and NielsenIQ. There are two versions for official PCE series to reflect a version prior to the 2023 comprehensive update and a version after the update. In this and subsequent figures comparing NielsenIQ results to official series, we often use *Nielsen* as a shorthand for expositional convenience in the labels in the legend.

Figure 4.3: Real Sales Index for Aggregated Tech



Note: The figure represents the real sales indices from 2017q1 to 2020q4 using Circana data in comparison to official PCE measures for aggregated Video, Audio, Photographic, and Other Information Processing Equipment. The Circana and PCE data exclude software and recordings from this aggregated category. For the Circana data we create 3 different real sales indices by deflating nominal sales by 83different price indices. We use 2 traditional price indices: geometric laspeyres and Tornqvist. The third price index is the hedonic Tornqvist price index.

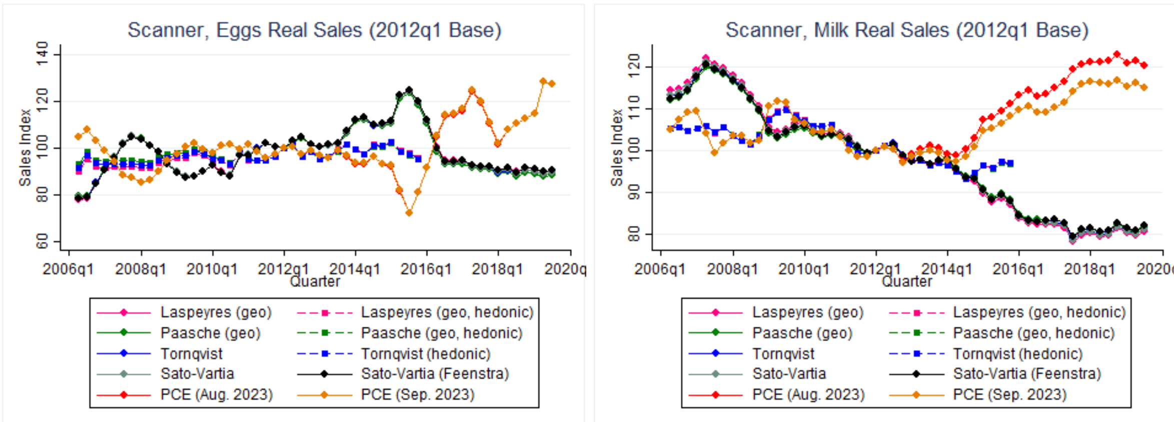
Figure 4.4: Real Sales Decomposition for Aggregated Tech



Price Index Sales Index

Note: These figures represent the price indices and nominal sales indices from 2017q1 to 2020q4 using Circana data in comparison to official PCE measures for aggregated Food and Beverages. The panel on the left has PCE official price index in addition to three different price indices computed using Circana data. The panel on the right has nominal sales from PCE and Circana. There are two versions for official PCE series to reflect a version prior to the 2023 comprehensive update and a version after the update.

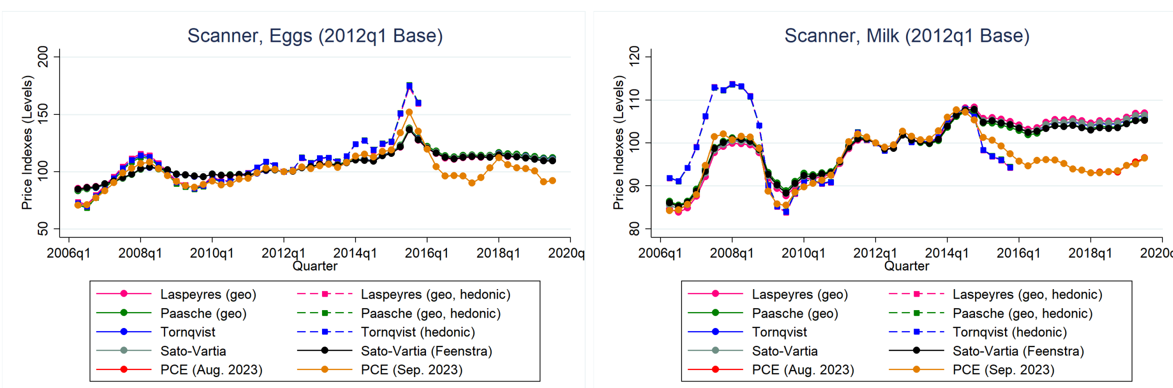
Figure 4.5: Real Sales Index for Disaggregated Food PCE Categories



Eggs Milk

Note: These figures represent real sales indices from 2006q2 to 2019q3 using NielsenIQ data in comparison to official PCE measures for two disaggregated PCE food categories: eggs and milk. The panel on the left has real sales indices for eggs. The panel on the right has real sales indices from milk. Each panel has eight series. Seven series are different price indices created using NielsenIQ POS data. The eight series is the official PCE price index.

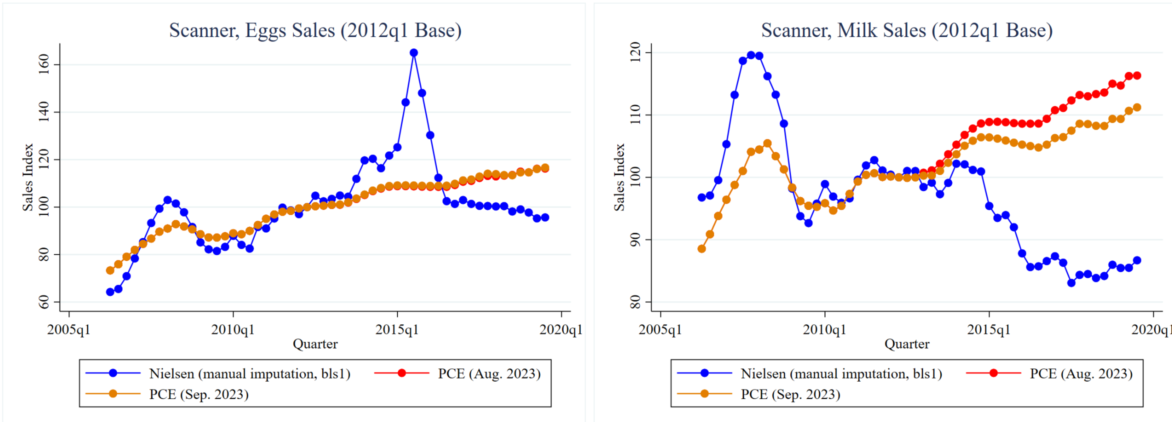
Figure 4.6: Disaggregated Food PCE Categories (Price Index)



Eggs Milk

Note: These figures represent price indices from 2006q2 to 2019q3 using NielsenIQ data in comparison to official PCE measures for two disaggregated PCE food categories: eggs and milk. The panel on the left has price indices for eggs. The panel on the right has price indices from milk. Each panel has eight series. Seven series are different price indices created using NielsenIQ POS data. The eight series is the official PCE price index.

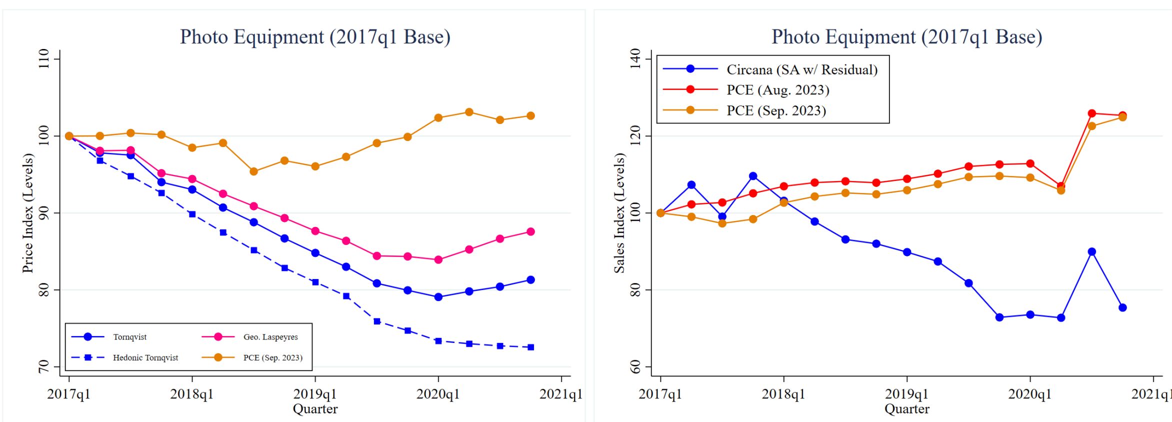
Figure 4.7: Disaggregated Food PCE Categories (Nominal Sales Index)



Eggs Milk

Note: These figures represent nominal sales indices from 2006q1 to 2019q3 using NielsenIQ data in comparison to official PCE measures for two disaggregated PCE food categories: eggs and milk. The panel on the left has nominal sales for eggs. The panel on the right has nominal sales from milk. Each panel has three series: two series for PCE nominal sales and one series for NielsenIQ sales. There are two versions for official PCE series to reflect a version prior to the 2023 comprehensive update and a version after the update.

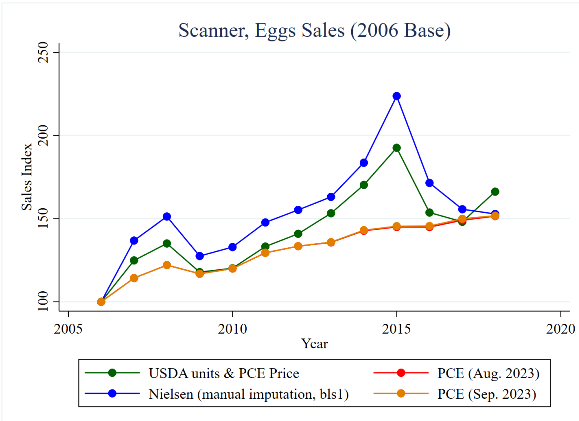
Figure 4.8: Disaggregated Tech: Photographic Equipment



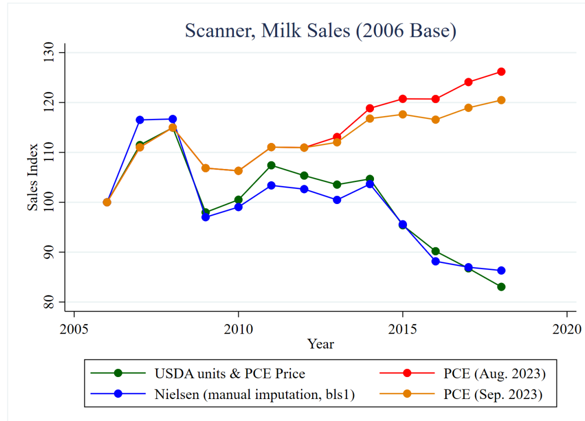
Price Index Sales Index

Note: These figures represent the price indices and nominal sales indices from 2017q1 to 2020q4 using Circana data in comparison to official PCE measures for a disaggregated tech category: photographic equipment. The panel on the left has PCE official price index in addition to three different price indices computed using Circana data. The panel on the right has nominal sales from PCE and Circana. There are two versions for official PCE nominal sales indices to reflect a version prior to the 2023 comprehensive update and a version after the update.

Figure 4.9: External Validity: Nominal Sales Index (Food)



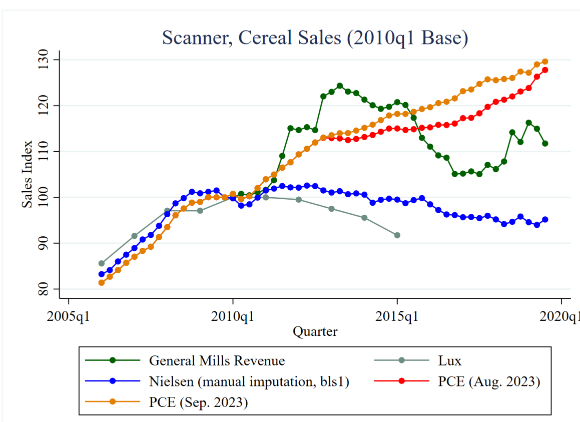
Eggs



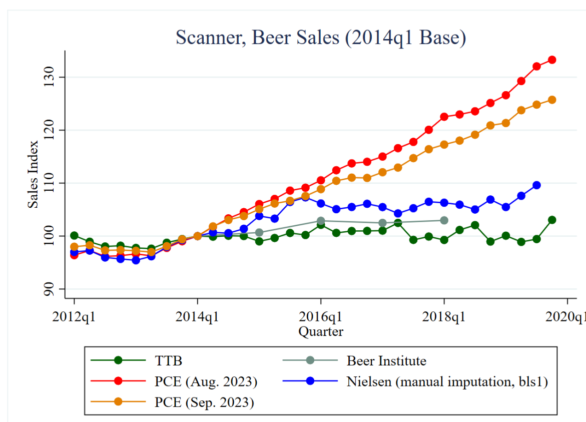
Milk

Note: These figures represent annual nominal sales indices from 2006 to 2018 using NielsenIQ data in comparison to official PCE measures for two disaggregated PCE food categories: eggs and milk. In addition to the POS and official PCE series, we provide nominal sales series based on units of eggs from the USDA and PCE price index. The panel on the left has nominal sales for eggs. The panel on the right has nominal sales for milk. Each panel has four series: two series for PCE nominal sales, one series for NielsenIQ sales, and one external series from the USDA. There are two versions for official PCE series to reflect a version prior to the 2023 comprehensive update and a version after the update.

Figure 4.10: External Validity: Nominal Sales Index (Food)



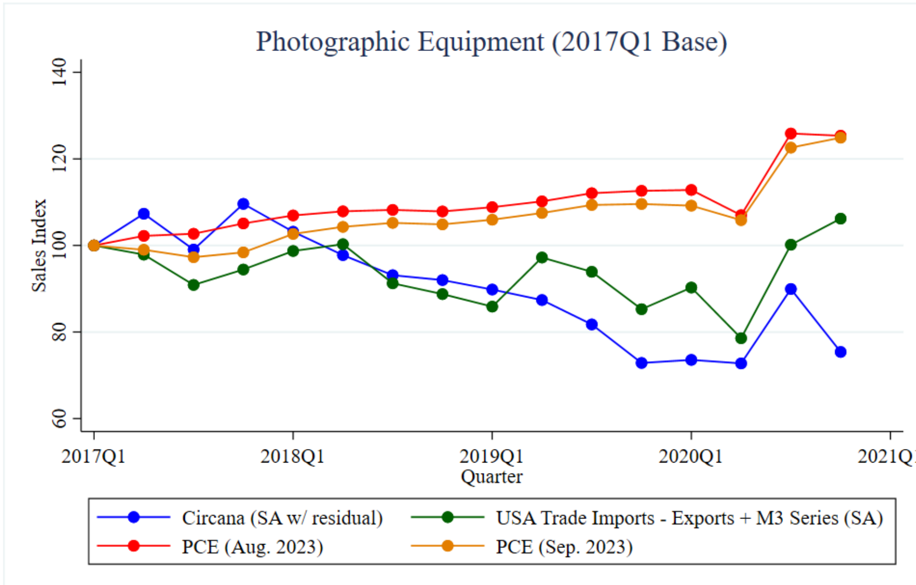
Cereal



Beer

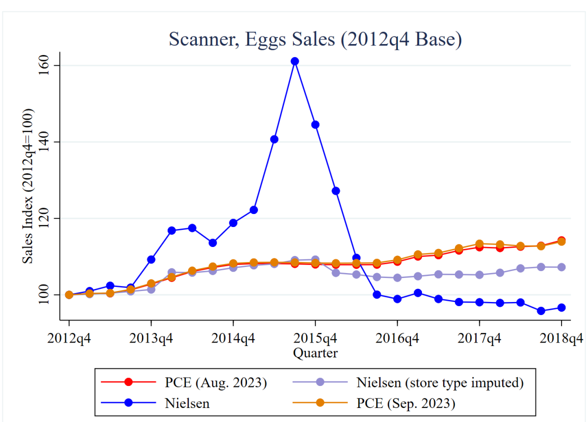
Note: These figures represent quarterly nominal sales indices from 2006q1 to 2019q3 using NielsenIQ data in comparison to official PCE measures for two disaggregated PCE food categories: cereal and beer. In addition to the POS and official PCE series, we provide external sources for nominal sales. We provide external sources from General Mills (2010q1 to 2019q3) and Lux (2006q1 to 2016q4) for cereal nominal sales. We provide external sources from the Beer Institute (2014-2018) and Tax and Trade Bureau (TTB) (2012q1 - 2019q3) for beer nominal sales. We use the PCE price index for beer to convert the TTB series from quantities produced to sales indices. The panel on the left has nominal sales for cereal. The panel on the right has nominal sales for beer. Each panel has four series: two series for PCE nominal sales, one series for NielsenIQ sales, and one external series from the USDA. There are two versions for the official PCE series to reflect a version prior to the 2023 comprehensive update and a version after the update.

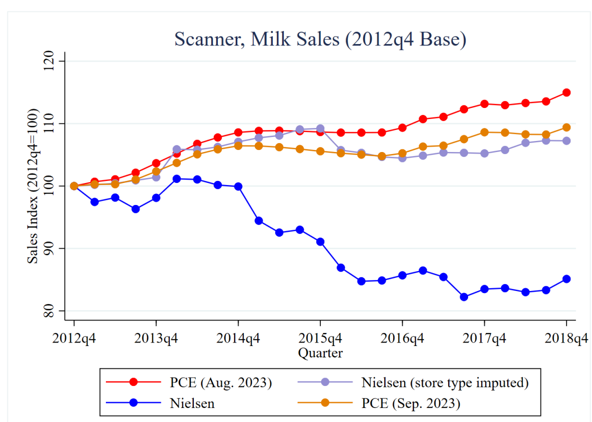
Figure 4.11: External Validity: Nominal Sales Index (Photo)



Note: This figure represents nominal sales indices from 2017q1 to 2020q4 using Circana data in comparison to official PCE measures for a disaggregated tech category: photographic equipment. The panel on the left has PCE official price index in addition to 3 different price indices computed using Circana data. There are four series: two series for PCE nominal sales, one series for NielsenIQ sales, and one external series from the M3 survey. There are two versions for official PCE series to reflect a version prior to the 2023 comprehensive update and a version after the update.

Figure 4.12: Counterfactual Exercise: Nominal Sales Index (Food)

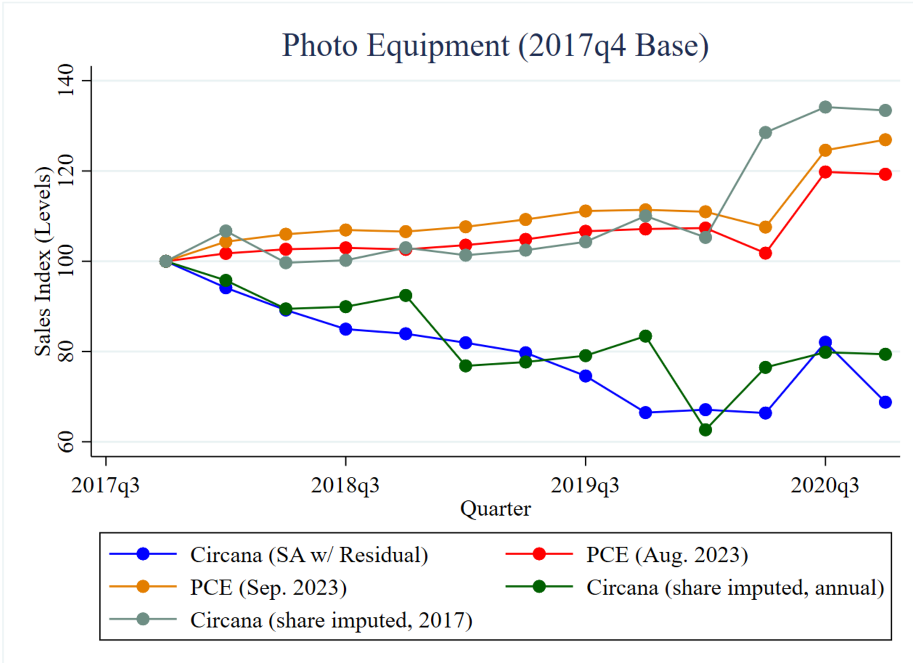




Eggs Milk

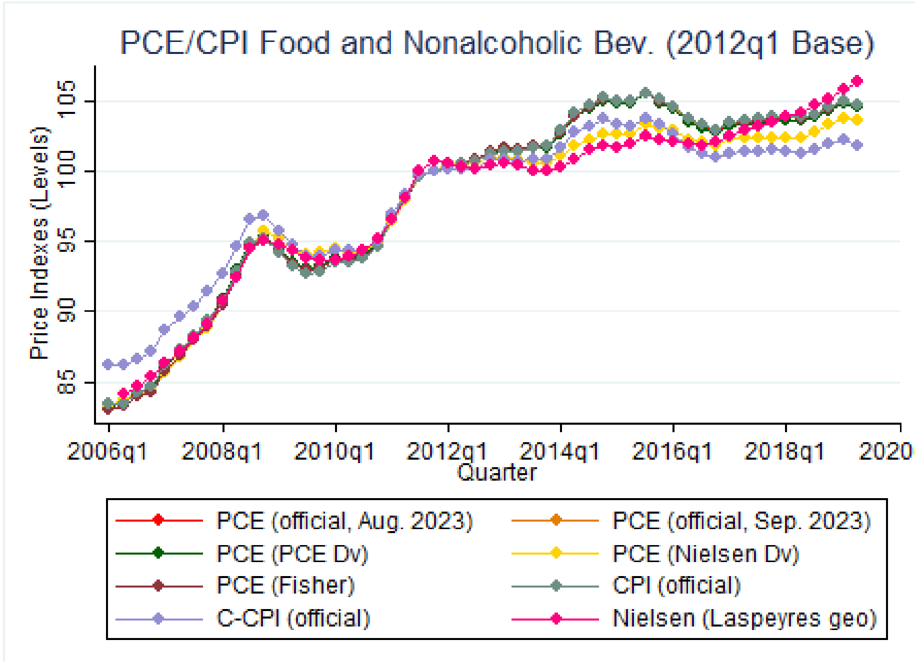
Note: These figures represent nominal sales indices from 2012q4 to 2018q4 using NielsenIQ data in comparison to official PCE measures for two disaggregated PCE food categories: eggs and milk. The panel on the left has nominal sales for eggs. The panel on the right has nominal sales from milk. Each panel has four series: two series for PCE nominal sales and two series for NielsenIQ sales. There are two versions for official PCE series to reflect a version prior to the 2023 comprehensive update and a version after the update. There are two versions of NielsenIQ to reflect a version created using the BEA approach and the other reflects our standard approach.

Figure 4.13: Counterfactual Exercise: Nominal Sales Index (Photo)



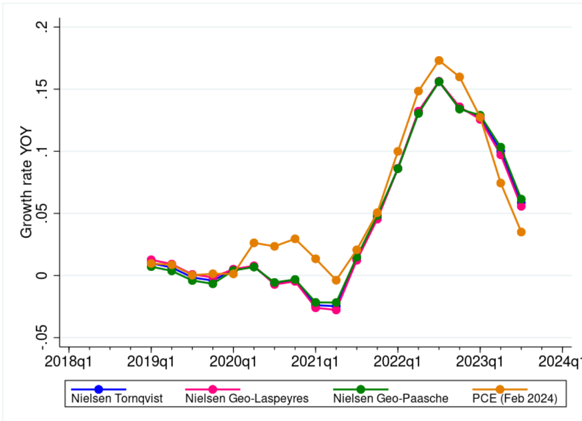
Note: This figure represents nominal sales indices from 2017q4 to 2020q4 using Circana data in comparison to official PCE measures for a disaggregated PCE food categories: photographic equipment. There are five series: two series for PCE nominal sales and three series for NielsenIQ sales. There are two versions for official PCE series to reflect a version prior to the 2023 comprehensive update and a version after the update. There are three versions of NielsenIQ to reflect two versions approximating the BEA approach and the other reflects our standard approach. The two approximations are created using a simple annual average of the photographic equipment share of aggregated tech and the other is the initial 2017 share of aggregated tech.

Figure 4.14: Counterfactual Aggregation Methods (Food)

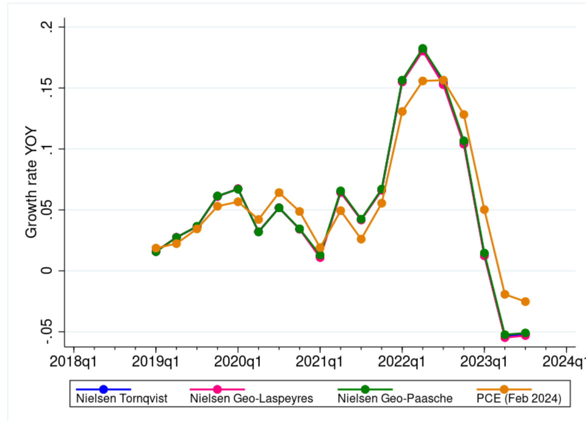


Note: This figure represents price indices for Food and Nonalcoholic beverages from 2006q2 to 2019q3. There are five standard series: (1) PCE (Aug. 2023), (2) PCE (Sep. 2023), (3) CPI (official), (4) C-CPI (official), and (5) NielsenIQ (Laspeyres). CPI corresponds to the BLS Consumer Price Index and C-CPI corresponds to the BLS Chained Consumer Price Index. To emphasize the importance of nominal sales of disaggregated series, we construct three counterfactual series: (1) PCE (PCE Dv), (2) PCE (NielsenIQ Dv), and PCE (Fisher). These counterfactual series take the 18 PCE disaggregated price indices and uses three different aggregation method or weight. Specifically PCE (PCE DV) uses PCE nominal sales as divisia weights, PCE (NielsenIQ Dv) uses NielsenIQ nominal sales as divisia weights, and PCE (Fisher) uses PCE nominal sales with a Fisher method of aggregation.

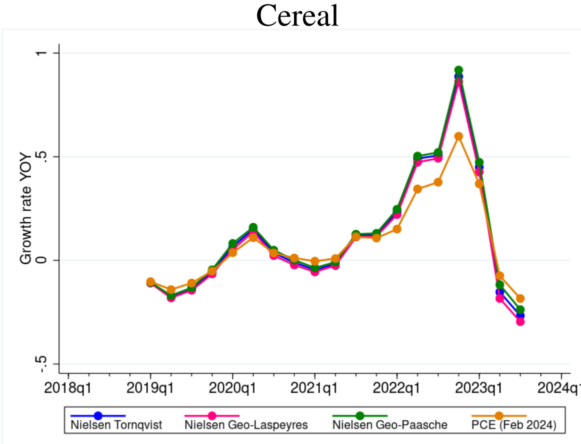
Figure 4.15: YOY Rates of Inflation over the Pandemic Period



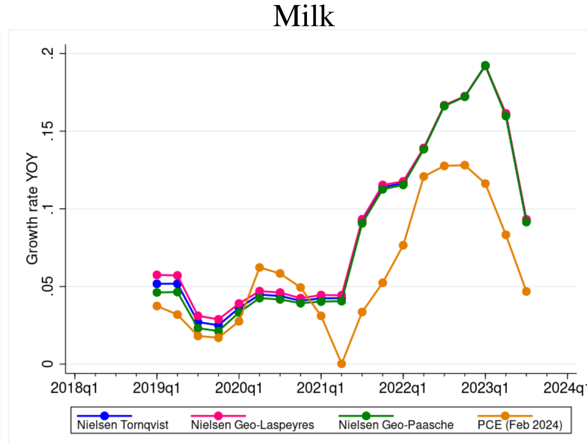
Cereal



Milk



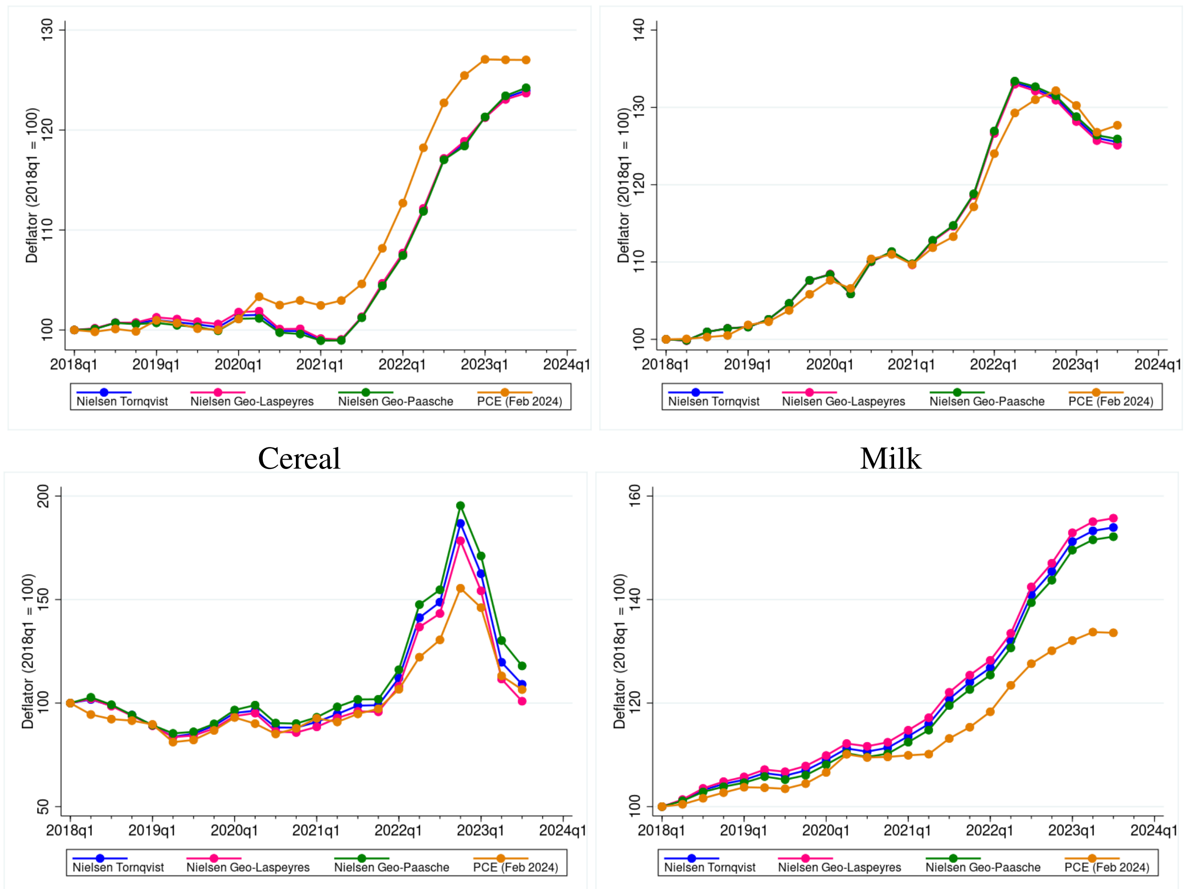
Eggs



Soda

Note: This figure represents year-over-year rates of inflation from 2019q1 to 2023q3 using NielsenIQ data in comparison to official PCE measures for a disaggregated PCE food categories: cereal, milk, eggs, and soda. There are four series: three series for NielsenIQ inflation and one series for PCE inflation. The PCE series is based on the February 2024 version. The three NielsenIQ inflation indices are based on Tornqvist, geometric Laspeyres and Paasche price indices.

Figure 4.16: Price Indices over the Pandemic Period



Cereal

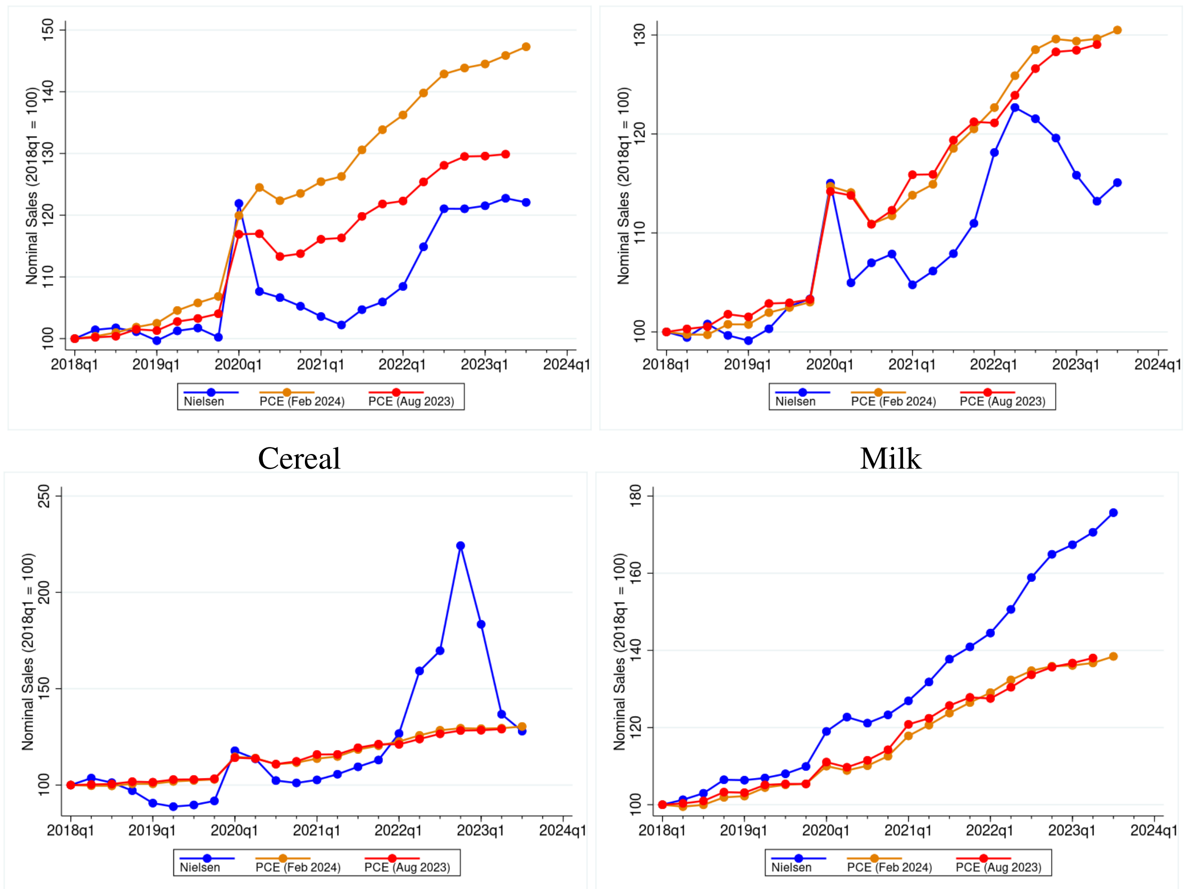
Milk

Eggs

Soda

Note: This figure represents price indices from 2018q1 to 2023q3 using NielsenIQ data in comparison to official PCE measures for a disaggregated PCE food categories: cereal, milk, eggs, and soda. There are four series: three series for NielsenIQ and one series for PCE indices. The PCE series is based on the February 2024 version. The three NielsenIQ price indices are based on Tornqvist, geometric Laspeyres and Paasche price indices.

Figure 4.17: Nominal Sales Indices over the Pandemic Period



Cereal

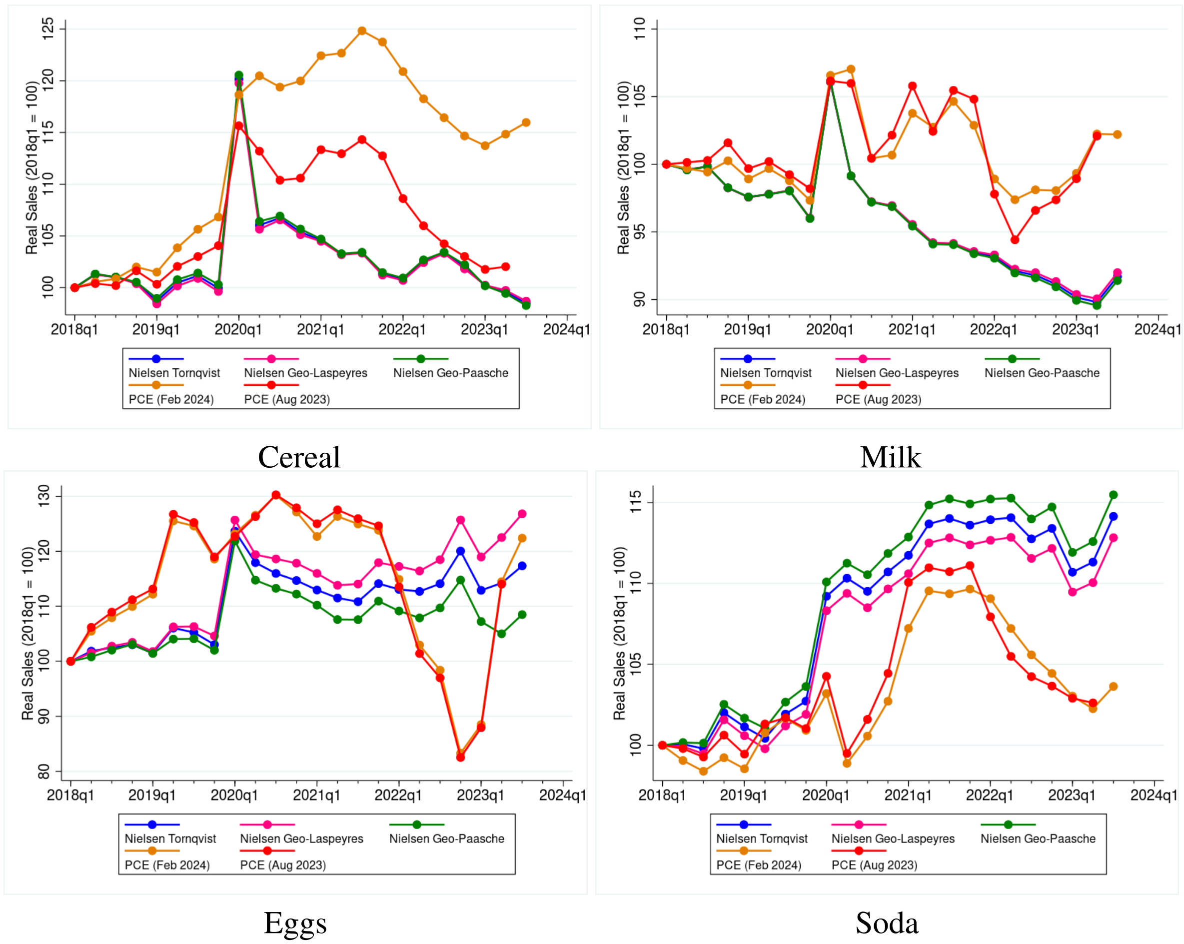
Milk

Eggs

Soda

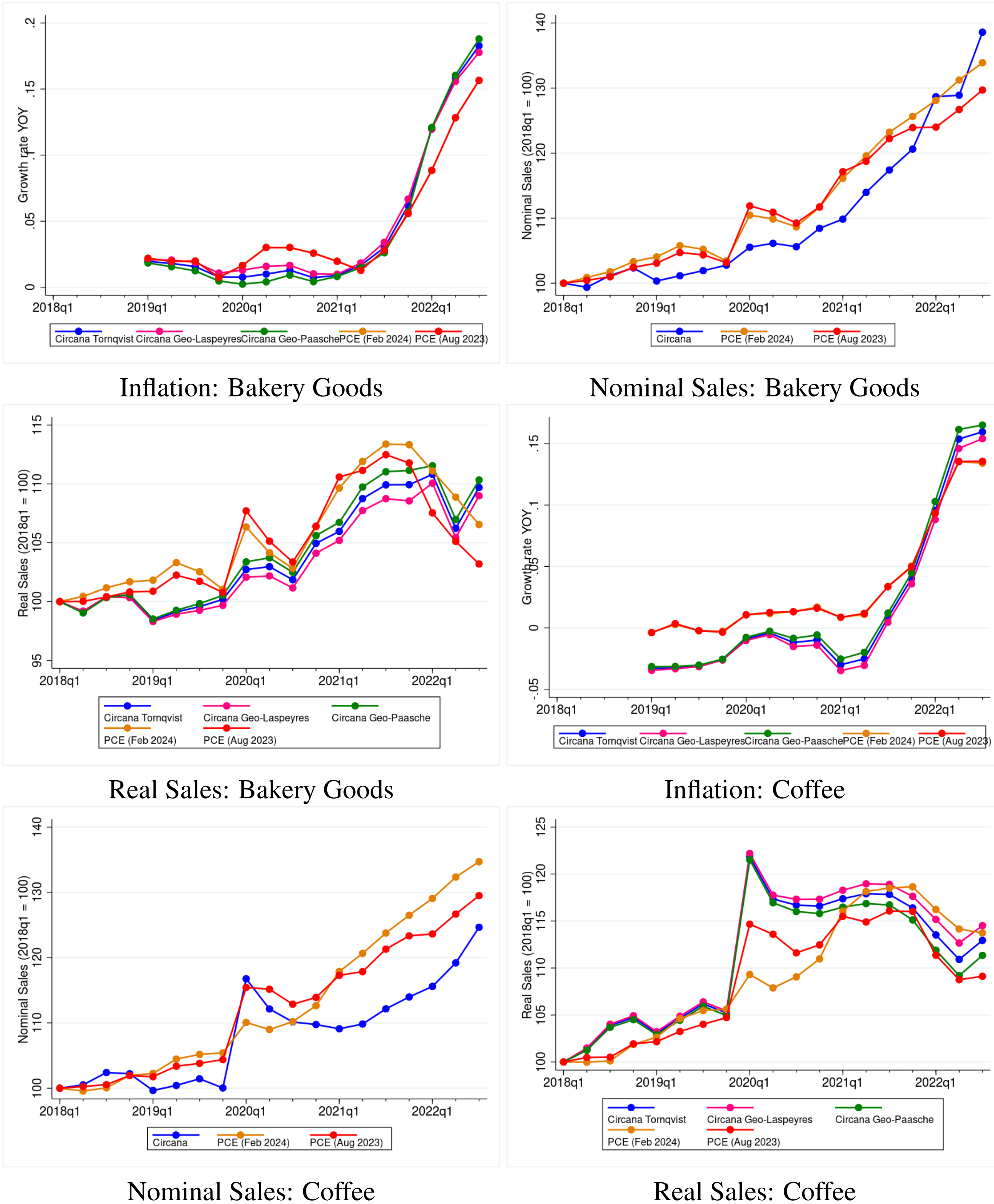
Note: This figure represents nominal sales indices from 2018q1 to 2023q3 using NielsenIQ data in comparison to official PCE measures for a disaggregated PCE food categories: cereal, milk, eggs, and soda. There are three series: one series for NielsenIQ and two series for PCE indices. There are two versions for PCE series based on February 2024 or August 2023.

Figure 4.18: Real Sales Indices over the Pandemic Period



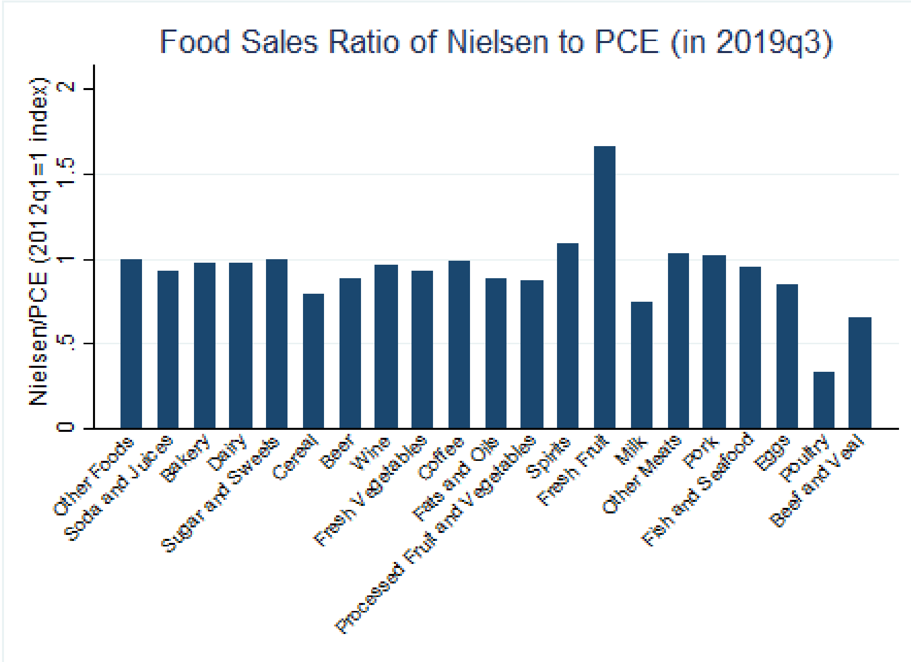
Note: This figure represents price indices from 2018q1 to 2023q3 using NielsenIQ data in comparison to official PCE measures for a disaggregated PCE food categories: cereal, milk, eggs, and soda. There are five series: three series for NielsenIQ and two series for PCE indices with different versions (February 2024 or August 2023). The three NielsenIQ price indices are based on Tornqvist, geometric Laspeyres and Paasche price indices.

Figure 4.19: Price, Nominal, Real Sales Indices over the Pandemic Period (Circana)



Note: This figure represents price, nominal, and real sales indices from 2018q1 to 2022q3 using Circana data in comparison to official PCE measures for a disaggregated PCE food categories: bakery goods and coffee. There are five series for each food item: three series for NielsenIQ and two series for PCE indices with different versions (February 2024 or August 2023). The three NielsenIQ price indices are based on Tornqvist, geometric Laspeyres and Paasche price indices.

Figure 4.20: Sales Ratio of NielsenIQ to PCE



Note: This figure represents the ratio of nominal sales of the NielsenIQ Retail Scanner Panel to the PCE for each disaggregated PCE food category. The measure is calculated in 2019q3, divided by the parallel ratio calculated in 2012q1. This bar indicates how faster the sales grows in NielsenIQ is relative to the PCE over this time period. (e.g., the values equal to one mean the sales in NielsenIQ grew at the same rate as the PCE over 2012q1-2019q3; the values less (greater) than one indicate the slower (faster) growth rates in NielsenIQ compared to PCE.)

B Tables

Table 4.1: Concordance between Economic Census (EC) and PCE

EC Category	PCE category
Eggs and Dairy (except ice cream)	Eggs
Eggs and Dairy (except ice cream)	Dairy
Eggs and Dairy (except ice cream)	Milk
Food Dry Goods and Other Foods	Cereals
Food Dry Goods and Other Foods	Other Foods
Food Dry Goods and Other Foods	Fats and Oils
Food Dry Goods and Other Foods	Processed Fruits and Vegetables
TVs and TV Equipment	TVs
Computers and Peripheral Equipment	PCs/Tablets and Peripheral Equipment

Note: This table provides a mapping of Economic Census (EC) categories to PCE categories. EC categories are more aggregated than PCE categories in general. This is particularly true for food categories.

Table 4.2: NielsenIQ vs PCE (food, nominal sales)

Item	Ratio (NielsenIQ/PCE)	Corr	Corr (trend)	Corr (detrended)	Ratio (SD)	Ratio (detrended SD)
Cereal	0.7266	0.4317	0.4375	0.7202	0.3816	1.0511
Coffee and Tea	0.7981	0.9698	0.9835	0.2415	0.6903	1.8146
Dairy	0.9358	0.9808	0.9873	0.7772	0.7835	1.0230
Eggs	0.7267	0.6712	0.8265	0.5878	1.4158	4.2136
Milk	0.6870	-0.3471	-0.9827	0.8823	1.5227	1.7611
Other Foods	0.8726	0.9877	0.9923	0.5576	0.6770	1.1679
Soda and Juices	0.9449	0.8903	0.9134	0.7514	0.8908	1.2597
Food and Beverages	0.9038	0.9700	0.9782	0.7191	0.7393	1.4015

Note: This table presents a set of summary statistics comparing NielsenIQ and PCE nominal sales indices normalized at 2008q1. For each food item, “Ratio (NielsenIQ/PCE)” indicates the ratio of NielsenIQ to PCE indices evaluated at 2019q3 (same as the numbers shown in Figure 4.20), “Corr” refers to the quarterly correlation between NielsenIQ and PCE series, “Corr (trend or detrended)” shows the quarterly correlation between the trends or detrended series of NielsenIQ and PCE, respectively,“Ratio (SD)” denotes the ratio of the quarterly standard deviation of NielsenIQ to that of PCE, and “Ratio (detrended SD)” represents the ratio of the quarterly standard deviation of the detrended NielsenIQ series to that of PCE.

Table 4.3: NielsenIQ vs PCE (food, price)

Item	Ratio (NielsenIQ/PCE)	Corr	Corr (trend)	Corr (detrended)	Ratio (SD)	Ratio (detrended SD)
Cereal	1.0653	0.9189	0.9291	0.8962	0.9844	0.8096
Coffee and Tea	0.9941	0.9757	0.9804	0.9526	1.4057	1.4161
Dairy	1.0692	0.9619	0.9750	0.9456	1.1471	0.8463
Eggs	1.2416	0.8409	0.8777	0.8987	0.7351	0.4888
Milk	1.1173	0.6841	0.5828	0.9508	1.1444	0.7475
Other Foods	1.0689	0.9764	0.9821	0.8966	1.2727	0.9420
Soda and Juices	1.1166	0.9283	0.9592	0.8289	1.8718	0.9494
Food and Beverages	1.0186	0.9850	0.9938	0.8753	0.9639	1.0262

Note: This table presents a set of summary statistics comparing NielsenIQ and PCE geometric Laspeyres indices normalized at 2008q1. For each food item, “Ratio (NielsenIQ/PCE)” indicates the ratio of NielsenIQ to PCE indices evaluated at 2019q3 (the counterpart to Figure 4.20 for price), “Corr” refers to the quarterly correlation between NielsenIQ and PCE series, “Corr (trend or detrended)” shows the quarterly correlation between the trends or detrended series of NielsenIQ and PCE, respectively,“Ratio (SD)” denotes the ratio of the quarterly standard deviation of NielsenIQ to that of PCE, and “Ratio (detrended SD)” represents the ratio of the quarterly standard deviation of the detrended NielsenIQ series to that of PCE.

Table 4.4: NielsenIQ vs PCE (food, real sales)

Item	Ratio (NielsenIQ/PCE)	Corr	Corr (trend)	Corr (detrended)	Ratio (SD)	Ratio (detrended SD)
Cereal	0.6821	-0.9281	-0.9831	0.0754	0.6835	0.8835
Coffee and Tea	0.8028	0.9768	0.9945	0.0926	0.6383	1.2137
Dairy	0.8752	0.9074	0.9377	0.0798	0.4854	0.5905
Eggs	0.5854	-0.6648	-0.3902	-0.8235	0.7337	0.7071
Milk	0.6149	-0.5294	-0.5910	-0.6470	1.9350	0.7568
Other Foods	0.8164	0.9210	0.9882	0.5163	0.2630	0.9876
Soda and Juices	0.8462	0.2846	0.2190	0.4845	0.3905	1.0266
Food and Beverages	0.8873	0.9145	0.9519	0.3195	0.4989	1.1712

Note: This table presents a set of summary statistics comparing NielsenIQ and PCE real sales indices (based on geometric Laspeyres) normalized at 2008q1. For each food item, "Ratio (NielsenIQ/PCE)" indicates the ratio of NielsenIQ to PCE indices evaluated at 2019q3 (the counterpart to Figure 4.20 for real sales), "Corr" refers to the quarterly correlation between NielsenIQ and PCE series, "Corr (trend or detrended)" shows the quarterly correlation between the trends or detrended series of NielsenIQ and PCE, respectively,"Ratio (SD)" denotes the ratio of the quarterly standard deviation of NielsenIQ to that of PCE, and "Ratio (detrended SD)" represents the ratio of the quarterly standard deviation of the detrended NielsenIQ series to that of PCE.

Table 4.5: Comparison of YOY Rates of Inflation (NielsenIQ vs PCE)

	Cereal		Milk		Eggs		Sodas	
Year	PCE	NielsenIQ	PCE	NielsenIQ	PCE	NielsenIQ	PCE	NielsenIQ
2019	0.5%	0.5%	3.2%	3.5%	-10.1%	-12.5%	2.6%	4.4%
2020	2.0%	0.0%	5.3%	4.6%	4.8%	4.8%	4.9%	4.4%
2021	2.0%	0.0%	3.7%	4.6%	5.7%	3.7%	2.9%	7.4%
2022	14.5%	12.8%	14.3%	14.8%	36.8%	51.2%	11.3%	14.9%
2023	7.9%	9.3%	0.2%	-3.2%	3.7%	-1.8%	8.2%	14.9%

Note: This table compares the yearly inflation rates between PCE and NielsenIQ for the years 2019 to 2023.

Chapter A: Appendix to Chapter 2

A Additional Robustness Checks

A.1 Republican Governor and Republican Vote Share

In my preferred specification for my consumption analysis, I have state fixed, day fixed effects, and the 2016 presidential Republican vote share interacted with Post as controls. In Table B.1, instead of using the 2016 presidential Republican vote share as a control, I use the Republican governor. Unlike the 2016 Republican vote share, there are statistically significant differences in Republican governor between COBOL and non-COBOL states. Columns 2 and 3 are significant at the 5 percent level and have a marginally higher point estimate, 3.0 ppt. relative decrease, than the baseline case with the Republican vote share. Columns 4 and 5 that introduce the state characteristics interacted with post noticeably decrease the point estimates to 2.1 ppt. and 2.0 ppt., respectively. These results are significant at the 10 percent level. I prefer using Republican vote share over Republican governor given that Republican vote share is a continuous variable while Republican governor is a binary variable.

However, one could view the Republican vote share and Republican governor as picking up different sources of variation. Republican vote share could capture COVID-19 cautiousness, while the Republican governor could capture different policies implemented during the pandemic. In table B.2 I use the Republican governor and Republican vote share as controls simultaneously. My results are robust across all specifications. The point estimates range from a 2.1 ppt decline to a 2.6 ppt. decline in specifications with controls other than fixed effects. All specifications are all significant at the 10 percent level and have have similar point estimates to the version without Republican governor. These new point estimates are slightly lower than the baseline, but are not statistically different than the baseline.

A.2 Randomization Inference

As another robustness check, I perform randomization inference where I conduct a permutation test. Specifically, I randomly assign COBOL status to 28 states 1,000 times and then run my preferred specification, column 2 of Table 2.2, with this random assignment of treatment. The specification has relative card consumption as the dependent variable, day fixed effects, state fixed effects, and 2016 Republican vote share interacted with Post. The treatment is the product of Post and the simulated COBOL variable. Post is a binary variable that takes the value of 1 after March 13 independent of the iteration. The simulated COBOL variable assigns COBOL status to 28 out of the 50 states and which 28 states are selected varies depending on the iteration. Similar to my regression results, I applied population weights. Figure B.1 shows that the permutation test yields significant results at the 10 percent level. The other specifications in Table 2.2 are all significant at the 10 percent level.[1]

A.3 Other Pandemic Transfer Programs

One concern one may have is that COBOL usage in UI benefit systems is correlated with other pandemic transfer programs, and that these transfer programs may well have had an independent effect on consumption. I focus on three pandemic transfer programs that were in effect in 2020: Paycheck Protection Plan (PPP), Economic Impact Payments (EIP), and Supplemental Nutrition Assistance Program (SNAP). As Faulkender et al. (2023) show PPP was intricately connected to UI given that PPP was partially implemented to help alleviate the stress that UI benefit systems were undergoing during the pandemic recession. SNAP and EIP were large pandemic transfer programs that were income based where high income earners were less likely to be eligible or at least eligible for lower amounts. EIP and PPP were federal programs, while SNAP

[1] Available upon request.

is a state administered program. The data for the PPP loans is provided by the U.S. Small Business Administration.[2] The data on the EIP is provided by the Internal Revenus Service's (IRS) Statistics of Income (SOI) program.[3] The data for SNAP is provided by the U.S. Department of Agriculture (USDA).[4]

Table B.3 shows summary statistics for COBOL and non-COBOL states for the following PPP measures, reported in per capita and per worker terms: initial loan amount, jobs reported, and total loans after dividing by either the 2019 labor force or by the 2019 population. These summary statistics correspond to PPP loans with approval dates before January 1, 2021. Specifically, COBOL states received $1,635.65 per capita in PPP funds, while non-COBOL states received $1,535.67. As reflected Table B.3, all the PPP outcomes have t-statistics that are statistically insignificant between COBOL and non-COBOL states.

In Table B.4, I created summary statistics for the first round of EIP that was disbursed in April 2020. Similar to PPP, there is no statistically significant difference between COBOL and non-COBOL states regardless of which of the two normalizations that I apply: (1) labor force (LF) or (2) population (Pop). Specifically, COBOL states received $851.76 per capita of EIP funds, while non-COBOL states received $841.79.

In Table B.5, I created summary statistics for SNAP benefits that were disbursed between March 2020 and December 2020. Similar to PPP and EIP, there is no statistically significant difference between COBOL and non-COBOL states regardless of which of the two normalizations that I apply: (1) labor force (LF) or (2) population (Pop). Specifically, COBOL states received $216.89 per capita in SNAP funds, while non-COBOL states received $219.03.

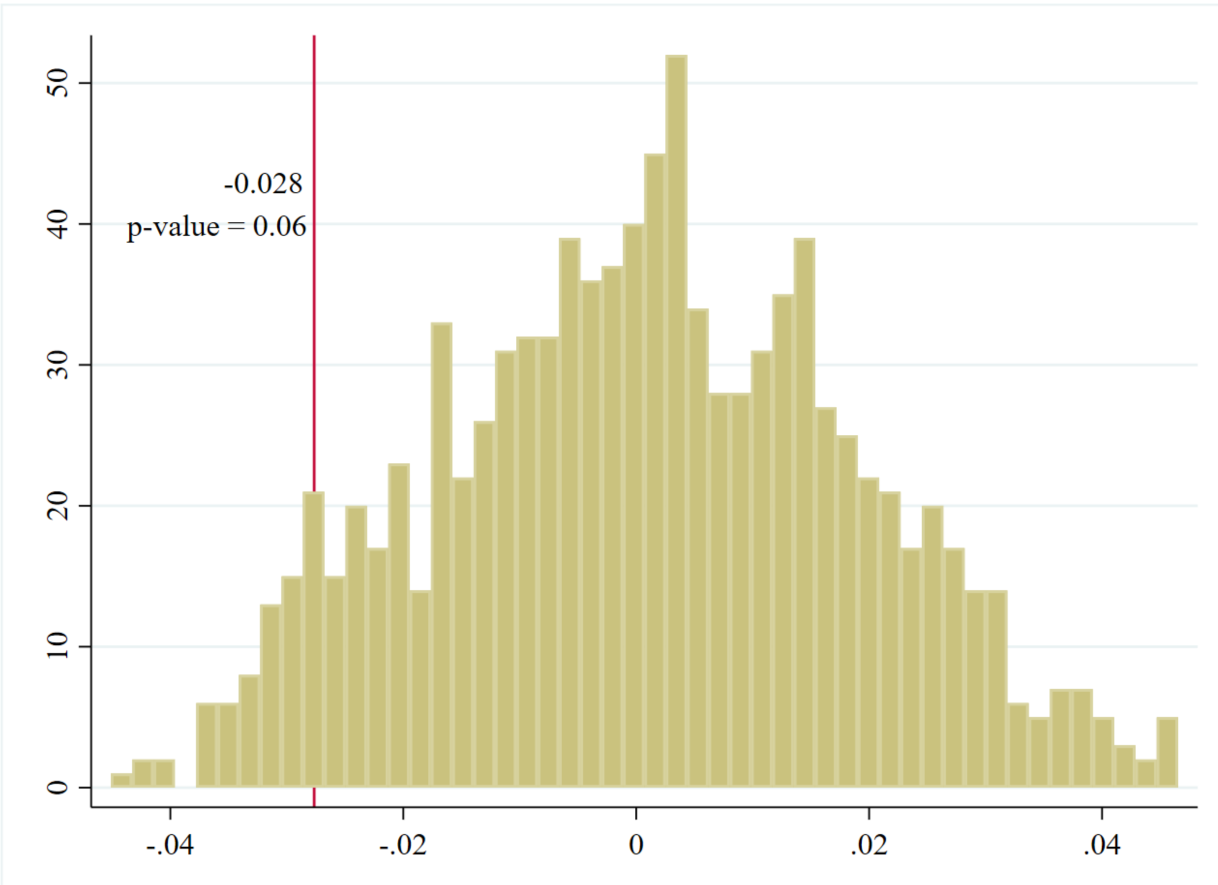
Overall, the analysis in Tables B.3, B.4, and B.5 demonstrates that there were not systematic differences in the disbursement of benefits between COBOL and non-COBOL states in regards to SNAP, EIP, and PPP. A full analysis of the impact of COBOL-induced UI issues on consumption with the inclusion of the transfer program benefit amounts would require a complex

[2]The data can be accessed from the SBA website.
[3]The data can be accessed from the IRS SOI website.
[4]The data can be accessed from the USDA website.

analysis across states and timing of disbursement. For example, Chetty et al. (2020) look at EIP affecting consumption only for 31 days starting from April 15, 2020. This complex analysis goes beyond the scope of this dissertation chapter.

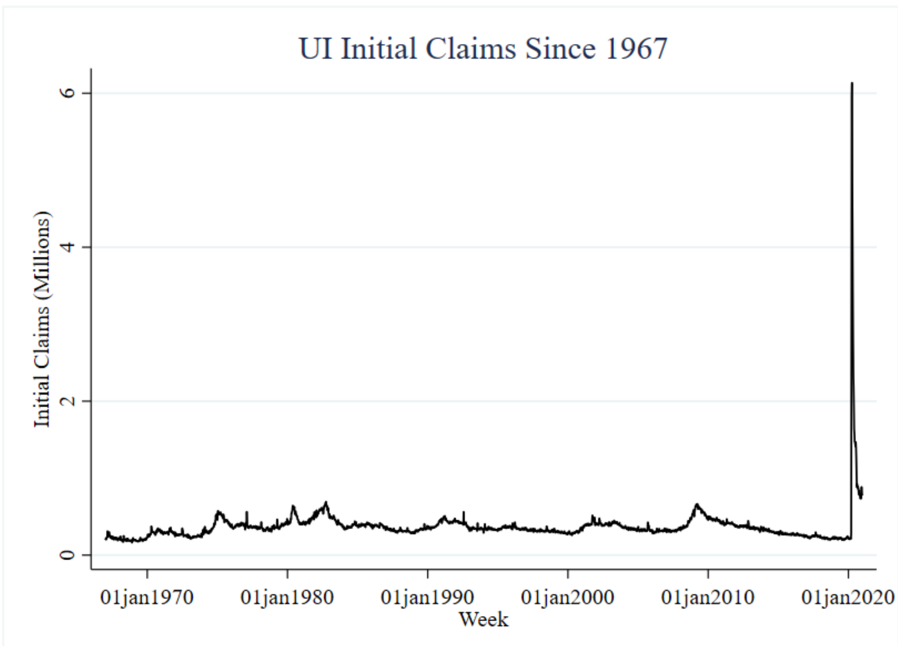
B Additional Tables and Figures

Figure B.1: Permutation Test for TWFE Estimator (1,000 Simulations)



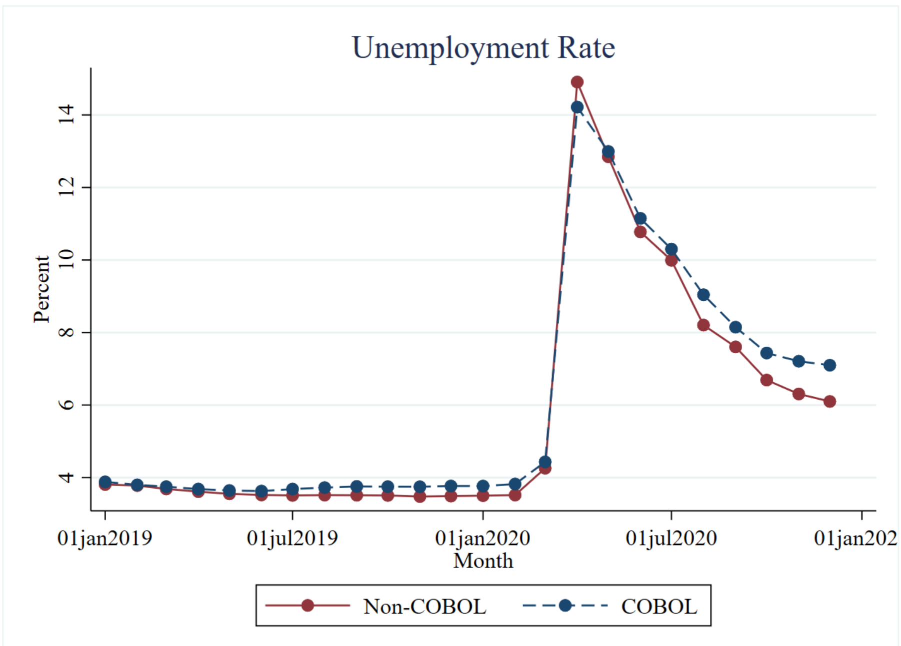
Note: The histogram shows the distribution of average treatment effects when treatment is randomly assigned across 28 of the 50 states 1,000 times using the TWFE specification with the 2016 Republican vote share interacted with Post. To be consistent, I apply population weights as analytic weights. The red line corresponds to actual treatment effect with the 28 COBOL states: a 2.8-percentage-point decline. This permutation test yields an effect that is significant at the 10% level.

Figure B.2: National UI Initial Claims



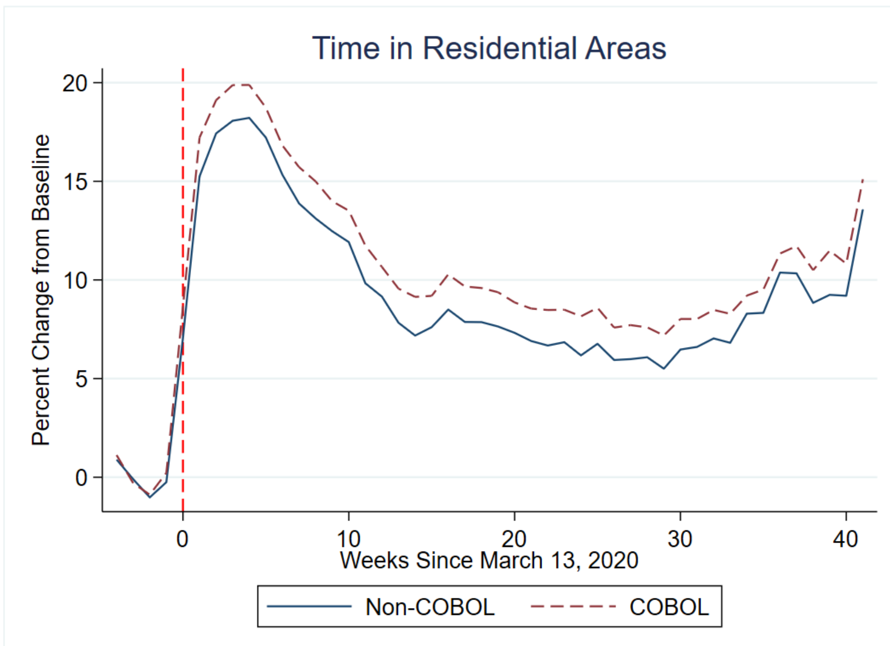
Note: This figure uses weekly initial claims data from 1967 to the end of 2020. The data used are from the Department of Labor Employment Training Administration (DOLETA). The peak in initial claims corresponds to the first week of April 2020 where there was just north of 6 million initial claims filed that week at the national level.

Figure B.3: Unemployment Rate by COBOL Usage



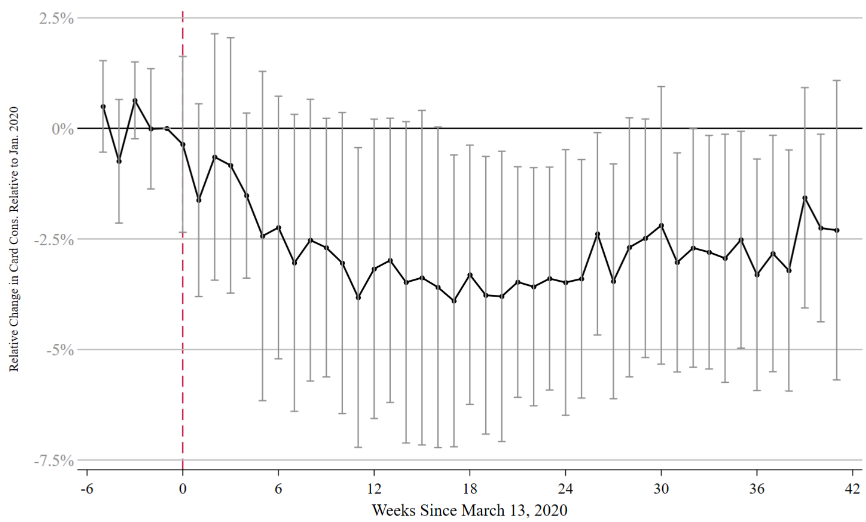
Note: This figure uses monthly seasonally adjusted state unemployment rate from the BLS. The data range from January 2019 to December 2020. The data is relative time spent in residential area relative to that's baseline period of time in residential areas. The baseline period corresponds to the first six weeks of 2020. The maroon line corresponds to non-COBOL states and the navy line corresponds to COBOL states. These two groups of states are aggregated using 2019 population weights.

Figure B.4: Potential COVID-19 Cautiousness



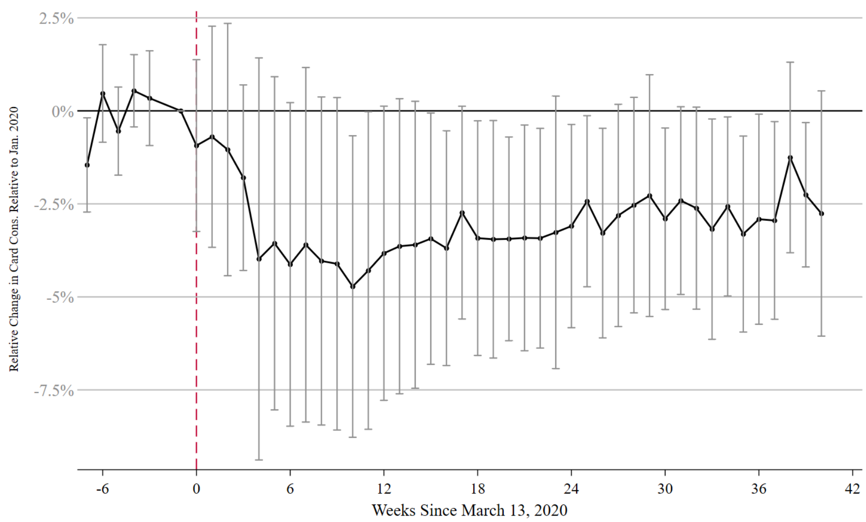
Note: This figure uses daily Google Mobility data from February 2020 to December 2020. I aggregate the data to the weekly frequency. Weeks a are determined relative to the week ending on March 13, 2020, which corresponds to the red vertical dashed line. The maroon line corresponds to non-COBOL states and the navy line corresponds to COBOL states. These two groups of states are aggregated using 2019 population weights.

Figure B.5: Relative-Consumption Weekly Event Study:
Relative Difference between COBOL and Non-COBOL States with Dynamic Rep. Share



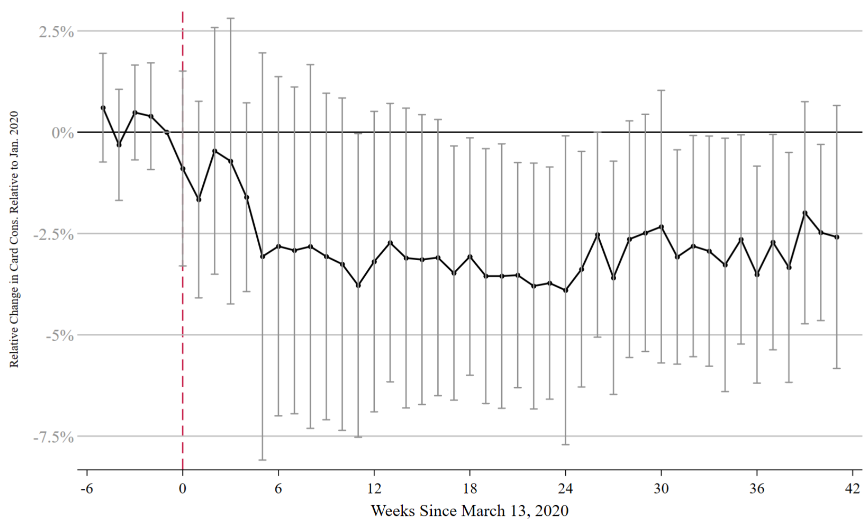
Note: The graph is a coefficient plot showing the coefficient on β_k from Equation 2.3. State and week fixed effects are included as well as COVID-19 controls. However, instead of using 2016 Republican vote share interacted with $Post_t$, I use the 2016 Republican vote share interacted with I_k. The red dashed line that goes through week zero corresponds to March 13, 2020. This figure shows that in each week after the week of the emergency declaration, COBOL states saw a larger decline in relative consumption than non-COBOL states even when dynamically interacting relative time spent at home. Relative time spent at home should be interpreted as a proxy for COVID-19 cautiousness. The bounds on each point estimate correspond to a 95% percent confidence interval. State populations from 2019 are applied as analytic weights. Standard errors are clustered at the state level.

Figure B.6: Relative-Consumption Weekly Event Study:
Relative Difference between COBOL and Non-COBOL States with Google Mobility (Control)



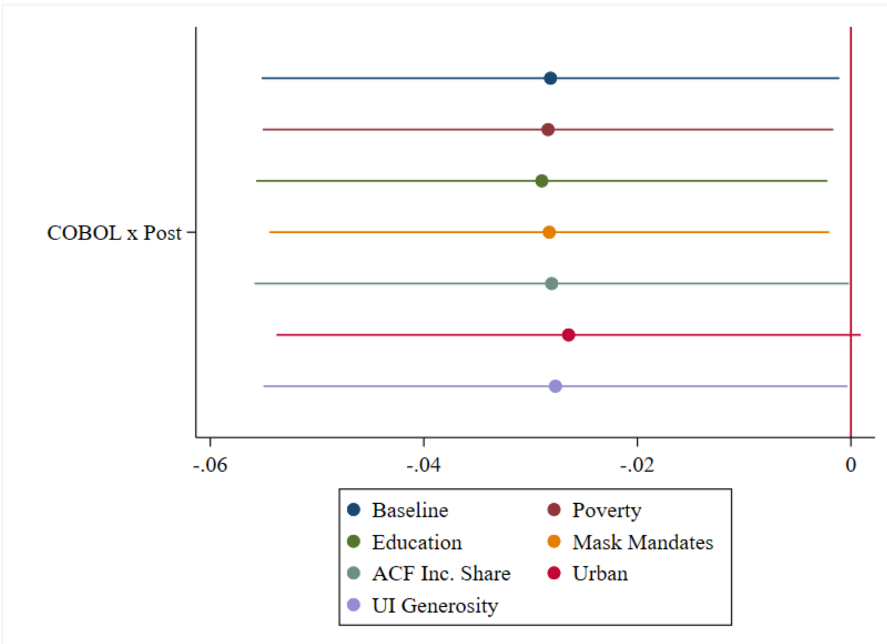
Note: The graph is a coefficient plot showing the coefficient on β_k from Equation 2.3. State and week fixed effects are included as well as COVID-19 controls. However, there instead of controlling for the 2016 Republican vote share, I control for relative time at home in state i in week k using Google Mobility data. The red dashed line that goes through week zero corresponds to March 13, 2020. This figure shows that in each week after the week of the emergency declaration, COBOL states saw a larger decline in relative consumption than non-COBOL states even when dynamically interacting relative time spent at home. Relative time spent at home should be interpreted as a proxy for COVID-19 cautiousness. The bounds on each point estimate correspond to a 95% percent confidence interval. State populations from 2019 are applied as analytic weights. Standard errors are clustered at the state level.

Figure B.7: Relative-Consumption Weekly Event Study:
Relative Difference between COBOL and Non-COBOL States with Google Mobility (Interacted Control)



Note: The graph is a coefficient plot showing the coefficient on β_k from Equation 2.3. State and week fixed effects are included as well as COVID-19 controls. However, there instead of controlling for the 2016 Republican vote share, I control for relative time at home in state i in week k and interact it with I_k. The red dashed line that goes through week zero corresponds to March 13, 2020. This figure shows that in each week after the week of the emergency declaration, COBOL states saw a larger decline in relative consumption than non-COBOL states even when dynamically interacting relative time spent at home. Relative time spent at home should be interpreted as a proxy for COVID-19 cautiousness. The bounds on each point estimate correspond to a 95% percent confidence interval. State populations from 2019 are applied as analytic weights. Standard errors are clustered at the state level.

Figure B.8: Coefficient Plot Interacting Potential Confounders Individually



Note: This figure plots the coefficient on $COBOL \times Post$ from the TWFE estimator (state fixed effects and day fixed effects). The baseline definition includes the COVID-19 controls and the interaction of COBOL state and the 2016 Republican presidential vote share. The remaining five coefficients build upon the baseline specification by adding one confounder, $Confounder_i$, and interacting it with $Post_t$. The figure plots coefficients on $COBOL \times Post$ ranging from a 2.6 percentage point decline in relative consumption to a 2.9 percentage point decline. The effect is significant at the 5 percent level in all specifications except for the one that adds percentage of the population living in an urban area (still significant at the 10 percent level). Standard errors are clustered at the state level.

Table B.1: TWFE COBOL Usage on All Card Consumption (Republican Governor, Replacement)

	(1) Rel Cons	(2) Rel Cons	(3) Rel Cons	(4) Rel Cons	(5) Rel Cons
COBOL × Post	-0.041**	-0.030**	-0.030**	-0.021*	-0.020*
	[0.020]	[0.015]	[0.015]	[0.011]	[0.011]
RepGov × Post		0.037*	0.038*	0.031	0.032
		[0.021]	[0.021]	[0.020]	[0.019]
UR					0.002
					[0.002]
State FE	Yes	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes	Yes
COVID-19 Controls	No	No	Yes	Yes	Yes
State Char. × Post	No	No	No	Yes	Yes
Days	335	335	335	335	335
States	50	50	50	50	50
Observations	16,750	16,750	16,750	16,750	16,750

Note: The table presents results from a TWFE estimator with day and state fixed effects. The dependent variable is the percentage-point change relative to the base period in credit and debit card consumption measured at the daily frequency. *Post* is a binary variable that takes the value 1 if the date is on or after March 13, 2020. *COBOL* is a binary variable that takes the value 1 if a state uses COBOL in its UI benefits system. The interaction term of interest is the product of *COBOL* and *Post*. *RepGov* is a binary variable corresponding to whether a state had a Republican governor in 2019. COVID-19 controls include new COVID-19 death rates and new COVID-19 case rates. Column 1 only includes state and day fixed effects. Column 2 adds the interaction of *Republican* and *Post*. Column 3 adds the COVID-19 controls. Column 4 adds five terms of *Post* interacted with another confounder: (1) income share in accommodation and food services (2019), (2) mask mandates in July 2020 (2020), (3) the percentage of the population living in poverty (2019), (4) the percentage of the population with at least a bachelor's degree (2019), and (5) UI generosity (Jan. 2020). Column 5 adds the monthly unemployment rate. These estimates cover the sample period of February 1, 2020, to December 31, 2020. State populations in 2019 are applied as analytic weights. Standard errors are clustered at the state level.

Table B.2: TWFE COBOL Usage on All Card Consumption (Republican Governor, Inclusion)

	(1) Rel Cons	(2) Rel Cons	(3) Rel Cons	(4) Rel Cons	(5) Rel Cons
COBOL × Post	-0.041**	-0.026*	-0.027*	-0.022**	-0.021*
	[0.020]	[0.014]	[0.014]	[0.011]	[0.011]
Republican Gov. × Post		0.006	0.007	0.018	0.019
		[0.016]	[0.016]	[0.018]	[0.017]
Republican × Post		0.003***	0.003***	0.003**	0.003***
		[0.001]	[0.001]	[0.001]	[0.001]
UR					0.002
					[0.002]
State FE	Yes	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes	Yes
COVID-19 Controls	No	No	Yes	Yes	Yes
State Char. × Post	No	No	No	Yes	Yes
Days	335	335	335	335	335
States	50	50	50	50	50
Observations	16,750	16,750	16,750	16,750	16,750

Note: The table presents results from a TWFE estimator with day and state fixed effects. The dependent variable is the percentage-point change relative to the base period in credit and debit card consumption measured at the daily frequency. *Post* is a binary variable that takes the value 1 if the date is on or after March 13, 2020. *COBOL* is a binary variable that takes the value 1 if a state uses COBOL in its UI benefits system. The interaction term of interest is the product of *COBOL* and *Post*. *Republican* is the Republican vote share in the 2016 presidential election. COVID-19 controls include new COVID-19 death rates and new COVID-19 case rates. Column 1 only includes state and day fixed effects. Column 2 adds the interaction of *Republican* and *Post* as well as interacting a binary variable indicating whether a state had a Republican governor and *Post*. Column 3 adds the COVID-19 controls. Column 4 adds five terms of *Post* interacted with another confounder: (1) income share in accommodation and food services (2019), (2) mask mandates in July 2020 (2020), (3) the percentage of the population living in poverty (2019), (4) the percentage of the population with at least a bachelor's degree (2019), and (5) UI generosity (Jan. 2020). Column 5 adds the monthly unemployment rate. These estimates cover the sample period of February 1, 2020, to December 31, 2020. State populations in 2019 are applied as analytic weights. Standard errors are clustered at the state level.

Standard errors: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table B.3: PPP Summary Statistics by COBOL Status (PPP 2020)

	Non-COBOL Mean	COBOL Mean	Diff	t-stat	Non-COBOL N	COBOL
Loan Amount per LF	3,056.37	3,212.00	-155.64	-1.25	22	28
Jobs Reported per LF	0.36	0.37	-0.00	-0.42	22	28
Total Loans per LF	0.03	0.03	-0.00	-0.03	22	28
Loan Amount per Pop	1,535.67	1,635.65	-99.99	-1.20	22	28
Jobs Reported per Pop	0.18	0.19	-0.01	-0.71	22	28
Total Loans per Pop	0.02	0.02	-0.00	-0.30	22	28

Note: The summary statistics table shows the differences in PPP loan amounts, total loans, and jobs reported between COBOL and non-COBOL states. These statistics are normalized by either the 2019 labor force or the 2019 population. This covers the period of PPP loans disbursed in 2020. These statistics are denominated as the total PPP amount disbursed in 2020 once the program started in April 2020. All statistics are in per capita or per worker terms.

Table B.4: EIP Summary Statistics by COBOL Status (First Round)

	Non-COBOL Mean	COBOL Mean	Diff	t-stat	Non-COBOL N	COBOL N
No. Pay per LF	1.01	0.99	0.02	0.75	22	28
No. Pay per Pop	0.50	0.50	0.00	0.26	22	28
Pay Amount per LF	1,715.88	1,669.02	46.87	1.02	22	28
Pay Amount per Pop	851.76	841.79	9.98	0.77	22	28

Note: The summary statistics table shows the differences in EIP payment amounts and number of EIP recipients. These statistics are either normalized by the 2019 labor force (LF) or the 2019 population (Pop). These statistics only correspond to the first wave of EIP, which occurred in April 2020. These statistics are denominated in total EIP disbursed in 2020 (one observation per state). All statistics are in per capita or per worker terms.

Table B.5: SNAP Summary Statistics by COBOL Status (Mar. 2020 to Dec. 2020)

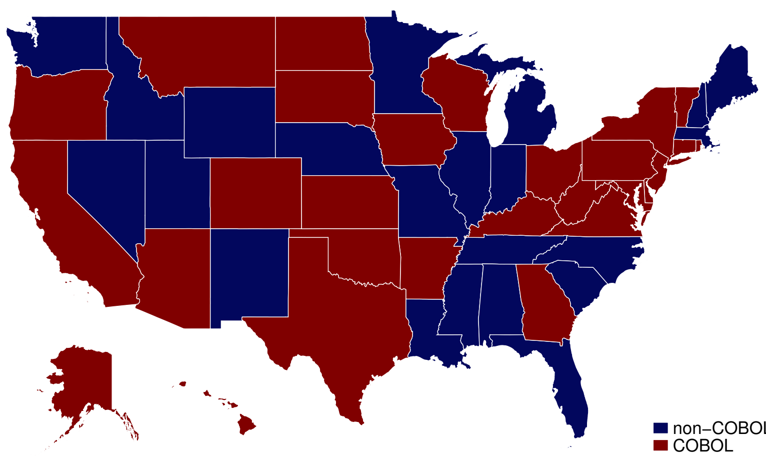
	Non-COBOL Mean	COBOL Mean	Diff	t-stat	Non-COBOL N	COBOL N
Pay Amount per LF	448.33	433.65	14.68	0.31	22	28
Pay Amount per Pop	219.03	216.89	2.14	0.10	22	28
HH Par. per LF	1.26	1.21	0.05	0.39	22	28
HH Par. per Pop	0.62	0.61	0.01	0.17	22	28
Ind. Par. per LF	2.49	2.36	0.13	0.53	22	28
Ind Par. per Pop	1.22	1.18	0.04	0.34	22	28

Note: The summary statistics table shows the differences in SNAP payment amounts, household participation, and individual participation. These statistics are normalized by either the 2019 labor force (LF) or the 2019 population (Pop). These statistics are denominated as the total SNAP amount disbursed from March 2020 to December 2020. All statistics are in per capita or per worker terms.

Chapter B: Appendix to Chapter 3

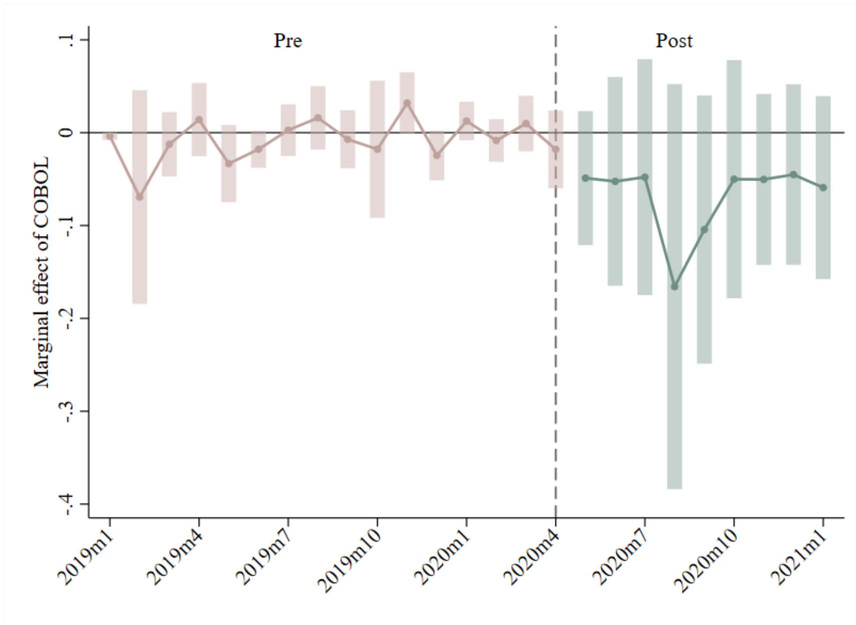
A Additional Tables and Figures

Figure A.1: COBOL vs. non-COBOL States



Notes: The map depicts which states use (versus not using) COBOL in their UI benefit systems in June 2020. The data on COBOL usage are collected in Chapter 2: primarily from emails, news articles, and information from the UI Information Technology Support Center.

Figure A.2: Marginal Effect of COBOL: CSDiD



Notes: This figure depicts the marginal effect of COBOL on log(business applications per capita) using the Callaway & Sant'Anna (2021) estimator with Republican governor and homeownership rate as controls.

Table A.1: Pandemic Related Restriction

	Coefficient on COBOL indicator			
Dependent Variable	**Est.**	**Std Err.**	**P-Value**	**Mean of Dep Var.**
Stay at home reqs.	0.028	(0.058)	(0.625)	0.210
School closings	0.043	(0.032)	(0.181)	0.934
Workplace closings	0.051	(0.087)	(0.558)	0.577
Cancellation of public events	-0.043	(0.082)	(0.597)	0.539
Restrictions on gatherings	-0.027	(0.073)	(0.717)	0.858
Public transport closures	-0.058	(0.068)	(0.395)	0.088

Notes: This table reports results from regressions where pandemic related policies in Column (1) are the dependent variables and the COBOL indicator is the independent variable. The source of all pandemic related policy variables is the Oxford COVID-19 Government Response Tracker (OxCGRT). For each state, OxCGRT tracks, at a daily frequency, whether for each policy category (stay at home requirements, school closings, etc.) there are no measures (0), recommended restrictions (1), or required restrictions of varying degrees of stringency (2, 3, and 4). We take the raw data for each policy category and create a dummy variable, which takes the value of 0 when the raw value is $\{0,1\}$ and takes the value of 1 when the raw value is $\{2,3,4\}$. We then calculate, for each policy category the cross-day average (between March 13 and December 31, 2020) at the state level. The resulting pandemic related policy variables are then regressed on the COBOL indicator. ***, **, and * indicate significance at the 1%, 5%, and 10% significance levels.

Table A.2: Balance of Characteristics: Employment Share By Firm Age, Firm Size, and Sector

Dependent Variable	Coefficient on COBOL indicator Est.	Std Err.	P-Value	Mean of Dep Var.
Firm Age				
0	-0.001	(0.001)	(0.46)	0.02
1-5	0.000	(0.004)	(0.93)	0.08
6-10	-0.004	(0.003)	(0.21)	0.07
11+	0.006	(0.008)	(0.50)	0.83
Firm Size				
1-19	0.003	(0.010)	(0.72)	0.18
20-499	0.009	(0.007)	(0.24)	0.31
500+	-0.012	(0.016)	(0.43)	0.51
Sector				
Agriculture, Forestry, Fishing and Hunting	-0.001	(0.001)	(0.29)	0.00
Mining, Quarrying, and Oil and Gas Extraction	0.001	(0.005)	(0.85)	0.01
Utilities	0.000	(0.001)	(0.52)	0.01
Construction	-0.003	(0.003)	(0.33)	0.06
Manufacturing	-0.010	(0.012)	(0.41)	0.10
Wholesale Trade	0.001	(0.002)	(0.64)	0.05
Retail Trade	-0.002	(0.004)	(0.52)	0.13
Transportation and Warehousing	0.002	(0.003)	(0.44)	0.04
Information	0.001	(0.002)	(0.53)	0.02
Finance and Insurance	0.005	(0.004)	(0.20)	0.05
Real Estate and Rental and Leasing	0.000	(0.001)	(0.95)	0.02
Professional, Scientific, and Technical Services	0.004	(0.006)	(0.49)	0.06
Management of Companies and Enterprises	0.001	(0.003)	(0.79)	0.02
Administrative and Support and Waste Management and Remediation	-0.005	(0.007)	(0.48)	0.08
Educational Services	0.001	(0.004)	(0.74)	0.03
Health Care and Social Assistance	0.009	(0.007)	(0.23)	0.17
Arts, Entertainment, and Recreation	0.001	(0.001)	(0.68)	0.02
Accommodation and Food Services	-0.006	(0.008)	(0.45)	0.12
Other Services (except Public Admin.)	0.001	(0.001)	(0.60)	0.04

Notes: This table reports results from regressions where each one of the state-level variables in Column (1) are dependent variables and the COBOL indicator is the independent variable. Employment share at each firm age, size, and sector category is 2019 values obtained from the Business Dynamics Statistics of the Census Bureau.

Table A.3: Balance of Characteristics: Employment Share By Occupation

Dependent Variable	Coefficient on COBOL indicator			
	Est.	**Std Err.**	**P-Value**	**Mean of Dep Var.**
Management	0.004	(0.003)	(0.28)	0.10
Business and Financial Operations	0.003	(0.002)	(0.20)	0.05
Computer and Mathematical	0.001	(0.003)	(0.70)	0.03
Architecture and Engineering	-0.001	(0.001)	(0.64)	0.02
Life, Physical, and Social Science	0.001	(0.001)	(0.51)	0.01
Community and Social Service	0.001	(0.001)	(0.12)	0.02
Legal	0.001	(0.001)	(0.26)	0.01
Educational Instruction and Library	0.001	(0.002)	(0.60)	0.06
Arts, Design, Entertainment, Sports, and Media	0.001	(0.001)	(0.42)	0.02
Healthcare Practitioners and Technical	0.001	(0.002)	(0.78)	0.06
Healthcare Support	0.001	(0.001)	(0.43)	0.03
Protective Service	0.001	(0.001)	(0.36)	0.02
Food Preparation and Serving Related	-0.004	(0.002)	(0.12)	0.06
Building and Grounds Cleaning and Maintenance	0.000	(0.002)	(0.93)	0.04
Personal Care and Service	0.000	(0.001)	(0.87)	0.03
Sales and Related	-0.002	(0.002)	(0.41)	0.10
Office and Administrative Support	0.000	(0.002)	(0.90)	0.11
Farming, Fishing, and Forestry	0.000	(0.002)	(0.82)	0.01
Construction and Extraction	-0.002	(0.003)	(0.50)	0.05
Installation, Maintenance, and Repair	-0.001	(0.001)	(0.48)	0.03
Production	-0.008	(0.006)	(0.16)	0.06
Transportation	-0.001	(0.003)	(0.71)	0.08
Material Moving	0.002	(0.001)	(0.21)	0.00

Notes: This table reports results from regressions where each one of the state-level variables in Column (1) are dependent variables and the COBOL indicator is the independent variable. Employment shares at each occupation category is obtained from the 2019 American Community Survey.

Table A.4: Balance of Characteristics For Share of In-person UI Claims Processing in 2019

	Coefficient on % In-person			
Dependent Variable	**Est.**	**Std Err.**	**P-Value**	**Mean of Dep Var.**
Demographics				
Log population	0.788	(1.938)	(0.69)	15.21
Median age	-5.455	(3.912)	(0.17)	38.34
High school or lower	0.045	(0.076)	(0.55)	0.41
Some college	-0.006	(0.045)	(0.89)	0.30
Bachelor's degree or higher	-0.039	(0.082)	(0.64)	0.28
% White	0.004	(0.194)	(0.98)	0.77
% Black	-0.230	(0.228)	(0.32)	0.10
% Hispanic	0.200	(0.130)	(0.13)	0.11
% Foreign born	0.013	(0.143)	(0.93)	0.12
Labor and Income				
Income per capita ($1,000)	-10.831	(13.045)	(0.41)	53.45
% Below poverty	0.056	(0.046)	(0.24)	0.13
Employment to population	-0.039	(0.067)	(0.56)	0.60
Labor force participation rate	-0.030	(0.064)	(0.64)	0.64
Self employment rate	-0.009	(0.026)	(0.74)	0.07
Unemployment risk exposure	0.004	(0.060)	(0.95)	0.08
Residential				
% Urban population	0.064	(0.294)	(0.83)	73.4
Homeownership rate	-0.072	(0.081)	(0.38)	0.62
% Houesholds w/ mortgage	-0.026	(0.076)	(0.74)	0.38
Median home value ($1,000)	-15.975	(2651.931)	(1.00)	617.96
Political Environment				
Republican governor	0.865	(0.827)	(0.30)	0.66
Republican vote share (2016)	0.010	(0.164)	(0.95)	0.10
Union membership rate (2018)	-0.024	(0.070)	(0.74)	0.49

Notes: This table reports results from regressions where each one of the state-level characteristics in Column (1) are dependent variables and the COBOL indicator is the independent variable. Variables under Demographics, Labor and Income, and Residential categories, except for income per capita and unemployment risk exposure, are obtained from the 2019 American Community Survey. We obtain income per capita in 2019 for each state from the Bureau of Economic Analysis. We calculate unemployment risk exposure using characteristics of those who became unemployed during April-July 2020 and local demographic characteristics in 2019. Percent urban is measured as of 2010 and is obtained from the U.S. Census Bureau, Republican governor share is measured as of 2018 and is obtained from the National Conference for State Legislatures, union membership is measured as of 2018 and is obtained from the Bureau of Labor Statistics. ***, **, and * indicate significance at the 1%, 5%, and 10% significance levels.

Table A.5: Top 10 Industries With Largest Growth in Business Applications

NAICS-3	NAICS Description	Percentage Growth
454	Nonstore Retailers	80.07
519	Other Information Services	50.48
812	Personal and Laundry Services	49.10
448	Clothing and Clothing Accessories Stores	49.06
492	Couriers and Messengers	48.36
339	Miscellaneous Manufacturing	45.81
425	Wholesale Electronic Markets and Agents and Brokers	42.30
315	Apparel Manufacturing	41.25
493	Warehousing and Storage	40.81
484	Truck Transportation	36.96

Notes: This table reports the top 10 NAICS-3 industries by growth in business applications from 2019 to 2020. NAICS-3 codes that specify no further detail are excluded. Weekly industry level data are obtained from the U.S. Census Bureau's Business Formation Statistics.

Chapter C: Appendix to Chapter 4

A Non-Food Coverage NielsenIQ

The nonfood sector in NielsenIQ is the aggregation of 27 PCE nonfood categories. The 27 nonfood categories are Cellphone, Clocks, Computer Software, Cosmetics, Dishes, Film, Household Cleaning Products, Household Paper Products, Infant Apparel, Lubricants and Fluids, Major Appliances, Medical Products, Miscellaneous Household Products, Newspaper, Nonelectric Cookware, Nonprescription Drugs, Other Appliances, Outdoor Equipment, Personal Care Products, Personal Computer, Pet Products, Photographic Equipment, Sewing, Sporting Equipment, Stationery, Tobacco, and Tools.

B Demand-based Indices

Another set of price indices is based on Constant Elasticity of Substitution (CES) demand systems. The CES utility function yields a tractable demand system with several computable price indexes that correspond exactly to the theoretical unit cost function faced by a representative consumer in the presence of product turnover and time-varying product appeal.

First, we estimate the consumer's elasticity of substitution ($\sigma > 1$) between products that is associated with the CES utility function. Our approach is to estimate a single elasticity for each of the PCE categories. We employ the method used by Feenstra (1994) and Redding & Weinstein (2020) for this purpose. The method double-differences the demand and supply curves sweeping out time and product group effects. The double-differenced demand and supply shocks are assumed to be uncorrelated but heteroskedastic across products. This yields a GMM specification for estimation.

Based on the estimates, we use the Sato-Vartia index (Sato, 1976; Vartia, 1976), the Feenstra-Adjusted Sato-Vartia index (Feenstra, 1995), referred to as "the Feenstra index," and the CES Unified Price Index of Redding & Weinstein (2020), which we will call the "CUPI."

The Sato-Vartia index (Sato, 1976; Vartia, 1976) is defined as:

$$\ln\left(\Phi^{SV}_{t-1,t}\right) = \sum_{k\in\mathbb{C}_t} \omega_{kt} \ln\left(\frac{p_{kt}}{p_{kt-1}}\right), \quad \omega_{kt} = \frac{s_{kt} - s_{kt-1}}{\ln(s_{kt}) - \ln(s_{kt-1})} \Bigg/ \left(\sum_{k\in\mathbb{C}_t} \frac{s_{kt} - s_{kt-1}}{\ln(s_{kt}) - \ln(s_{kt-1})}\right), \tag{C.1}$$

where it abstracts from product turnover.

The Feenstra index (Feenstra (1994)) builds on Sato-Vartia as follows:

$$\Phi^{Feenstra}_{t-1,t} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}} \left(\Phi^{SV}_{t-1,t}\right), \tag{C.2}$$

where $\Phi^{SV}_{t-1,t}$ is the Sato-Vartia price index defined over common varieties, and other terms are defined as:

$$\lambda_{t,t-1} = \frac{\sum_{k\in\mathbb{C}_t} p_{kt}c_{kt}}{\sum_{k\in\Omega_t} p_{kt}c_{kt}}, \qquad \lambda_{t-1,t} = \frac{\sum_{k\in\mathbb{C}_t} p_{kt-1}c_{kt-1}}{\sum_{k\in\Omega_{t-1}} p_{kt-1}c_{kt-1}}. \tag{C.3}$$

The Feenstra-adjusted Sato-Vartia index is exact in the presence of product turnover, but time-varying product appeal is still absent.

The CUPI by Redding & Weinstein (2020) is under more general conditions of product turnover and time variation in product appeal. Redding & Weinstein (2020) presents the exact-price index in this setting (the CUPI) as:

$$\Phi^{CUPI}_{t-1,t} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}} \frac{\tilde{P}^*_t}{\tilde{P}^*_{t-1}} \left(\frac{\tilde{S}^*_t}{\tilde{S}^*_{t-1}}\right)^{\frac{1}{\sigma-1}}. \tag{C.4}$$

The first term is the Feenstra (1994) adjustment factor for product turnover. The second term is the traditional Jevons index, where $\tilde{P}^*_t$ denotes the geometric mean of prices across common goods in period t and $\tilde{P}^*_{t-1}$ represents the same object in period $t-1$ defined over the set of common goods.

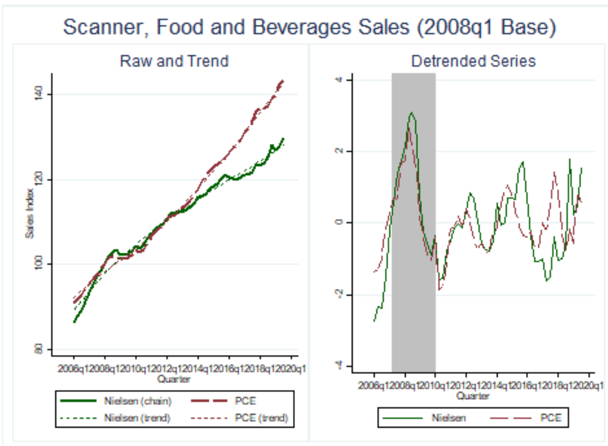
$$\tilde{P}^*_t = \left(\prod_{k \in \mathbb{C}_t} p_{kt} \right)^{\frac{1}{N_{\mathbb{C}_t}}}, \qquad \tilde{P}^*_{t-1} = \left(\prod_{k \in \mathbb{C}_t} p_{kt-1} \right)^{\frac{1}{N_{\mathbb{C}_t}}}. \tag{C.5}$$

The third term is defined as the unweighted geometric average expenditure shares on common varieties in periods $t-1$ and t:

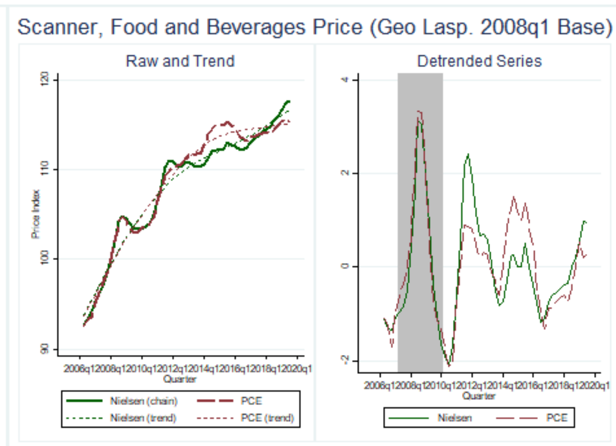
$$\tilde{S}^*_t = \left(\prod_{k \in \mathbb{C}_t} s_{kt} \right)^{\frac{1}{N_{\mathbb{C}_t}}}, \qquad \tilde{S}^*_{t-1} = \left(\prod_{k \in \mathbb{C}_t} s_{kt-1} \right)^{\frac{1}{N_{\mathbb{C}_t}}}. \tag{C.6}$$

C Trend and Detrended Series

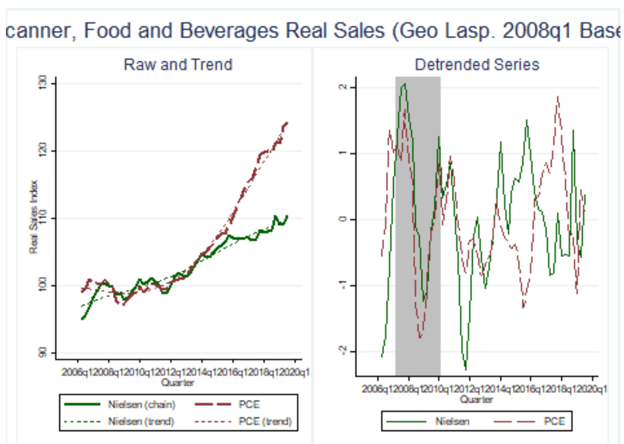
Figure C.1: Trend and Detrended Series: Food and Beverages



Sales Index



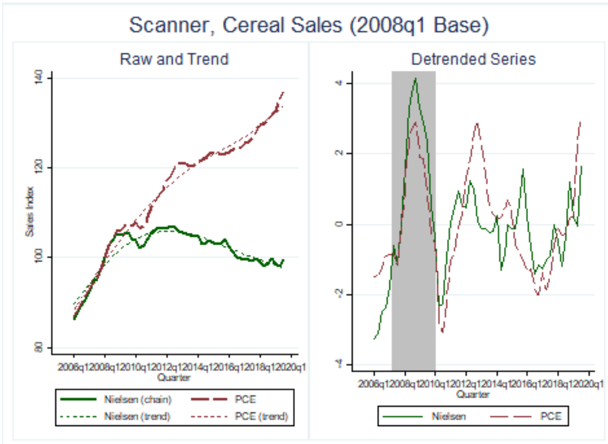
Price Index



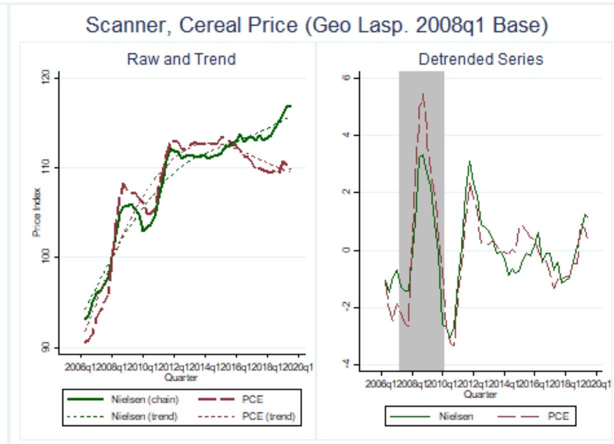
Real Sales Index

Note: This figure represents the trend and detrended nominal sales (the top left panel), prices (the top right panel), and real sales (the bottom panel) indices for Food and Beverages. The indices are normalized at 2008Q1, and price index is based on the geometric Laspeyres.

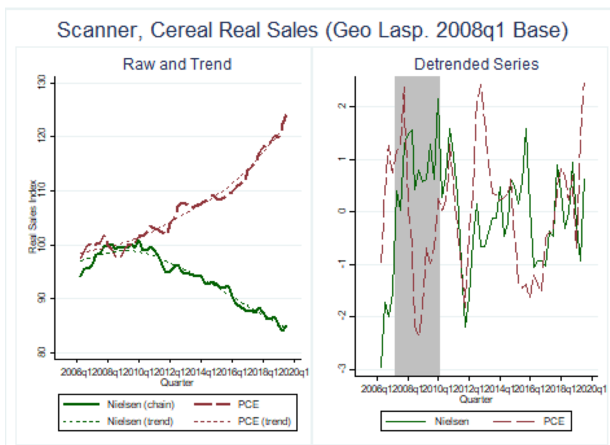
Figure C.2: Trend and Detrended Series: Cereal



Sales Index



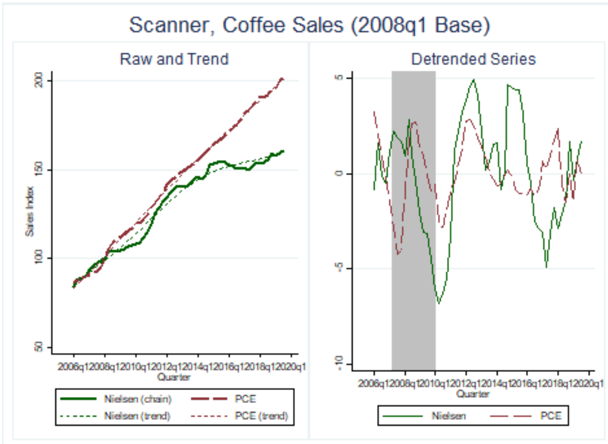
Price Index



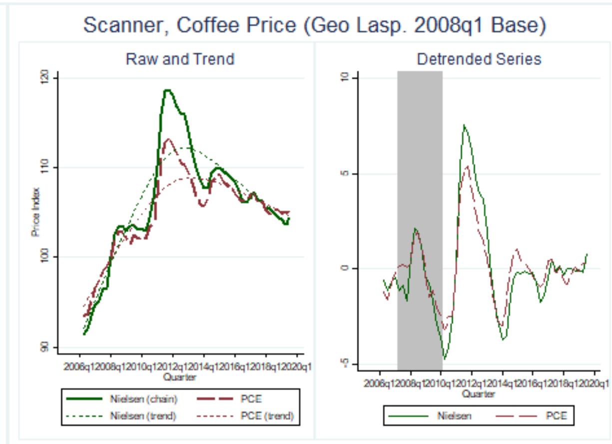
Real Sales Index

Note: This figure represents the trend and detrended nominal sales (the top left panel), prices (the top right panel), and real sales (the bottom panel) indices for Cereal. The indices are normalized at 2008Q1, and price index is based on the geometric Laspeyres.

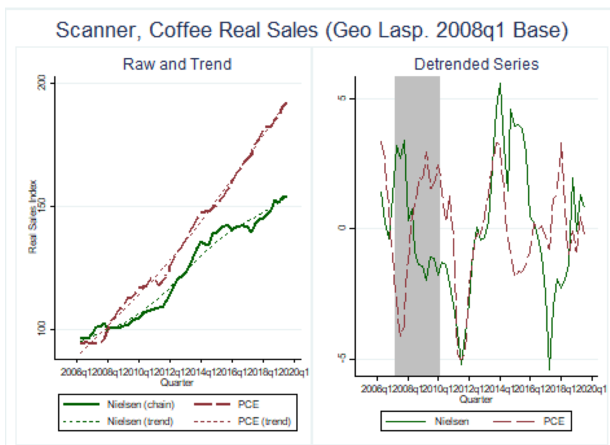
Figure C.3: Trend and Detrended Series: Coffee and Tea



Sales Index



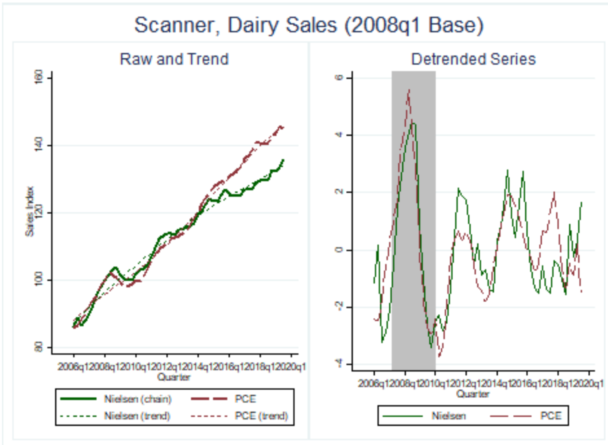
Price Index



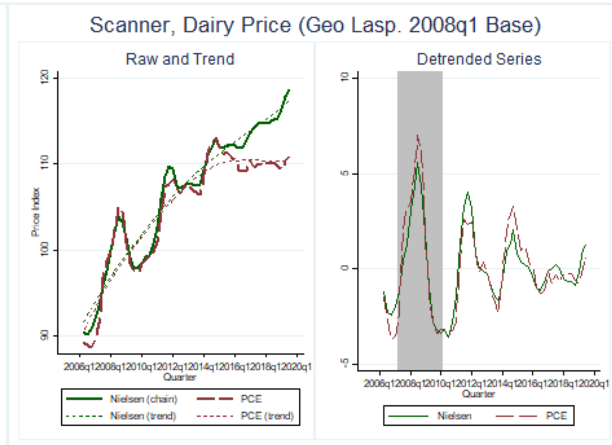
Real Sales Index

Note: This figure represents the trend and detrended nominal sales (the top left panel), prices (the top right panel), and real sales (the bottom panel) indices for Coffee and Tea. The indices are normalized at 2008Q1, and price index is based on the geometric Laspeyres.

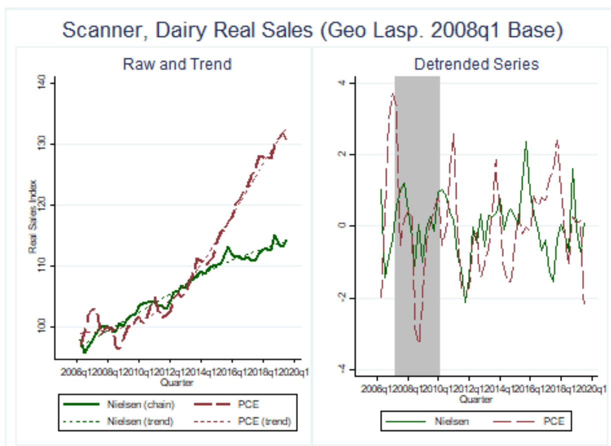
Figure C.4: Trend and Detrended Series: Dairy



Sales Index



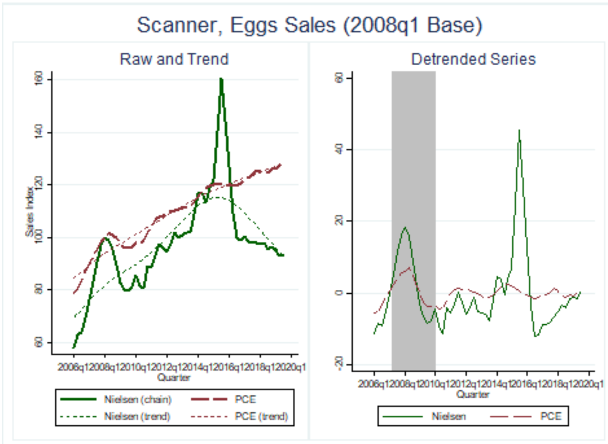
Price Index



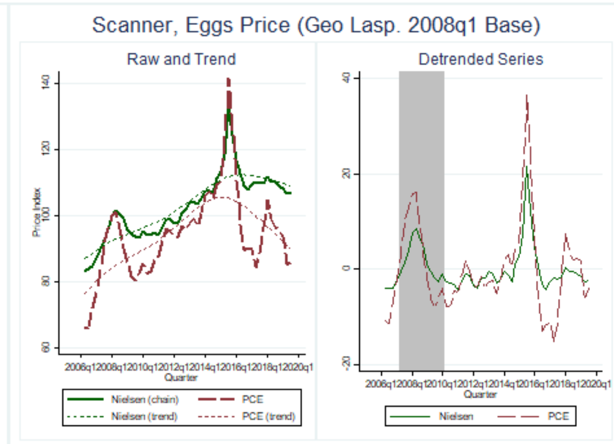
Real Sales Index

Note: This figure represents the trend and detrended nominal sales (the top left panel), prices (the top right panel), and real sales (the bottom panel) indices for Dairy. The indices are normalized at 2008Q1, and price index is based on the geometric Laspeyres.

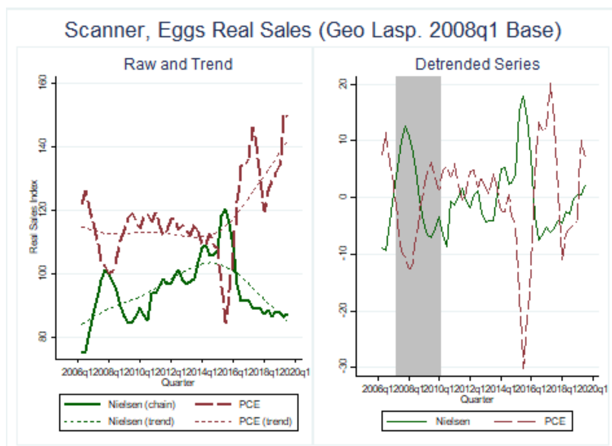
Figure C.5: Trend and Detrended Series: Eggs



Sales Index



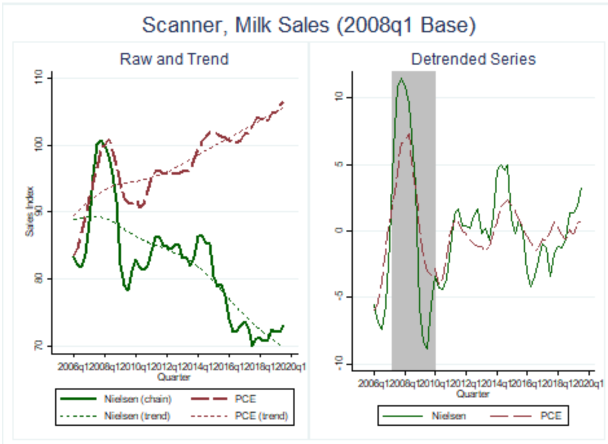
Price Index



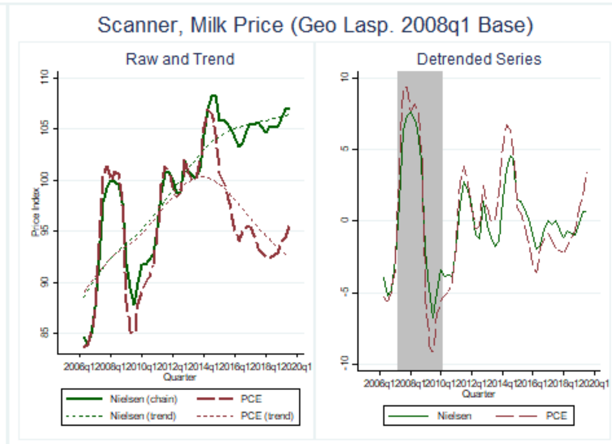
Real Sales Index

Note: This figure represents the trend and detrended nominal sales (the top left panel), prices (the top right panel), and real sales (the bottom panel) indices for Eggs. The indices are normalized at 2008Q1, and price index is based on the geometric Laspeyres.

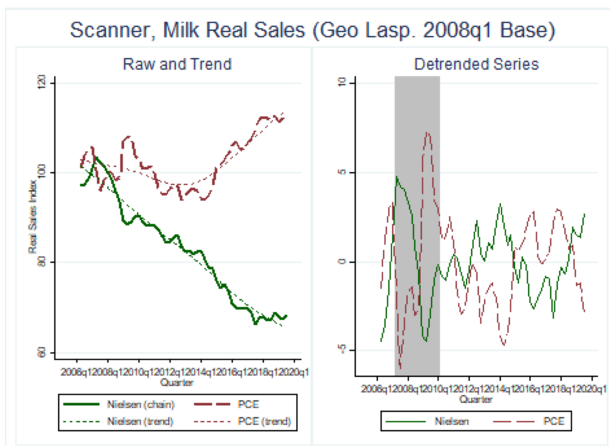
Figure C.6: Trend and Detrended Series: Milk



Sales Index



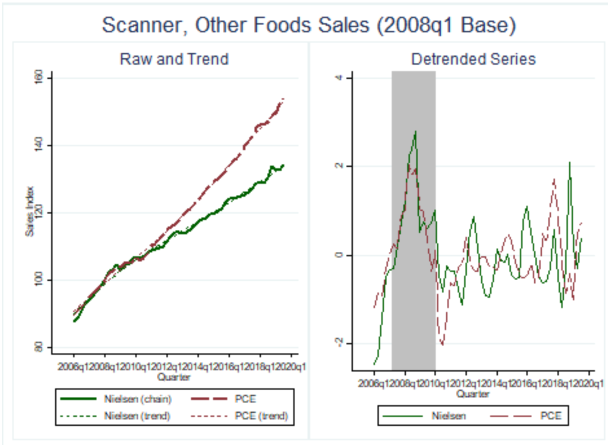
Price Index



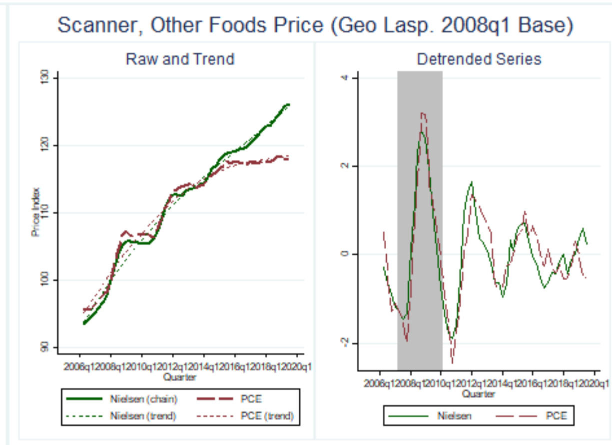
Real Sales Index

Note: This figure represents the trend and detrended nominal sales (the top left panel), prices (the top right panel), and real sales (the bottom panel) indices for Milk. The indices are normalized at 2008Q1, and price index is based on the geometric Laspeyres.

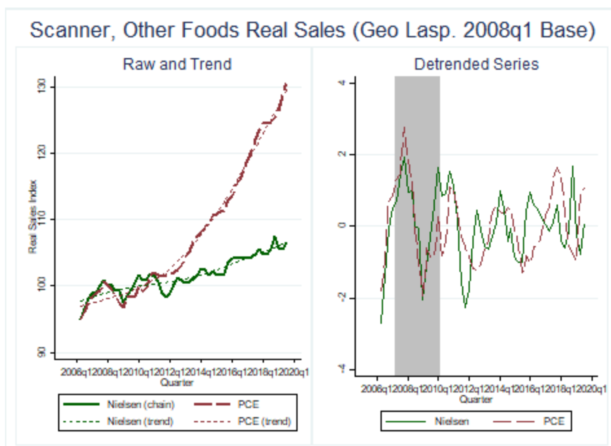
Figure C.7: Trend and Detrended Series: Other Foods



Sales Index



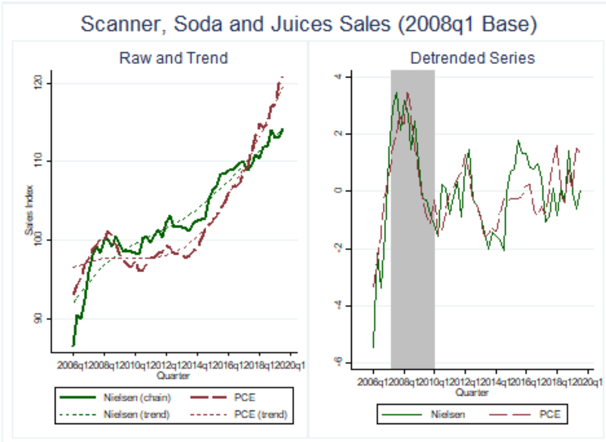
Price Index



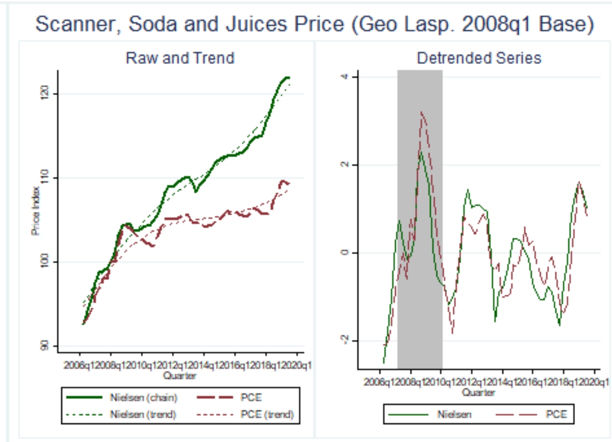
Real Sales Index

Note: This figure represents the trend and detrended nominal sales (the top left panel), prices (the top right panel), and real sales (the bottom panel) indices for Other Foods. The indices are normalized at 2008Q1, and price index is based on the geometric Laspeyres.

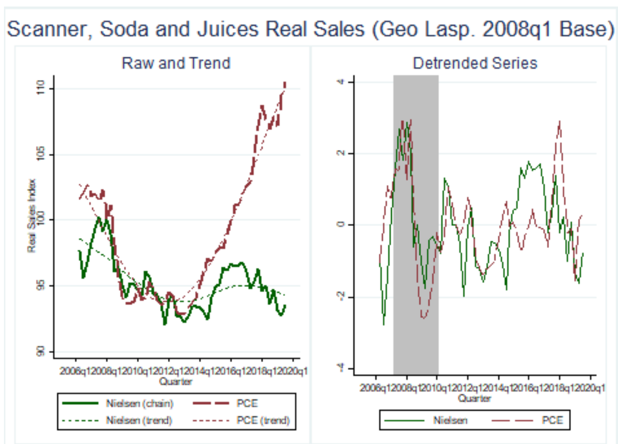
Figure C.8: Trend and Detrended Series: Soda and Juices



Sales Index



Price Index



Real Sales Index

Note: This figure represents the trend and detrended nominal sales (the top left panel), prices (the top right panel), and real sales (the bottom panel) indices for Soda. The indices are normalized at 2008Q1, and price index is based on the geometric Laspeyres.

Bibliography

Abadie, Alberto & Jeremy L'Hour (2021), "A Penalized Synthetic Control Estimator for Disaggregated Data," *Journal of the American Statistical Association* 116(536): 1817–1834, doi: 10.1080/01621459.2021.1971535, URL https://doi.org/10.1080/01621459.2021.1971535.

Adams-Prassl, Abi, Teodora Boneva, Marta Golin & Christopher Rauh (2020), "Inequality in the impact of the coronavirus shock: Evidence from real time surveys," *Journal of Public economics* 189: 104245.

Akcigit, Ufuk & Sina T Ates (2023), "What happened to US business dynamism?" *Journal of Political Economy* 131(8): 2059–2124.

Aladangady, Aditya, David Cho, Laura Feiveson & Eugenio Pinto (2022), "Excess Savings During the COVID-19 Pandemic," *FEDS Notes* (2022-10): 21.

Alon, Titan, David Berger, Robert Dent & Benjamin Pugsley (2018), "Older and slower: The startup deficit's lasting effects on aggregate productivity growth," *Journal of Monetary Economics* 93: 68–85, ISSN 0304-3932, doi:https://doi.org/10.1016/j.jmoneco.2017.10.004, URL https://www.sciencedirect.com/science/article/pii/S0304393217301113.

Amburgey, Aaron, Serdar Birinci et al. (2020), "Which Earnings Groups Have Been Most Affected by the COVID-19 Crisis," *Economic Synopses* (37).

Andrews, R.J., Catherine Fazio, Jorge Guzman, Yupeng Liu & Scott Stern (2022), "The Startup Cartography Project: Measuring and Mapping Entrpreneurial Ecosystems," *Research Policy* 51.

Atkinson, Anthony B. & John Micklewright (1991), "Unemployment Compensation and Labor Market Transitions: A Critical Review," *Journal of Economic Literature* 29(4): 1679–1727.

Bayard, Kimberly, Emin Dinlersoz, Timothy Dunne, John Haltiwanger, Javier Miranda & John Stevens (2018), "Early-Stage Business Formation: An Analysis of Applications for Employer Identification Numbers," NBER Working Paper No. 24364.

Boone, Christopher, Arindrajit Dube, Lucas Goodman & Ethan Kaplan (2021), "Unemployment Insurance Generosity and Aggregate Employment," *American Economic Journal: Economic Policy* 13(2): 58–99.

Buera, Francisco J (2009), "A dynamic model of entrepreneurship with borrowing constraints: theory and evidence," *Annals of finance* 5: 443–464.

Cajner, Tomaz, Leland D. Crane, Ryan A. Decker, John Grigsby, Adrian Hamins-Puertolas, Erik Hurst, Christopher Kurz & Ahu Yildirmaz (2020a), "The U.S. Labor Market during the Beginning of the Pandemic Recession," Working Paper 27159, National Bureau of Economic Research, doi:10.3386/w27159, URL http://www.nber.org/papers/w27159.

Cajner, Tomaz, Andrew Figura, Brendan Price, David Ratner & Alison Weingarden (2020b), "Reconciling unemployment claims with job losses in the first months of the COVID-19 crisis," Working paper, Board of Governors of the Federal Reserve System, URL https://doi.org/10.17016/FEDS.2020.055.

Callaway, Brantly & Pedro H.C. Sant'Anna (2021), "Difference-in-Differences with multiple time periods," *Journal of Econometrics* 225: 200–230.

Camarero Garcia, Sebastian & Michelle Hansch (2021), "The Effect of Unemployment Insurance Benefits on (Self-)Employment: Two Sides of the Same Coin?" Bundesbank Discussion Paper 18/2021.

Card, David, Andrew Johnston, Pauline Leung, Alexandre Mas & Zhuan Pei (2015), "The Effect of Unemployment Benefits on the Duration of Unemployment Insurance Receipt: New Evidence from a Regression Kink Design in Missouri, 2003-2013," *American Economic Review* 105(5): 126–130.

Chetty, Raj, John N Friedman, Nathaniel Hendren, Michael Stepner & The Opportunity Insights Team (2020), "How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data," Working Paper 27431, National Bureau of Economic Research, doi:10.3386/w27431, URL http://www.nber.org/papers/w27431.

Chodorow-Reich, Gabriel, John Coglianese & Loukas Karabarbounis (2019), "The Maro Effects of Unemployment Benefits Extensions: A Measurement Error Approach," *The Quarterly Journal of Economics* 134(1): 227–279.

Coibion, Olivier, Yuriy Gorodnichenko & Michael Weber (2020), "The Cost of the Covid-19 Crisis: Lockdowns, Macroeconomic Expectations, and Consumer Spending," Working Paper 27141, National Bureau of Economic Research, doi:10.3386/w27141, URL http://www.nber.org/papers/w27141.

Corradin, Stefano & Alexander Popov (2015), "House prices, home equity borrowing, and entrepreneurship," *The Review of Financial Studies* 28(8): 2399–2428.

Cox, Natalie, Peter Ganong, Pascal Noel, Joseph Vavra, Arlene Wong, Diana Farrell, Fiona Greig & Erica Deadman (2020), "Initial impacts of the pandemic on consumer behavior: Evidence from linked income, spending, and savings data," *Brookings Papers on Economic Activity* 2020(2): 35–82.

Davis, Steven J. & John Haltiwanger (forthcoming), "Dynamism Diminished: The Role of Housing Markets and Credit Conditions," *American Economic Journal: Macroeconomics* .

Decker, Ryan & John Haltiwanger (2023), "Surging Business Formation in the Pandemic: Causes and Consequences," Fall 2023 Brookings Papers on Economic Activity.

Decker, Ryan, John Haltiwanger, Ron Jarmin & Javier Miranda (2014), "The role of entrepreneurship in US job creation and economic dynamism," *Journal of Economic Perspectives* 28(3): 3–24.

Di Maggio, Marco & Amir Kermani (2016), "The Importance of Unemployment Insurance as an Automatic Stabilizer," Tech. rep., National Bureau of Economic Research.

Dinlersoz, Emin, Timothy Dunne, John Haltiwanger & Veronika Penciakova (2021), "Business Formation: A Tale of Two Recessions," *AEA Papers and Proceedings* 111: 253–257.

Dinlersoz, Emin, Timothy Dunne, John Haltiwanger & Veronika Penciakova (2023), "The Local Origins of Business Formation," Federal Reserve Bank of Atlanta Working Paper No. 2023-9.

Dolls, Mathias, Clemens Fuest & Andreas Peichl (2012), "Automatic stabilizers and economic crisis: US vs. Europe," *Journal of Public Economics* 96(3): 279–294, ISSN 0047-2727, doi: https://doi.org/10.1016/j.jpubeco.2011.11.001, URL https://www.sciencedirect.com/science/article/pii/S0047272711001642.

Ehrlich, Gabriel, John Haltiwanger, Edward Olivares, Matthew D. Shapiro & Laura Yi Zhao (2023), "Demand Based Quality-Adjusted Price Indices: Theory vs Evidence," Tech. rep., Census, Maryland and Michigan.

Eilbott, Peter (1966), "The Effectiveness of Automatic Stabilizers," *American Economic Review* 56(3): 450–465, ISSN 00028282, URL http://www.jstor.org/stable/1823778.

Evans, David S & Boyan Jovanovic (1989), “An estimated model of entrepreneurial choice under liquidity constraints,” *Journal of political economy* 97(4): 808–827.

Faberman, R. Jason, Andreas I Mueller & Ayşegül Şahin (2022), “Has the Willingness to Work Fallen during the Covid Pandemic?” Working Paper 29784, National Bureau of Economic Research, doi:10.3386/w29784, URL http://www.nber.org/papers/w29784.

Farber, Henry S., Jesse Rothstein & Robert G. Valletta (2015), “The Effect of Extended Unemployment Inusrance Benefits: Evidence from the 2012-2013 Phase-out,” *American Economic Review* 105(5): 171–176.

Faulkender, Michael W, Robert Jackman & Stephen Miran (2023), “The job preservation effects of paycheck protection program loans,” *Available at SSRN 3767509* .

Fazio, Catherine E., Jorge Guzman, Yupeng Liu & Scott Stern (2021), “How is COVID Changing the Georgraphy of Entrepreneurship? Evidence from the Startup Cartography Project,” NBER Working Paper No. 28787.

Feenstra, Robert C (1994), “New Product Varieties and the Measurement of International Prices,” *The American Economic Review* 84(1): 157–177.

Feenstra, Robert C (1995), “Exact Hedonic Price Indexes,” 77(4): 634–653.

Gaillard, Alexandre & Sumudu Kankanamge (2023), “Gross Labor Market Flows, Self-Employment, and Unemployment Insurance,” Working Paper.

Ganong, Peter, Fiona E. Greig, Pascal J. Noel, Daniel M. Sullivan & Joseph S. Vavra (2022a), “Spending and Job-Finding Impacts of Expanded Unemployment Benefits: Evidence from Administrative Micro Data,” NBER Working Papers 30315, National Bureau of Economic Research, Inc, URL https://ideas.repec.org/p/nbr/nberwo/30315.html.

Ganong, Peter, Fiona E Greig, Pascal J Noel, Daniel M Sullivan & Joseph S Vavra (2022b), "Spending and Job-Finding Impacts of Expanded Unemployment Benefits: Evidence from Administrative Micro Data," Tech. rep., National Bureau of Economic Research.

Ganong, Peter, Pascal Noel & Joseph Vavra (2020a), "US unemployment insurance replacement rates during the pandemic," *Journal of Public Economics* 191: 104273, ISSN 0047-2727, doi: https://doi.org/10.1016/j.jpubeco.2020.104273, URL https://www.sciencedirect.com/science/article/pii/S0047272720301377.

Ganong, Peter, Pascal Noel & Joseph Vavra (2020b), "US Unemployment Insurance Replacement Rates during the Pandemic," *Journal of Public Economics* 191: 104273, ISSN 0047-2727, doi:https://doi.org/10.1016/j.jpubeco.2020.104273, URL https://www.sciencedirect.com/science/article/pii/S0047272720301377.

Guzman, Jorge & Scott Stern (2020), "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 U.S. States, 1988–2014," *American Economic Journal: Economic Policy* 12(4): 212–243.

Hale, Thomas, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar et al. (2021), "A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)," *Nature human behaviour* 5(4): 529–538.

Herd, Pamela & Donald P. Moynihan (2018), *Administrative Burden: Policymaking by Other Means*, Russell Sage Foundation, ISBN 9780871544445, URL http://www.jstor.org/stable/10.7758/9781610448789.

Hombert, Johan, Antoinette Schoar, David Alexander Sraer & David Thesmar (2020), "Can Unemployment Insurance Spur Entrepreneurial Activity? Evidence from France," *Journal of Finance* 75: 1247–1285.

Hopenhayn, Hugo, Julian Neira & Rish Singhania (2022), "From population growth to firm demographics: Implications for concentration, entrepreneurship and the labor share," *Econometrica* 90(4): 1879–1914.

Hottman, Colin J, Stephen J Redding & David E Weinstein (2016), "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics* 131(3): 1291–1364.

Katz, Lawrence F. & Bruce D. Meyer (1990), "Unemployment Insurance, Recall Expectations, and Unemployment Outcomes," *The Quarterly Journal of Economics* 105(4): 973–1002.

Kekre, Rohan (2022), "Unemployment Insurance in Macroeconomic Stabilization," *Review of Economic Studies* 90(5): 2439–2480, ISSN 0034-6527, doi:10.1093/restud/rdac080, URL https://doi.org/10.1093/restud/rdac080.

Kerr, Sari Pekkala, William R. Kerr & Ramana Nanda (2022), "House prices, home equity and entrepreneurship: Evidence from U.S. census micro data," *Journal of Monetary Economics* 130: 103–119.

Kerr, William R & Ramana Nanda (2011), "8 Financing constraints and entrepreneurship," *Handbook of Research on Innovation and Entrepreneurship* p. 88.

Lachowska, Marta, Alexandre Mas & Stephen A. Woodbury (2022), "Poor Performance as a Predictable Outcome: Financing the Administration of Unemployment Insurance," *AEA Papers and Proceedings* 112: 102–06, doi:10.1257/pandp.20221073, URL https://www.aeaweb.org/articles?id=10.1257/pandp.20221073.

Lalive, Rafael, Jan van Ours & Josef Zweimuller (2006), "How Changes in Financial Incentives Affect the Duratin of Unemployment," *Review of Economic Studies* 73(4): 1009–1038.

Lastrapes, William D., Ian Schmutte & Thor Watson (2022), "Home equity lending, credit constraints and small business in the US," *Economic Inquiry* 60(1): 43–63.

Marinescu, Ioana, Daphné Skandalis & Daniel Zhao (2021), "The Impact of the Federal Pandemic Unemployment Compensation on Job Search and Vacancy Creation," *Journal of Pub-

lic Economics 200: 104471, ISSN 0047-2727, doi:10.1016/j.jpubeco.2021.104471, URL https://www.sciencedirect.com/science/article/pii/S0047272721001079.

McKay, Alisdair & Ricardo Reis (2016), "The Role of Automatic Stabilizers in the U.S. Business Cycle," *Econometrica* 84(1): 141–194, doi:https://doi.org/10.3982/ECTA11574, URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA11574.

Meyer, Bruce D. (1990), "Unemployment Insurance and Unemployment Spells," *Econometrica* 58: 757–782.

Nekoei, Arash & Andrea Weber (2017), "Does Extending Unemployment Benefits Improve Job Quality?" *American Economic Review* 107(2): 527–561.

Nichols, Albert L. & Richard J. Zeckhauser (1982), "Targeting Transfers through Restrictions on Recipients," *American Economic Review* 72(2): 372–377, ISSN 00028282, URL http://www.jstor.org/stable/1802361.

Parker, Jonathan A., Nicholas S. Souleles, David S. Johnson & Robert McClelland (2013), "Consumer Spending and the Economic Stimulus Payments of 2008," *American Economic Review* 103(6): 2530–53, doi:10.1257/aer.103.6.2530, URL https://www.aeaweb.org/articles?id=10.1257/aer.103.6.2530.

Parton, Brent (2023), "Announcement of a Grant to Support State Unemployment Insurance (UI) Information Technology (IT) Modernization Activities under the American Rescue Plan Act (ARPA)," Online source, Brookings Institution, URL https://www.dol.gov/sites/dolgov/files/ETA/advisories/UIPL/2023/UIPL%2007-23/UIPL%2007-23.pdf.

Pugsley, Benjamin Wild & Ayşegül Şahin (2019), "Grown-up business cycles," *The Review of Financial Studies* 32(3): 1102–1147.

Redding, Stephen & David E. Weinstein (2018), “Measuring Aggregate Price Indexes with Taste Shocks: Theory and Evidence for CES Preferences,” *Quarterly Journal of Economics* 135: 503–560.

Redding, Stephen J & David E Weinstein (2020), “Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences,” *The Quarterly Journal of Economics* 135(1): 503–560, ISSN 0033-5533, doi:10.1093/qje/qjz031.

Sato, Kazuo (1976), “The Ideal Log-Change Index Number,” *The Review of Economics and Statistics* ISSN 00346535, doi:10.2307/1924029.

Schmalz, Martin C, David A Sraer & David Thesmar (2017), “Housing collateral and entrepreneurship,” *The Journal of Finance* 72(1): 99–132.

Schmieder, Johannes F., Till von Wachter & Stefan Bender (2012), “The Effects of Extending Unemployment Insurance Over the Business Cycle: Evidence from Regression Discontinuity Estimates Over 20 Years,” *The Quarterly Journal of Economics* 127(2): 701–752.

Simon-Mishel, Julia, Maurice Emsellem, Michele Evermore, Andrew Leclere, Ellen Stettner & Martha Coven (2020), “Centering Workers–How to Modernize Unemployment Insurance Technology,” Tech. rep., The Century Foundation, URL https://tcf.org/content/report/centering-workers-how-to-modernize-unemployment-insurance-technology/.

Skandalis, Daphné, Ioana Marinescu & Maxim N. Massenkoff (2022), “Racial Inequality in the US Unemployment Insurance System,” Tech. rep., National Bureau of Economic Research.

Vartia, Yrjo (1976), “Ideal Log-Change Index Numbers,” *Scandinavian Journal of Statistics* 3(3): 121–126.