ESSAYS ON INDUSTRIAL ORGANIZATION AND HEALTH ECONOMICS

by

SICHAO JIANG

A DISSERTATION

Presented to the Department of Economics
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

March 2022

DISSERTATION APPROVAL PAGE

Student: Sichao Jiang

Title: Essays on Industrial Organization and Health Economics

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Economics by:

Keaton Miller	Co-Chairperson
Wesley W. Wilson	Co-Chairperson
Edward Rubin	Core Member
N. Narasimha Murthy	Institutional Representative

and

Krista Chronister	Vice Provost of Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded March 2022

DISSERTATION ABSTRACT

Sichao Jiang

Doctor of Philosophy

Department of Economics

March 2022

Title: Essays on Industrial Organization and Health Economics

This thesis is composed of three essays and provides empirical contributions to the Industrial Organization literature, especially in the field of transportation and health economics. It aims to understand different issues related to economics by applying various empirical methods.

The first essay (chapter 2) examines firm exit in Canadian markets, specifically the grain elevator market. There is a long line of previous literature that finds capacity, vintage, multi-plant ownership affect exit. In this paper, a choice model is used to examine a firm's decision to shut down a grain elevator in terms of these variables, but also develops measures of spatial competition, local economic conditions and linkages to the transportation markets. In all cases, these variables are statistically important and point to results that reinforce previous studies, but also direct to new explanations on the determinants of plant exit.

The second essay (chapter 3) examines the effects of marijuana legislation change on the agricultural labor market. The paper uses differences-in-differences with a synthetic control methodology to identify the effects of labor market outcomes from marijuana legalization. This method aims to avoid substantial labor market spillovers in neighboring states and to construct a decent parallel trend for the pre-treatment time period with pretty varied agricultural markets in the U.S.

The results show that cannabis legalization is associated with an increase in overall employments that people are flushing into the industry, but no increase in per-employee wages in both the retailer and agricultural labor market.

The third essay (chapter 4) looks into the accuracy of firms' prediction errors in the context of Medicare Advantage, where insurers receive subsidies from the government and compete to provide health insurance to seniors. The results show that on average firms overestimate future costs. Overestimation in forecast error decreases with the experience of the firm. Firms in more competitive markets (as measured by the number of other firms present) form more accurate estimates. Firms with higher costs than expected generally offer plans that feature greater patient cost sharing (i.e. higher deductibles and copays).

This dissertation includes both previously published/unpublished and co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Sichao Jiang

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR, USA
Texas Tech University, Lubbock, TX, USA

DEGREES AWARDED:

Doctor of Philosophy, Economics, 2022, University of Oregon
Master of Science, Economics, 2018, University of Oregon
Bachelor of Science, Economics & Mathematics, 2017, Texas Tech University

AREAS OF SPECIAL INTEREST:

Industrial Organization
Applied Econometrics
Applied Microeconomics

PROFESSIONAL EXPERIENCE:

Graduate Employee, Department of Economics, University of Oregon, 2018-2022
Summer Associate, Analysis Group, 2021
Research Consultant, Western Economic Association International,2020-2021

GRANTS, AWARDS AND HONORS:

Graduate Teaching Fellowship, University of Oregon, 2018-2022
Edward G. Daniel Scholarship, University of Oregon, 2018
Kleinsorge Fellowship Award and Distinctive Scholar Award, University of Oregon, 2017
Graduate School First-Year Fellowship, University of Oregon, 2017

PUBLICATIONS:

Jiang, S., Nolan, J. & Wilson, W.W. (2021). Exit Decisions in the Canadian Grain Elevator. *Journal of Industry, Competition and Trade*, 1-19.

TABLE OF CONTENTS

LIST OF FIGURES

Figure | Page

Figure | Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

This thesis is composed by three essays and provides empirical contributions to the Industrial Organization literature, specially in the field of transportation and health economics. It aims to understand different issues related to economics by applying various empirical methods which include models of exit, policy effects on labor market, and prediction errors in the context of Medicare Advantage settings. The empirical models include conventional techniques logit modeling and piecewise regression models, causal inference methods i.e., different in different and synthetic control, and some machine learning applications like LASSO, MICE, regression trees and etc.

In chapter 2, jointly with Dr. James Nolan and Dr. Wesley W. Wilson, we examined firm exit in Canadian markets, specifically the grain elevator market. Grain elevators play a central role in the movement of grain to market and to rural economies in terms of employment and investment. Over the last three decades, the grain elevator industry in Canada has experienced a major decline in the number of elevators as older and technologically obsolete elevators have been replaced by larger and more technologically advanced elevators. We develop a model of exit in the Canadian grain elevation industry using data from 1999 to 2016 collected at the individual elevator level. Our specification explains elevator exit based on traditional variables used in the industrial organization literature such as capacity, multi-plant ownership, and vintage. But, we also include a measure of vertical linkages in the industry (i.e., the effects of vertical investments in transportation infrastructure) as well as spatial measures to account for local demand, supply and competition. The results provide strong evidence that exit in this key agricultural and trade industry is affected by whether an elevator is a recent entrant (vintage), its size, vertical linkages, local demand and supply conditions, and spatial competition.

In chapter 3, jointly with Dr. Keaton Miller, we examined the effects of marijuana legislation change on the agricultural labor market. Over the past several years, cannabis has become legal for recreational use in several U.S. states and jurisdictions around the world. The opening of these markets has led to the establishment of hundreds of cannabis production and retail firms with accompanying demand for labor, leading to concerns about spillover effects on wages from incumbents. We study the markets for agricultural and retail labor in Washington and Colorado, early legalizers with now-established cannabis markets. Using a synthetic control technique to account for the possibility of border-state spillover effects and machine learning techniques for data imputation and variable selection, we find no evidence that cannabis legalization is associated with increases in per-employee wages, neither within industries most similar to cannabis production or retail, nor in more broad industry categories. We conclude that cannabis legalization is unlikely to negatively impact incumbent firms through the labor market channel.

In chapter 4, jointly with Dr. Keaton Miller we investigated the accuracy of the predictions in the context of Medicare Advantage, where insurers receive subsidies from the government and compete to provide health insurance to seniors. As part of the "bidding" process, firms must submit forecasts of their costs. Insurers have incentives to report accurately, as these predictions are used to determine both the level of the subsidy and (implicitly) the degree to which that subsidy can be spent on various plan features. We collect data on predictions and realized expenses per member per month at the plan-service-category level from 2008-2015, and document three stylized facts. Our results show that first, on average firms overestimate future costs. Second, this overestimation decreases with the experience of the firm. Firms in more competitive markets (as measured by the number of other firms present) form more accurate estimates. We show that firms with higher costs than expected generally offer plans that feature greater patient cost sharing (i.e. higher deductibles and copays).

CHAPTER II

EXIT DECISIONS IN THE CANADIAN GRAIN ELEVATOR INDUSTRY

From *Jiang, S., Nolan, J. & Wilson, W.W. (2021). Exit Decisions in the Canadian Grain Elevator.* Journal of Industry, Competition and Trade*, 1-19.*

2.1 Introduction

Grain elevators have long been a visible and important component in rural communities. They not only provide employment opportunities and invest in the local areas, they are also the central local gathering point for agricultural products and link the local markets to destinations. Elevators not only receive grain, they also store, blend and/or treat grains, ultimately loading the stored grain for shipment to terminals and processing plants. In Canada, since the first Prairie elevator was built in 1881 in Gretna, Manitoba there has been enormous investment in rural elevators to accommodate the long-term growth of the Prairie grain industry.[1]

Over the last few decades, the industry has transitioned as older (mostly built of wood) and smaller elevators have given way to more modern (mostly built of concrete) larger elevators. In this paper, we examine this transition by developing and estimating a model of exit that captures traditionally used determinants such as scale, multi-plant firms, plant vintage, but also develop and include a measure of vertical linkages as well as local measures of demand, supply, and spatial competition. We develop a panel dataset of Western Canadian elevators in operation during 1999-2016 time period. These data provide a comprehensive unbalanced panel that allow patterns of exit to be described and

[1] `https://www.thecanadianencyclopedia.ca/en/article/grain-elevators.`

allow the estimation of a unique model of exit to be estimated. We first estimate a model motivated by the industrial organization literature on exit. The model includes variables such as capacity, multi-plant ownership, and whether or not an elevator was a recent entrant to the market to capture elevator vintage. We then develop and include in the model, measures of the vertical relationship between the elevators and the transportation markets as well as a set of variables to reflect the spatial setting of the elevator. These latter include local measures such as demand, supply and spatial competition. The results provide strong evidence consistent with the literature that capacity and time of entry (i.e. more recent, newer technology) have negative effects on exit. We also find that vertical relationships as well as spatial measures of demand, supply and spatial competition have expected and strong effects on the probability that an elevator in the sample exited the market.

The current industry in Canada consists of four types of elevators. These are: primary, forwarding, process, and terminal elevators. Primary elevators receive grain directly from producers for storage and forwarding. Process elevators receive and store grain for direct manufacture or processing into other grain products. Terminal elevators receive grain after official inspection and weighing, cleaning and storing grain before moving it along the supply chain; and transfer elevators transfer grain that has been officially inspected and weighed at another elevator. Transfer elevators can also receive, clean, and store domestic or foreign grain.[2] Primary elevators still dominate the grain elevator industry in Canada. In 1999, there were 976 primary elevators, and a total of 57 other types of elevators. The number of primary elevators generally decreases throughout the range of the data (1999-2016). Yet since 2004, total capacity of primary grain elevators has increased

[2] https://en.wikipedia.org/wiki/Grain_elevator

in most Canadian provinces. For example, Alberta has seen total primary elevator capacity increases from 1,685,250 to 1,834,160 tonnes within the time frame of 1999 to 2016. British Columbia is the only province where inland elevator capacity fell, from 46,030 in 1999 to 41,130 tonnes in 2016.[3] Overall, the industry now has fewer but larger elevators, while these remain mostly primary elevators. Concurrently, the number of process and terminal elevators have increased while their average capacity has dropped. And since 2013, no transfer elevators have been operational in this market. We focus the analysis on primary elevators, given their dominance in the industry.

There is considerable research on industrial entry and exit.[4] Generally, this literature finds that inefficient firms/plants tend to exit the market either due to the lack of scale economies or due to inherent inefficiencies.[5] Other research points to strategic motivations for exit e.g., Ghemawat and Nalebuff (1985), Ghemawat and Nalebuff (1990), and others. In addition, recent research provides evidence that firm characteristics at time of entry have an important effect on exit (T. Dunne, Klimek, and Roberts (2005), while other research theoretically points to the role of multiplant ownership e.g., Reynolds (1988) and Ghemawat and Nalebuff (1990) and and empirically (e.g., Audretsch and Mahmood (1995), Mata et al. (1995),

[3]The data comes from the Canadian Grain Commission, `https://www.grainscanada.gc.ca/wa-aw/geic-sgc/summary-sommaire-eng.asp`

[4]There is a related literature that examines the effects of financial information on firm failure e.g., Altman (1968, 1973), Zingales (1998). Zingales (1998) finds, for example, that highly leverages trucking firms are less likely to survive following deregulation. In the case of the Canadian grain elevator industry, elevator success in Western Canada is tied to several factors, including financial/operational elements (e.g. turnover, capacity, pricing, etc.) as well as transportation (i.e. rail and road) connectivity (Lawrence, Nolan, & Schoney, 2016).

[5]See, for example, Franklin (1974), Jovanovic (1982),T. Dunne, Roberts, and Samuelson (1988, 1989), Lieberman (1990),Audretsch (1991, 1995), P. Dunne and Hughes (1994), Mata, Portugal, and Guimaraes (1995), Gibson and Harris (1996), Audretsch, Houweling, and Thurik (2000), Segarra and Callejón (2002), Elston and Agarwal (2004), K. S. Miller and Wilson (2018)and etc.

and K. S. Miller and Wilson (2018)).[6] Previous research has shown that the effects of scale can have a positive effect on the likelihood of exit. That is, in declining demand markets. Ghemawat and Nalebuff (1985) show theoretically that a small single plant firm can profitably "hang on" longer than a large firm, with the result that the larger firm exits first. In a subsequent paper, Ghemawat and Nalebuff (1990) allow partial capacity adjustments, but similarly find that large firms reduce capacity before small firms and then both small and large firms reduce capacity until their plants are equally sized, subsequently reducing capacity at the same rate. When firms operate multiple plants, Whinston (1988) find that the large plant can improve a multi-plant firm's strategic position in the survival game, with the result that this firm will not necessarily be the first to exit or cut capacity. Fudenberg and Tirole (1986) analyze a model in which two firms possess asymmetric information about each other's fixed costs, but hold symmetric expectations. They identify a unique subgame perfect equilibrium where high-cost firms exit earlier than low-cost firms. We conclude that overall, the implications of plant capacity on exit decisions are mixed. That is small and/or inefficient firms can be "shaken" out earlier, but in declining markets it may be that small firms "stakeout" the market with the result that larger firms exit first. (Lieberman (1990), and Blonigen, Liebman, and Wilson (2007))

In this paper, we control for what have become standard exit related variables such as capacity, multi-plant ownership, and whether the plant (individual elevator) is a recent entrant (to capture vintage effects). But given the nature of the industry, we also introduce industry specific variables that may also have an

[6] There are other studies that reinforce these findings, covering a broad range of countries. These include Italy(Colombo & Delmastro, 2000), the United States (Bernard & Jensen, 2007), Belgium (Van Beveren, 2007), Sweden (Bandick, 2007), Japan (Kimura & Kiyota, 2006), New Zealand (Gibson & Harris, 1996), Chile (Alvarez & Görg, 2009), etc.

effect on elevator exit. Since elevators are part of an extensive grain supply chain in Canada, their individual relationships with the grain transportation market can also influence their long term viability. In this sense, rail transportation in particular represents a vertical linkage connecting individual grain elevators to final markets. Plants/elevators supported by well developed transportation infrastructure would seem less likely to exit the market. Some elevators now also have considerable capacity to load hopper cars, which means they can ship large quantities, leading to lower negotiated freight rates than elevators with smaller capacities. We capture this effect with a rail carload capacity variable and find that it has a significant negative effect on the probability of a given elevator exiting the market. In addition, we also introduce a set of variables to account for differences in local demand and supply conditions, as well as spatial competition from other proximate elevators. The latter measure is a measure of capacity of other proximate elevators (calculated using weights inversely related to distance to the neighboring elevators).

There is a limited amount of research that applies to vertical relations in firms' exit decisions. As an example, Chen (2002) use a duration model to find that vertical integration reduces the likelihood of survival for US petroleum plants. de Figueiredo and Silverman (2012) examine exit rates in the US laser printer and manufacturing industry, finding that the density of a vertically related population has an adverse effect on exit rate. In our market, a clear vertical linkage exists between grain elevators and rail transportation. While some elevators can load only a few cars at a time, others can load dozens of rail cars in short order. Railroads will typically offer rate discounts for multiple rail car shipments over smaller rail car movements, a situation that places elevators having high car loading capacity with a substantial competitive advantage.[7] Due to this, car loading capacity is our measure of vertical linkage. In the case of grain elevators, Sarmiento and Wilson (2005) show that large elevators have a greater tendency to adopt a shuttle train technology than smaller ones, while the size of a rival has a negative impact on adoption decisions. They also find shuttle train technologies tend to be adopted in regions with high production and less competition.

The remainder of this paper is organized as follows: Section 2 contains some general background of the grain elevation industry in Canada and provides a review of related academic literature; Section 3 describes the data on Western Canadian grain elevators, providing more details about firm entry and exit during the study period; Section 4 presents the results of various econometric specifications and then discusses our findings; Section 5 concludes. At the very end of the paper, we have added an Appendix of the basic model estimates as Appendix A.

2.2 Background

Grain elevators have long been crucial to Canadian rural agricultural communities, not only in terms of the services provided but also in terms of employment, investment, local purchases, etc. Their primary role is to provide a convenient collection point for local grain, including storage and processing, as well as providing a connection with rail transportation that allows access to both domestic and international markets. In Canada, most grain operations on

[7] While there remains limited regulation on grain rates in Canada, the importance of rate discounts for larger shippers was buttressed with the so-called Great Northern Grain (GNG) case of 2007 (filed with the Canadian Transportation Agency (2007)) whereby a relatively small grain shipper (GNG) filed a rate complaint against a major railway. The case concerned the level of rate discounts being offered by railways at that time to elevators that had high loading capacities. GNG argued this was effectively undue discrimination against smaller grain shippers like itself who could not easily improve their loading situation. GNG won the case with the regulator, with the consequence that the major railways were forced to revise their grain rate schedules to render them less discriminatory against smaller grain shippers.

the Prairies are linked to the Class 1 railroad network. An interesting feature of Canadian elevators is their geographic dispersion. As discussed by Selyem (2000), Canadian Prairie towns were historically located approximately 6-10 miles apart, a distance based mostly on the limitation of transportation modes at the time, as well as the availability of farm inputs such as water and fuel. While much fewer in number today, grain elevators were and remain a significant business activity in many rural communities, especially in Western Canada. While rooted in history, change now characterizes the modern grain elevator industry in Canada. As an example, up until 2012, Canadian farmers sold their grain through the Canadian Wheat Board (CWB), the quasi-governmental agency acting as the sole marketer of Prairie grains destined for export from Canada.[8] While the CWB no longer exists, farmers still need to deliver and sell their grain through a licensed grain elevator company. In turn, grain companies gain comparative advantages through procuring grain in local markets, coupled with other factors such as elevator capacity, ownership, and rail carload capacity.

In this market, historical smaller capacity elevators have gradually given way to fewer but larger and more modern successors. Average elevator capacity has nearly tripled in recent years; in 1999 capacity was 6,558.34 tonnes but grew to 20,568.66 tonnes by 2016. Modern elevators offer higher-speed loading and unloading facilities, fast grain cleaning capabilities, unit train loading ability, and substantial storage space.

The history of the ownership of Canadian grain elevators is of interest and is important to understand the transition to the modern era. In early pioneer days individuals living in Western Canada's prairie towns often built their own grain

[8] https://www.thecanadianencyclopedia.ca/en/article/canadian-wheat-board

elevators (as co-operatives), and this process gradually brought in private grain companies as competition.[9] While much of the 20th century was dominated by the provincial (Alberta, Saskatchewan and Manitoba) pool elevator companies, by the mid-1990s falling costs of grain processing led to increased consolidation through mergers.[10] Within our data set, in fact, several major mergers occurred. These include Agricore United taking over United Grain Growers in 2001, and subsequently Agricore United was taken over by the Saskatchewan Wheat Pool in 2007. The latter merger created the largest grain handler in Canada, re-named as Viterra Inc.[11] Mergers in this industry have been mostly approved by Canadian competition authorities, but the latter merger was subject to some regulatory intervention due to competitive concerns (i.e. creation of more market power) in several areas, including at the Port of Vancouver. In summary, over the last 25 years, the grain elevator industry in Canada has been characterized by relatively small number of multi-plant (elevator) firms.

Bulk transportation has been a driving force in agriculture and in particular the grain industry in Canada. When the Canadian Pacific Railway (CPR) linked the Pacific province of British Columbia to the rest of Canada in 1871, one goal of the Federal government at that time was that this would help Canada expand its nascent grain export markets (Larson & James, 2007). Railroads were an important part of Western Canadian expansion, with farming and rail access going hand in hand for new immigrants looking to settle the West. To this end, grain transportation rates were regulated by the Canadian government ever since the subsidy given to help CPR build the final rail linkage to British Columbia. But

[9] https://en.wikipedia.org/wiki/Grain_elevator

[10] https://en.wikipedia.org/wiki/Grain_elevator

[11] https://en.wikipedia.org/wiki/Viterra

as a mature industry, by the 1970s the Canadian government had also started to subsidize various mainline rail upgrades to support on-going grain shipments. However, around this time the government also began to allow the two Canadian Class 1 railways to sell or abandon rail lines that were deemed to be uneconomic (Vachal & Bitzan, 1997).

Many of the abandoned lines were in fact so-called "grain dependent" branch lines, track that in some cases served long lines of individual town grain elevators through distant parts of the Prairies. The peak period of this abandonment process was between 1984 to 1996, where the total length of so-called grain-dependent branch lines in Western Canada dropped by about 14 percent (Thraves, 2007). These abandonments hastened the demise of the old and small wooden elevators. The trend was reinforced with the repeal of Crow rate regulations in 1996 (J. F. Nolan & Kerr, 2012) and then slowed when new and stricter rate regulations were introduced in 2000 as the railways and grain companies had streamlined into more efficient grain handling networks (Brewin, Schmitz, Nolan, & Gray, 2017).

Concurrently, the ownership of one major Class 1 railroad underwent dramatic change with the privatization of the formerly publicly owned Canadian National Railway in 1995. Prior to this, the operations of the multi-modal Canadian Pacific Limited had been devolved into five independent companies. This ultimately left two private railroads carrying Canadian grain at regulated rates. Change was visible in other ways. The average number of rail cars that could be loaded by remaining grain elevators increased with the consolidation and modernization of the grain handling system. For example, over the duration of the data set we found average car load capacity for all elevators nearly tripled from 22.97 to 62.84.

In comparing the industry in Canada and the U.S., it is worth noting that Canada is a considerably smaller grain producer, with a much greater focus on export markets, whereas the U.S. grain industry splits between domestic and export markets, with domestic grain markets dominating elevator operations in most states. To this end, previous research on the grain elevator industry and industry evolution focuses on the U.S. market. Relevant works include Frittelli (2005) who find that between 1980 to 1998, the number of farms decreased by 15%, but farm size increased by 11%; concurrently, the number of terminal elevators increased, but the total number of grain elevators dramatically fell, mostly due to country elevators exiting the market. Other research has looked into vertical relations within the grain industry in the U.S. Schmiesing, Blank, and Gunn (1985) find that increases in the use of (large) unit grain trains in turn gives elevators access to larger and in some cases more distant markets, improving their price efficiency. Huang (2003) determined factors affecting shuttle[12] adoption, and as of the late 1980s with changes to how railroads marketed to grain shippers, shuttle trains have been increasingly adopted by the elevator industry. Prater, Sparger, Bahizi, and O'Neil (2013) highlighted the importance of grain train shuttles to railroad efficiencies. Local grain elevators that have been unable to accommodate shuttle-train shipments (for example, because they had small siding or loading capacity) have mostly gone out of business. In essence, many believe that with the demise of the CWB, the Canadian grain elevator industry is becoming similar

[12]A type of bulk train movement. A shuttle is a dedicated set of 75 to 110 covered grain hopper cars that carry just grain from one destination to another. `https://www.up.com/customers/ag-prod/shuttle/index.htm`

to that in the US, but “is approximately 20 years behind the grain movement system now operating in the United States.(D. Wallace, 1997).” Finally, Vachal and Bitzan (1997) examine survey data of Canadian grain elevators. At that time, they concluded that industry parties were in fact expecting a declining number of elevators in Canada in the short run. However, respondents also expected an increase in production as well as in overall elevator capacity.

2.3 Data

The data we use contains information on all grain handling facilities in Western Canada that were licensed from 1999 to 2016 (through the Canadian Grain Commission). The data was downloaded directly from the Canadian Grain Commission website.[13] In total, 1346 elevators operated over this time period. Many elevators are observed over time which allows us to ascertain whether an individual facility exited the market (or not) and when it exited. For each elevator, the data contains details such as storage capacity, ownership, the town closest to the facility, the geographical coordinates of the town, the type of elevator, the railroad(s) that serve the elevator, as well as the type of elevator. As we shall see, due to the importance of proximate grain production to an elevator, we also added data on regional agricultural production.

Initial examination of the data highlights the major changes in the industry, especially at the beginning of the sample. As alluded to earlier, the number of primary grain elevators decreased markedly through the 1990s as a result of rail line abandonment and the repeal of Crow rates (discussed previously). Some of the decline of the 1990s is reflected in our data from 1999 to 2002 (Figure 1)

[13] `https://www.grainscanada.gc.ca/en/grain-research/statistics/grain-deliveries/`

and indicates the magnitude of the declines. But since 2002, the total number of elevators in the region has been relatively stable.[14]

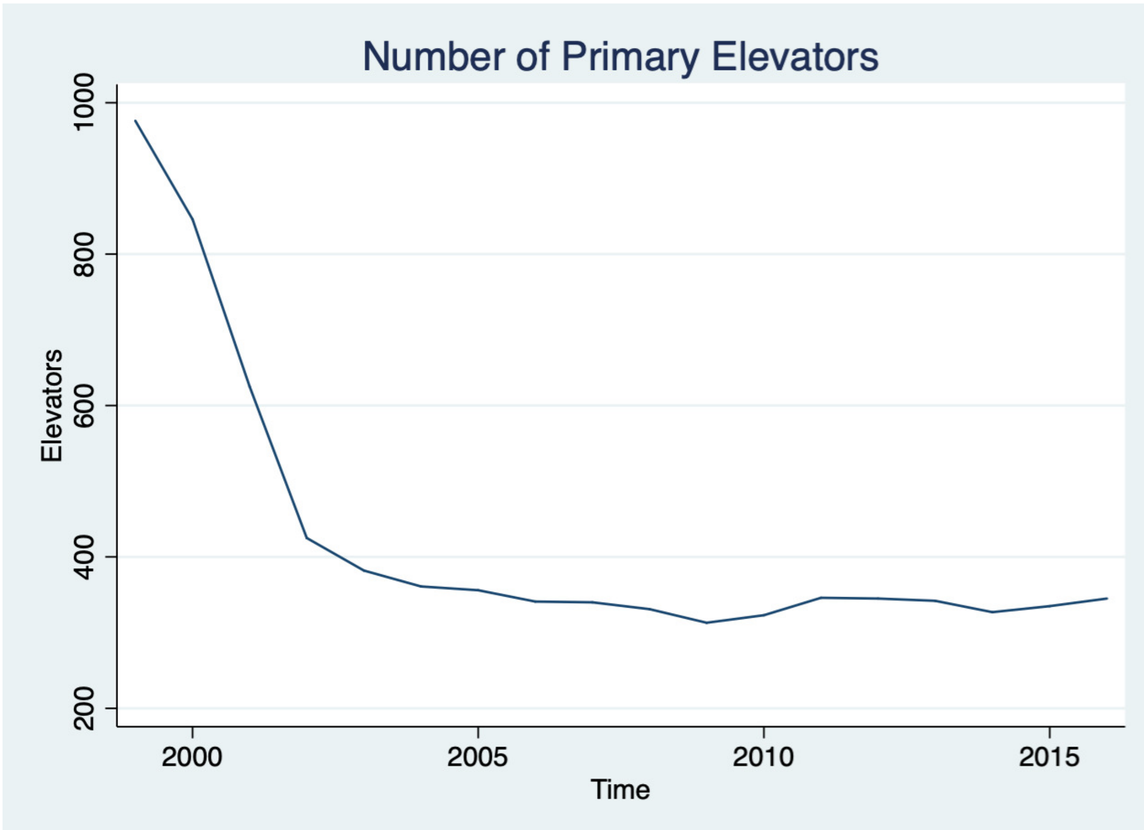


Figure 1. Number of Primary Elevators Over Time

The dramatic decline in elevators was the result of more cost efficient transportation options for grain movement as well as industry technological change. Ceteris paribus railroads prefer higher loading and shipment volume, creating the so called "unit trains".[15] These lead to lower unit costs. As a result of increased

[14]In our later analysis, we model the data using both the whole sample period as well as data after 2002, and the results are qualitatively equivalent as well as numerically similar.

[15]A unit train, also called a block train or a trainload service, is a train in which all cars (wagons) carry the same commodity and are shipped from the same origin to the same destination, without being split up or stored en route. `https://en.wikipedia.org/wiki/Unit_train`

freedoms to conduct line abandonments at that time, Canadian railroads chose to abandon numerous low-density branch lines on the Prairies. Concurrently, small-town grain elevators were gradually replaced by more dispersed but larger and more efficient terminals, which effectively caused farmers to truck their grain over considerably longer distances (J. Nolan, 2007). Since older grain elevators remain a nostalgic symbol for many western Canadians, some towns have succeeded in preserving elevators by switching them into museums or art galleries. But preservation is not the norm and most have been deteriorated or dismantled. Over the last 20 years, newer grain elevators have tended to have much larger capacities and are more efficient and durable than their predecessors.[16] These high-efficiency grain elevators not only facilitate loading/unloading grain more quickly than previously, but they also help maintain higher grain quality on a large scale (Simmins, 2004).

The dependent variable in our analysis is a discrete variable that reflects whether or not an elevator exits in the subsequent time period. This variable reflects a determination by the owning grain company that the long-run profit of a given individual elevator does not support keeping it open. In effect, these profits are assumed to be a function of the elevator capacity, whether the elevator is owned by a company that owns other elevators, whether the elevator entered the data after the first year of the data (i.e., new entrant[17]), rail car loading/siding capacities, and local demand and supply conditions, including the degree of spatial competition.

---

[16] `https://www.farmprogress.com/grain-handling/new-innovations-grain-storage-systems-higher-capacities-and-better-grain-quality`

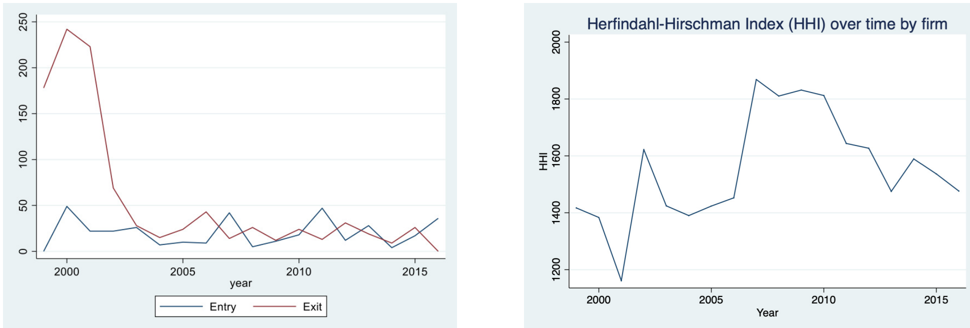
[17] In our case the first year of data is 1999. We do not observe when the plants that existed in 1999 first appeared. But, we are able to identify plants that first appeared after 1999 which is the basis for the entrant dummy and follows an approach used by T. Dunne et al. (2005).

The size of a grain elevator plays an essential role on exit behavior. Over time the average capacity has increased. In 1999, the average capacity was 6,558 tonnes, increasing to 16,723 tonnes in 2009, and to 20,568 tonnes in 2016 (Table 1).

Elevator ownership is a discrete variable that takes a value of one if it is the single elevator owned by a firm or zero if it is owned by a firm operating multiple elevators. The number of single ownership elevators is relatively small, but has increased modestly through the time period (Table 1). We also observe from the data that the status of grain elevator ownership (i.e self-owned-plant or a multi-plant firm) usually does not change given that the elevator remains in the market. Statistically, we only have a single observation that switched its (plant) ownership. In fact, this elevator was owned by a multi-plant firm at the beginning of the sample period, but later became a self-owned elevator, and went out of business in 2006.

Since we have the capacity of each licensed primary elevator throughout the sample, if an elevator is not part of the 1999 data then we record it as an entrant and, as such, these are more likely to be more technologically advanced than elevators that have been in the market longer. In principle, such elevators tend not to exit (at least immediately after entry), which is consistent with T. Dunne et al. (2005), who found that entry barriers are also exit barriers. The total number of new entrants since 1999 increased to 247 by 2016, as shown in Table 1. Figure 2 ((a)) illustrates the number of entrants and exits throughout the data. As noted earlier, there was dramatic exit from 1999-2003. Since then the number of exits has been relatively stable and remain about the same over time.

The number of firms in the industry has also changed over time. But, the change varies over time. In 1999, there were a total of 32 firms in our data



(a) Entrants and Exits (b) HHI

Figure 2. Entrant, Exits and HHI

increasing to 68 by 2011 and remaining somewhat constant since then. In 1999, the top three elevators owners in our data are Saskatchewan Wheat Pool (303 elevators with 1,523,880 tonnes capacity), Agricore Cooperative Ltd. (258 elevators and 1,371,140 tonnes capacity), and United Grain Growers (126 elevators and 820820 tonnes capacity). The number of elevators owned by top firms within the industry dropped largely while the total capacity increased. In 2016, the largest three elevator owners were Viterra Inc. (72 elevators and 1,884,570 tonnes capacity), Richardson Pioneer (58 elevators and 153,640 tonnes capacity), and Patterson Grain (28 elevators and 710,750 tonnes capacity) leading the market.

The number of firms is changing over time and has generally increased over the time span of the data, but the average capacity held by firms has also changed. The average capacity per firm, was initially large in 1999, fell dramatically over the next five years and then increased, leaving the concentration in the market somewhat mixed. As illustrated in Figure 2(b), the Hirshmann-Herfindahl Index (HHI) for the industry generally increased from 1999 to 2007, but has fallen somewhat since then.

Vertical linkage of a grain elevator to the transportation sector is measured through car loading capacities (in rail cars) associated with each elevator. Elevators that can load greater numbers of rail cars tend not to leave the industry. The average car loading number for each of the elevator has increased from 22.96 in 1999 to 62.83 in 2016 as shown in Table 1. From this table, it is also clear that surviving elevators enlarge their grain and car capacities.

Total agricultural production data[18] in the area was collected at the elevator station level.[19] We also have data on the third level of census division, or subdivision of Canada, and we aggregate total production in the subdivision to provide a measure of local demand for elevator services.[20] Additionally, we aggregate the total amount of elevator capacity in each of the subdivisions as a measure of the supply of elevator services. From Table 1, we find that agricultural production per subdivision, per elevator, and per unit capacity are all increasing throughout the sample period.

The data contain the geographical coordinates of each elevator. We use the geographic coordinates to first calculate distances between each of the facilities. Then using the distances between facilities, we construct an inverse distance weighted mean value of elevator capacity (Liu et al., 2021). The idea is that nearby elevators have a larger influence than elevators located further away. The general formula for inverse weighted means is given below. Different exponent values relate to the decay associated with distance. To this end, we considered $\frac{1}{2}$, 1 and 2. Further, we restricted the range of a fixed distance. In our case, the calculations were based on a distance of 20 miles, but we also considered distances that ranged

[18] `https://www150.statcan.gc.ca/n1/en/type/data?MM=1`

[19] The original data has the station geographic coordinates in the smallest administrative division in Canada, such as cities, towns, villages, townships, and parishes and etc. `https://en.m.wikipedia.org/wiki/Administrative_divisions_of_Canada`

[20] `https://en.wikipedia.org/wiki/Census_geographic_units_of_Canada`

from 10 to 150 miles. The formula is given by:

$$\hat{z}(\mathbf{x}_0) = \sum_{i=1}^{n} z(\mathbf{x}_i) \cdot d_{ij}^{-r} / \sum_{i=1}^{n} d_{ij}^{-r}$$

Our empirical findings are consistent across varying exponents and distances. The reported results are based on an exponent of $\frac{1}{2}$ and we have included all elevators in the sample within 20 miles of the reference point.[21]

Table 1. Descriptive Data of Primary Elevators

Variables	1999	2009	2016
Number of Elevators	976	314	345
Exit	178	12	0
Capacity (tonnes)	6558	16723	20569
Single Elevators	18	23	39
Entrant	0	151	247
Car Loading Capacities	22.96	54.87	62.83
Agriculture Production (thousands of tonnes)	63.5	210	280
Weighted Capacity of 20 mile range	7093.88	16705.81	20221.44
Agriculture Production Per Elevator (in subdivision)	3170	25852	34549
Agriculture Production Per Unit of Capacity (in subdivision)	0.486	1.79	1.78

2.4 Econometric Specification and Results

In this section, we examine exit behavior under a logit choice specification. In particular, we know that conceptually firms will exit if the long-run profits of the firm with exit are greater than the long-run profits of the firm if they do not exit. Our approach is common in the literature e.g., Blonigen et al. (2007); T. Dunne et al. (2005); Lieberman (1990), and K. S. Miller and Wilson (2018). As discussed in K. S. Miller and Wilson (2018), the approach can emerge naturally from common dynamic models of entry and exit e.g., Ericson and Pakes (1995) where firms make exit decisions based on expected future profits which are a function of their individual state, the state of their competitors as well as market conditions. In our

[21] We also explored different exponents (r=1 and 2) and examined different distances (d=10,20,$\cdots$,150 miles). These results are generally consistent with those presented through the rest of the paper.

model, we capture individual states with measures of capacity, whether the elevator is recent entrant, whether it is part of a multi-plant firm, the state of competitors is captured in inverse weighted capacity measures (spatial competition), and market conditions are captured in the local demand and supply measures. The specification allows us to identify characteristics that contribute to elevator longevity in this market. Effectively, for the ith elevator, we define the latent variable as Y^*

$$Y_i^* = \beta \times X_i + \varepsilon_. \tag{2.1}$$

As discussed earlier, we do not observe profits, but we do observe whether the elevator or firm exits, which is represented by Y_i.[22] The dependent variable equals one if the elevator exits the market at the end of the year, and has a value of zero if it did not.

The explanatory variables are represented by X_i, while β is a vector of parameters to be estimated. The variables considered include grain capacity, if the elevator is part of a multi-plant firm, if the elevator is an entrant, car loading amounts, agricultural production at the subdivision level, weighted capacity, subdivision capacity, as well as agricultural production per unit of capacity within the subdivision.[23] The detailed explanations are:

[22] Another approach that could be considered is based on hurdle rates which reflect the minimum rate of return to gauge where or not to pursue (or maintain) a project e.g., Brigham (1975), Liesch and Knight (1999), These are, of course, very comparable in concept in that decisions can be interpreted as the result of long-run profits. In our case, we simply do not have access to the cost and returns necessary to implement a hurdle approach, and instead follow the bulk of the literature to use the decision to reflect the long-run profits.

[23] The pairwise correlations amongst the right-hand side variable are generally quite small and the empirical results are generally robust across a wide range of specifications suggesting the any issues related to multicollinearity are non-existent or quite small. Further, a referee suggested that we consider the possible endogeneity of capacity. However, given that the market shares defined over sub-divisions are quite small (These average about 12 percent overall, and only 6 percent for existing plants, with only about a 2.6 percent median value). Given the relatively small market shares and the stability of coefficients, endogeneity does not appear to be a major issue.

LOG_CAP	=	Logged capacity of the elevator
ELEV_OWNERSHIP	=	One if the grain elevator is owned by a single plant firm; zero if the elevator is owned by multiplant firms.
ENTRANT	=	One if the grain elevator is not operating at the start of the observation year but enter the market later on; zero if the firm is operating at the beginning
LOG_CAR	=	Logged elevator car loading capacities
AG_Production	=	Total agriculture production within each subdivision level.
Weighted_CAP_20_Mile	=	Weighted average capacity with inverse distance of all elevators from the sample excluding the reference point, center elevator, within 20 miles
Subdiv_CAP	=	Total grain elevator capacity within a subdivision area
Ag_Production/Subdiv_Capacity	=	Average Agriculture Production per total elevator capacity in a subdivision area

We present two sets of results. Table 2 contains a base model with elevator, capacity, ownership, and whether the elevator is a new entrant (after 1999) with different fixed treatments (by time and subdivision). Column 1 does not contain any fixed effects and is the base specification; Column 2 includes time fixed effects; Column 3 includes subdivision fixed effects; and Column 4 contains both time and subdivision fixed effects. Table 3 adds measures of rail car-loading capacity (vertical linkages) and measures of regional demand, supply and local competition. All of the specifications in this model include time fixed effects.

The results reported in Table 2 are largely consistent with the previous literature.[24] That is, the effects of capacity and whether the elevator entered since the beginning of the data are both negative and statistically significant. This

[24]Given the large degree of exit in the first few time periods, we estimated the models excluding the initial three time periods where a lot of the exits were observed. The results are quite comparable with similar results and is attached in the appendix, Table A.1

means that larger and more modern elevators are less likely to exit. The effect on elevator ownership (a binary variable) has a negative coefficient, but is statistically significant only in Column 1 (which does not contain any fixed effects). A negative value here means that single ownership elevators are less likely to exit the market. While the overall results are mixed in terms of significance, as discussed earlier, the industry is still dominated by grain companies owning multiple elevators.

Table 3 incrementally adds rail car loading capacity, regional demand, regional capacity, and spatial competition. In all cases, the results related to capacity and whether or not the elevator is an entrant are quite similar to those found in Table 2. Both capacity and being an entrant have a negative effect on the likelihood of exit. In column 1, we add car loading capacity as a measure of vertical linkages. In this specification as well as the ensuing specifications this coefficient is negative and statistically significant. In column 2, we add total agricultural production of the subdivision in which the elevator is located, while in column 3, we add the total elevator capacity in the region. The former reflects demand conditions and the latter supply conditions. The coefficients on the demand measure are negative, while the coefficients on the supply measure are positive, which provide strong evidence that elevator in areas with strong demand and weak supply are less likely to exit the market. In Column 5, we add the ratio of the demand measure to the supply measure, and we find that the coefficient is negative, reinforcing our findings. Finally in Columns 4 and 5 we include our measure of spatial competition, that is the inverse distance weighted measure of rival elevator capacities. These, as discussed, capture the presence of rivals in the nearby area. This estimated coefficient in both specifications is positive and statistically significant, providing more evidence that the likelihood of exit is positively affected by the presence of spatial competition.

Table 2. Basic Model Results

	(1) exit	(2) exit	(3) exit	(4) exit
LOG_CAP	-0.941^{***}	-0.889^{***}	-1.017^{***}	-0.986^{***}
	(0.0474)	(0.0506)	(0.0516)	(0.0553)
Elev_Ownership	-0.364^{*}	-0.223	-0.308	-0.208
	(0.203)	(0.209)	(0.216)	(0.221)
Entrant	-1.307^{***}	-0.799^{***}	-1.390^{***}	-0.931^{***}
	(0.110)	(0.125)	(0.114)	(0.131)
Time Fixed Effect	✗	✓	✗	✓
Subdivision Fixed Effect	✗	✗	✓	✓
Constant	6.623^{***}	5.928^{***}	6.980^{***}	6.419^{***}
	(0.406)	(0.427)	(0.563)	(0.587)
N	7662	7317	7660	7315
Log Likelihood	-2552.8261	-2375.8799	-2489.3525	-2324.9201

standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Overall, multiple specifications are presented in Tables 2 and 3, yet there is a remarkable similarity among the various models. In all models (both Table 2 and Table 3), capacity has a negative effect on the likelihood of exit which is consistent with the notion that small country elevators are disappearing. Second, elevators that entered since the beginning of the data are less likely to exit. Our interpretation is that these elevators have been built with increasingly modern technology, so our results highlight the importance of elevator modernization as necessary in order to compete in this market. The results with respect to ownership are somewhat mixed. The estimated effects are negative in all specifications, but are mixed in terms of statistical significance. This may be the result, as mentioned, that the industry in Canada is dominated by multi-elevator firms.

Table 3. Vertical Linkages and Spatial Competition Model Results

	(1) exit	(2) exit	(3) exit	(4) exit	(5) exit
LOG_CAP	-0.660^{***}	-0.461^{***}	-0.483^{***}	-0.468^{***}	-0.328^{***}
	(0.061)	(0.061)	(0.062)	(0.062)	(0.068)
Elev_Ownership	-0.138	-0.413^{*}	-0.447^{**}	-0.371^{*}	-0.474^{**}
	(0.220)	(0.222)	(0.224)	(0.224)	(0.235)
Entrant	-0.821^{***}	-0.849^{***}	-0.887^{***}	-0.899^{***}	-0.847^{***}
	(0.127)	(0.129)	(0.131)	(0.131)	(0.134)
LOG_CAR	-0.371^{***}	-0.386^{***}	-0.391^{***}	-0.380^{***}	-0.267^{***}
	(0.062)	(0.062)	(0.063)	(0.063)	(0.067)
LOG(AG_Production)		-0.108^{***}	-0.113^{***}	-0.115^{***}	-0.274^{***}
		(0.008)	(0.008)	(0.008)	(0.019)
LOG(Subdiv_CAP)			0.271^{***}	0.231^{***}	0.219^{***}
			(0.052)	(0.053)	(0.054)
LOG(Weighted_CAP_20_Mile)				1.876^{***}	1.650^{***}
				(0.457)	(0.475)
LOG(Ag_Production/Subdiv_Capacity)					-0.255^{***}
					(0.027)
Constant	5.034^{***}	4.818^{***}	1.855^{***}	-14.450^{***}	-13.741^{***}
	(0.435)	(0.430)	(0.715)	(4.037)	(4.186)
Time Fixed Effect	✓	✓	✓	✓	✓
N	7297	7297	7297	7297	7231
Log Likelihood	-2340.7356	-2244.6586	-2230.1823	-2221.7229	-2138.394

standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Rail car loading capacity has a negative effect on the likelihood of exit in all specifications. This means that elevators set up for large scale rail shipment are less likely to exit and suggests that another avenue for elevator survival would be to invest in rail car capacity. The regional demand and supply measures are as expected. Elevators in regions with strong demand (large agricultural production) are less likely to exit, while elevators in regions with considerable capacity are more likely to exit. Finally, we introduced a measure of spatial competition and found it has a positive effect on the likelihood of exit. Hence, elevators in areas with large nearby competitors are more likely to exit the market.

The results in Table 3, Column 5 appear to be have the best overall results based on the log-likelihood. In addition, generally, the coefficient estimates are remarkably stable across specifications and consistent with prior expectations. The overriding conclusion for the collective results is that the industry has transitioned from one of small capacities and dated technologies to one of large elevator sizes and more modern technologies. Elevators with a strong tie to transportation (rail car capacity), in areas with strong demands (agricultural production) and little capacity are more likely to remain in the market than those with limited car loading capacity, those located in low demand, and high capacity regions. In all models, elevators that existed at the beginning of the data (1999) were more likely to exit than elevators that entered later. This is consistent with the findings of T. Dunne et al. (2005). Finally, the role of spatial competitors in elevator exit decisions makes intuitive sense and is underscored by the empirical results. In all pertinent specifications, the effect of the inverse distance measure of competitor capacity is statistically important and has a strong positive effect on the probability of exit.

It is clear from the results in tables 2 and 3 whether an entrant has a strong negative effect on exit. As discussed in section 3, entrants are more likely to be more technologically advanced than non-entrants. In tables 2 and 3, we simply have entrant or not. In tables 4 and 5, we replicate the results in table 2 and 3, but introduce a measure of entrant age. All qualitative results remain the same and are numerically similar. The effects of age are positive (as expected) and significant in all specifications, suggesting that plants that enter later are less likely to exit the market.

Table 4. Basic Model with Age Included

	(1) exit	(2) exit	(3) exit	(4) exit
LOG_CAP	−0.942***	−0.895***	−1.023***	−0.997***
	(0.048)	(0.051)	(0.052)	(0.056)
Elev_Ownership	−0.358*	−0.185	−0.293	−0.163
	(0.204)	(0.210)	(0.216)	(0.222)
Entrant	−1.352***	−1.017***	−1.512***	−1.931***
	(0.178)	(0.182)	(0.182)	(0.187)
Age	0.00836	0.0478*	0.0235	0.0584**
	(0.026)	(0.028)	(0.027)	(0.028)
Time Fixed Effect	✗	✓	✗	✓
Subdivision Fixed Effect	✗	✗	✓	✓
Constant	6.635***	5.978***	7.028***	6.523***
	(0.408)	(0.428)	(0.566)	(0.590)
N	7662	7317	7660	7315
Log Likelihood	−2552.77	−2374.41	−2488.97	−2322.83

standard errors in parentheses

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

2.5 Conclusion

In Western Canada, elevators remain a key part of the modern grain supply chain. But over the last few decades, the number of grain elevators on the Prairies has fallen dramatically. In this paper, we develop and estimate a model of elevator exit that encompasses relationships from the extant industrial organization literature, but also measures of vertical linkages to transport markets and the spatial setting of the elevator.

Our findings are largely consistent with prior literature in that we find the likelihood of exit is negatively affected by elevator size and also whether a given elevator was an entrant over the duration of the sample. But unlike prior literature, we find mixed evidence that multi-plant (elevator) ownership have a significant effect on exit decisions. To this basic specification, we also added

Table 5. Vertical Linkages and Spatial Competition Model Results with Age included

	(1) exit	(2) exit	(3) exit	(4) exit	(5) exit
LOG_CAP	-0.668^{***}	-0.469^{***}	-0.491^{***}	-0.476^{***}	-0.333^{***}
	(0.061)	(0.061)	(0.062)	(0.062)	(0.067)
Elev_Ownership	-0.0982	-0.363	-0.393^{*}	-0.317	-0.430^{*}
	(0.221)	(0.224)	(0.226)	(0.225)	(0.236)
Entrant	-1.061^{***}	-1.227^{***}	-1.280^{***}	-1.293^{***}	-1.129^{***}
	(0.184)	(0.188)	(0.190)	(0.191)	(0.193)
LOG_CAR	-0.370^{***}	-0.387^{***}	-0.391^{***}	-0.380^{***}	-0.266^{***}
	(0.062)	(0.063)	(0.063)	(0.063)	(0.067)
LOG(AG_Production)		-0.110^{***}	-0.116^{***}	-0.117^{***}	-0.275^{***}
		(0.008)	(0.008)	(0.008)	(0.019)
LOG(Subdiv_CAP)			0.274^{***}	0.234^{***}	0.220^{***}
			(0.052)	(0.053)	(0.054)
LOG(Weighted_CAP_20_Mile)				1.883^{***}	1.654^{***}
				(0.458)	(0.476)
LOG(Ag_Production/Subdiv_Capacity)					-0.254^{***}
					(0.027)
Age	0.0523^{*}	0.0836^{***}	0.0866^{***}	0.0868^{***}	0.0625^{**}
	(0.028)	(0.028)	(0.028)	(0.028)	(0.029)
Constant	5.092^{***}	4.910^{***}	1.910^{***}	-14.46^{***}	-13.72^{***}
	(0.437)	(0.431)	(0.717)	(4.045)	(4.194)
Time Fixed Effect	✓	✓	✓	✓	✓
N	7297	7297	7297	7297	7231
Log Likelihood	-2339.00	-2240.39	-2225.65	-2217.17	-2136.15

standard errors in parentheses

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

variables intended to capture vertical linkages to the freight transportation market and to accommodate local demand and supply conditions (including spatial competition). We find that our measure of vertical linkages has a strong negative effect on the likelihood of exit, indicating this is still a factor critical for elevator investment decisions. We also find that local measures of demand, supply and spatial competition matter in the exit decision. Simply put, as local agricultural production levels increase, spatial competition falls, and production relative to capacity increases, we find elevators were less likely to exit.

As an important industry to the economy of Western Canada, grain elevators have a long history, characterized by a series of ongoing changes to its organization. Based upon relatively recent industry data, the focus of this analysis was on obtaining a better understanding of grain company decisions about opening or closing individual grain elevators. In spite of various factors unique to this industry that have affected its evolution, our basic findings still strongly accord with prior work on industrial exit. To this end and due to the inherently spatial nature of grain elevator markets, we found that conditioning on market factors the degree of localized spatial elevator competition was a significant determinant of an exit decision. The latter raises concerns as per the continued evolution of the industry with respect to potential mergers/consolidation and the degree of local market power in various key production regions.

One factor that could not be accounted for in this analysis is the overall effect that elevator closures have had on primary agriculture in Canada. For example, in moving grain to their "local" elevator, Prairie farmers have seen this average distance grow considerably over the past 30 years (J. Nolan, 2007). Aside from the social costs of rural road damage (Larson & Nolan, 2007) and considering the benefits of economies of scale, elevator exit and consolidation have shifted some logistics costs over to individual farmers and has thus affected the long-term sustainability and structure of Prairie agriculture.

Unlike its close counterpart in the U.S., grain elevators in Canada have always been characterized by just a few co-operatives or private firms. With the demise of the Canadian Wheat Board in 2012, along with climate change, the lengthening of the crop season, more arable land and increased crop diversity in the region, the Canadian Prairie elevator industry is poised to see major changes in the near future. Our findings on exit decisions support the need for vigilance on the part of competition authorities in shaping how the industry moves forward. This must be done in order to avoid the monopolizing effects of economies of scale in elevation coupled with the identified objective of minimizing spatial competition.

CHAPTER III

WATCHING THE GRASS GROW: DOES RECREATIONAL CANNABIS LEGALIZATION AFFECT RETAIL AND AGRICULTURAL WAGES?

The synthetic procedure described in this chapter was a collaborative work where I did the data collection, curation, analysis, and visualization and Dr. Miller and I wrote and edited the manuscript together.

3.1 Introduction

The long-standing landscape of cannabis prohibition is rapidly changing. In the past decade, the median American voter moved from opposing to supporting legalization (Motel, 2015), more than a dozen U.S. states legalized the substance for adult use, and jurisdictions around the world loosened restrictions. One argument employed by supporters of legalization is the assertion that policy liberalization would lead to the creation of new jobs across multiple sectors (see e.g. Keys, 2020; A. Wallace, 2020). Indeed, according to Statistics Canada, the industry generated over 10,000 jobs within a year of Canada's federal-level legalization, with average hourly wages above the national average, and Barcott and Whitney (2019) estimate that the U.S. cannabis industry (including both medical and adult-use cannabis) directly employed more than 200,000 workers in 2019.

Cannabis, however, does not exist in a vacuum – the labor involved in cannabis production and retail is similar to that involved in other agricultural and retail markets and so cannabis legalization may induce workers to substitute between employers. Indeed, farmers of other crops in many areas have expressed concerns about the potential for upward pressure on agricultural labor wages as a consequence of adult-use cannabis laws (RCLs) (Smith, Powell, Mungeam, &

Emmons, 2019; Stoicheff, 2018; Valachovic, Quinn-Davidson, Stackhouse, Butsic, et al., 2019; Washburn, 2020). In this paper, we investigate these concerns by measuring the impact of recreational cannabis legalization on wages using data collected from the U.S. Census Bureau. We focus on Washington and Colorado due to their early adoption of legalization policies and therefore the longest post-legalization period during which to measure any changes in labor markets. We focus on agricultural and retail labor markets as those are plausibly the most likely to be affected by the opening of adult-use cannabis markets.

While this policy change may seem like a relatively clean quasi-experiment—both Washington and Colorado legalized adult-use through ballot initiatives and while the opportunity to generate tax revenue likely played a role in the success of these efforts, it is unlikely that the timing of these ballot initiatives or their implementation was driven by labor market conditions—and an opportunity for a differences-in-differences approach, we must overcome a number of challenges.

The first is data-related: cannabis is not separately categorized by the North American Industry Classification System (NAICS) and so we cannot measure the level of employment in the cannabis industry directly, but must instead infer it from changes in some larger category. Using data from the Quarterly Census of Employment, we identify NAICS categories that experience changes in the number of firms and employees that match state regulator data on cannabis firms. These categories differ across states as a consequence of differing regulatory frameworks. These data limitations create potential limitations in our ability to answer questions: if we observe a large increase in wages in the NAICS categories which contain cannabis firms, we cannot be certain that those higher wages are being paid to other workers in those categories without either additional assumptions or additional data. We address this in part by defining broader categories of retail and agriculture firms over which cannabis firms play a small role; if we observe an increase in wages in these broader categories, we can more reasonably conclude that incumbent firms are paying higher wage bills.

Second, given the spillover effects of legalization efforts both in terms of geography (Hansen, Miller, & Weber, 2020b) and in product space (K. Miller & Seo, 2021), as well as the mobility of (particularly agricultural) labor (Holmes, 2013; Thomas-Lycklama-Nijeholt, 2012), it is difficult to choose an appropriate control group *a priori*. We therefore follow Hansen, Miller, and Weber (2020a), who study the impact of cannabis legalization on traffic fatalities, and use a synthetic control approach. We create a control group by choosing weights for states without legal cannabis markets to match moments characterizing each state in the pre-legalization period. By comparing post-legalization employment and wages in the treated states to their synthetic controls, we can estimate the causal impact of legalization on these outcomes of interest.

Implementing this approach for the retail sector is relatively straightforward – the elements of retail sectors which drive labor market outcomes (i.e. household income and population density) do so in a consistent way across states (Blakely & Leigh, 2013; Neumark, Zhang, & Ciccarella, 2008). Agricultural sectors in different states, however, are significantly different due to variation in growing conditions and the characteristics of arable land. While many detailed industry measures are available, the set of measures changes frequently and often are not available for all states. Faced with a need to both select variables and impute certain values, we follow the approach of White, Reiter, and Petrin (2018) and implement machine learning techniques to accomplish these tasks algorithmically. In particular, we use LASSO for variable selection and classification and regression trees (CART) to impute missing values.

Our primary finding is a null result: we find little evidence of a significant difference in weekly wages per worker in the most directly substitutable NAICS categories. Furthermore, though our estimates are noisier, we do not find evidence of changes in weekly wages per worker in our broader definitions of the retail and agricultural sectors.

This paper adds to the growing literature investigating the legalization of cannabis for adult (recreational) use and its effects on outcomes thought to be related to cannabis consumption. Smart and Pacula (2019) summarizes many of the policy implications of cannabis legalization. Specific examples include studies on student performance (A. M. Miller, Rosenman, & Cowan, 2017), traffic fatalities (Aydelotte et al., 2017; Hansen, Miller, & Weber, 2020a), crime (Dragone, Prarolo, Vanin, & Zanella, 2019; Hao & Cowan, 2020; Hughes, Schaible, & Jimmerson, 2020) and the consumption of other "sin" goods and cannabis substitutes (Baggio, Chong, & Kwon, 2018; Chan, Burkhardt, & Flyr, 2020; Hansen, Miller, Seo, & Weber, 2020; Kerr, Bae, Phibbs, & Kern, 2017; K. Miller & Seo, 2021).

Our analysis hinges on the assumption that labor supply conditions are largely unaffected by cannabis legalization. Since all states which have legalized cannabis for adult use have previously legalized cannabis for medical use, the effects of both policies are relevant to our study. Ullman (2017) finds that medical cannabis laws (MCLs) reduce the number of absences due to sickness, while Sabia and Nguyen (2018) employ a synthetic control approach and find "no evidence that [MCLs] affect employment, hours, or wages among working-age adults". Nicholas and Maclean (2019) find evidence that MCLs "lead to increases in older adult labor supply, with effects concentrated on the intensive margin" and Ghimire and Maclean (2020) provide evidence that workers' compensation claims fall following the adoption of MCLs. On the adult-use side, Maclean, Ghimire, and Nicholas (2021) argue that RCLs increase Social Security disability claims, while Abouk, Ghimire, Maclean, and Powell (2021) find that workers' compensation benefits decline after RCL adoption. Taken together, these results suggest that our assumption is reasonable to a first-order approximation, though we discuss the way in which increases in labor supply driven by RCL adoption would influence our results in our conclusion.

More recently, the literature has begun to examine the cannabis industry as an economic entity of interest in and of itself and as a tool to investigate long-standing questions in industrial organization and policy design: Hansen, Miller, and Weber (2017) investigate the impact of a change in Washington's tax structure throughout the cannabis supply chain, Thomas (2018) considers the effect of Washington's licensing quota system, Hollenbeck and Uetake (2019) estimate the level and effects of market power in the industry, and Berger and Seegert (2020) use the cannabis industry to analyze the effects of financial exclusion on firms.

Within the literature, the closest effort to that of our own is that of Chakraborty, Doremus, and Stith (2020), who study the effects of Colorado's legalization on labor market outcomes at the county level exploiting the timing of retail entry across counties. Ultimately, they find, as we do, that while the entry of legal cannabis employers leads to increases in the number of employees in the relevant sectors, the impact on equilibrium wages is approximately zero. Relative to that work, we aggregate to the state level to avoid concerns about intra-state labor mobility, use states without legal cannabis markets as the bases for synthetic controls to avoid inter-state spillover effects, and add an additional treated unit (Washington).

We proceed in Section 2 by describing labor in the cannabis industry relative to other agricultural and retail industries. In Section 3, we describe our data on labor market outcomes and our methodology. In Section 4, we present our findings. We conclude in Section 5 with a discussion of the policy implications and suggestions for future research.

3.2 Labor in the Cannabis Industry

Relative to many commodity agriculture crops such as corn and wheat, cannabis production is labor intensive owing in large part of the dioecious nature of plants in genus *Cannabis*. Buds with high concentrations of the psychoactive cannabinoids tetrahydrocannabinol (THC) and cannabidiol (CBD) (among others) are only produced by female plants prior to pollination (Chandra, Lata, Khan, & ElSohly, 2017). Thus, in contrast to other dioecious agriculture operations, such as fruiting trees where males are necessary for fruit production, cannabis growers must identify and remove male cannabis plants from growing areas as even a small number of male plants can provide pollen for an entire crop, triggering seed production in females, a diminished set of flowers, and a corresponding reduction in cannabinoid production. This labor is necessary even when farmers plant "feminized" seeds or clones of female plants as the costs of a single male plant are high enough that growers use labor resources to identify and destroy male buds (see e.g. Schaneman, 2019). A relevant analogy in traditionally-legal agricultural products is hops (*Humulus lupulus*); producers of hops remove male plants to prevent pollination (Shepard, Parker, Darby, & Ainsworth, 1999).

The prevalence of indoor growing facilities complicates direct comparisons between cannabis and other plants. According to an industry report, 60% of legal producers operate indoor facilities, and 41% operate greenhouses – only 12% of firms grow cannabis in the outdoors alone (Cannabis Business Times, 2020). The use of indoor and greenhouse spaces allows for more precise control of the growing environment, leading to more potent output (Aizpurua-Olaizola et al., 2016), and enables production regardless of the outdoor agricultural season. However, the amount of labor hours needed per pound produced is likely higher for indoor and greenhouse operations than for outdoor operations (Caulkins, 2010).

After budding, plants must be harvested and trimmed of buds – a process which takes four to six hours per pound manually (Cervantes, 2006). While mechanized trimmers are available, hand-trimmers are able to extract higher quality buds from plants which can command higher prices from consumers; the majority of products sold to consumers (by revenue) consists of dried and cured buds and thus the visual appearance of the buds is directly relevant to demand (K. Miller & Seo, 2021). The remaining plant material undergoes extraction processes to produce concentrate and edible products which are generally sold at a lower price per weight of plant input. As a consequence, skilled trimmers can earn more than twice the average hourly wage of other laborers in crop, nursery, and greenhouse operations (Krissman, 2017).

These features of the cannabis industry imply that it is at least plausible that a small number of cannabis producers (relative to the number of other agricultural producers using greenhouses) could sufficiently impact the aggregate demand for agricultural labor to significantly change equilibrium wages. However, relative to other agricultural products, the market for cannabis labor is tightly regulated. In each state with an operating recreational market, individuals must pass a background check before working for a cannabis producer – and to pass that check, the worker must have legal immigration status and (in most states) must not have recent felony convictions related to Schedule I or Schedule II drugs. According to the U.S. Department of Labor, approximately 47% of the U.S. agricultural labor industry are undocumented immigrants, though agricultural industry sources estimate the share is closer to 75% (Jordan, 2020). If the labor markets are bifurcated due to immigration status, the effects of legalization on wages may be minimal at best. Furthermore, as the highest wages available within the cannabis industry are paid to workers with cannabis-specific skills, the substitutability of that labor (and therefore the upwards pressure on equilibrium wages) may be limited.

The process of retail sales of cannabis products also differ from most retail businesses. In most jurisdictions, psychoactive cannabis inventory must be strictly and securely separated from the sales floor, which is often required to be separated from pedestrian access through secure doors so that customer ages can be verified before entry. Inventory must be tracked in real-time for compliance with federal guidelines and state seed-to-sale traceability regulations. Audits are frequent and penalities for non-compliance include civil and criminal liability for firm owners and managers (Hansen, Miller, & Weber, 2018). These additional layers of security and related regulations imply that, relative to other retailers with similar footprints, cannabis retailers may demand additional labor hours.

Finally, though Colorado and Washington set up recreational markets in the same time period, the regulatory structures vary in ways relevant to our analyses; see Hansen, Miller, and Weber (2021a) for more details about the regulatory structures in the various states which have legalized cannabis for adult use. First, while Washington required vertical separation between production and retail, Colorado initially required retailers to produce 70% of the products they sell through vertically integrated production facilities, often located close to the retailer (Hansen, Miller, & Weber, 2021b). As a consequence, while firms in both Washington and Colorado set up production operations, production facilities in Washington, which were both more geographically dispersed and more specialized, arguably competed more directly with other greenhouse agricultural facilities for labor. Second, Colorado initially limited adult-use licenses to existing medical dispensaries, which may limit the number of new establishments entering at the time Colorado's market opened. Finally, Colorado allows home cultivation, which Washington bans. While this may affect demand for cannabis on the margin, we note that that to-date, the cannabis industry in Colorado has generated more revenue per resident than Washington's industry.

3.3 Data and Methodology

We begin our analysis of the relationship between cannabis legalization and labor market outcomes by obtaining labor market data from the Quarterly Census of Employment and Wages compiled by the U.S. Bureau of Labor Statistics (BLS). BLS categorizes employers according to the North American Industry Classification System (NAICS) – a system of 2-6 digit codes which classifies employers in narrowing groups according to their output or primary business activity. Our outcomes of interest include the number of establishments, the total number of workers, the total real wages, and the average weekly real wage per worker. We collect these outcomes at the NAICS-state-quarter level from 2000-2019, aggregate to the annual level, and deflate to 2019 dollars using the Consumer Price Index.

To capture time-varying characteristics of labor markets which may influence outcomes, we collect demographic data from the U.S. Census Bureau and Department of Education including state-level high school and college graduation rates, population density, the aggregate unemployment rate, and per-capita GDP. Agricultural labor markets differ widely from state to state due to differences in the characteristics of arable land and growing seasons and therefore to capture other time-varying characteristics of agricultural markets which may influence relevant labor market outcomes, we additionally collect state-year-level survey data from the National Agricultural Statistics Service from 2000-2015 and state-level data from the U.S. Censuses of Agriculture for 2002, 2007, and 2012 (i.e. pre-treatment covariates). A challenge we face in using this data is the prevalence of missing values which stem in part from changes in the survey questions from year to year. To create a panel data set for analysis, we focus on variables for which there are at least 30 state-level observations per year. These variables largely sort into clear topic areas: demographics, land statistics including rental prices, counts of farm establishments, and variables capturing output for corn, wheat, hay, and fruits and vegetables.

Despite this restriction, the data still contain many missing values complicating any analysis effort. Following White, Reiter, and Petrin (2012); White et al. (2018), we use the Van Buuren, Brand, Groothuis-Oudshoorn, and Rubin (2006) modification of the Classification and Regression Trees (CART) algorithm to impute missing values. The algorithm uses a Gibbs sampling procedure to generate a plausible value for each missing value. Key to our application, the algorithm uses "chained" imputation: for each unit of observation (i.e. each state-year observation), the most recent generated imputation for each column is used as a predictor for the next column to minimize bias (Michalowsky, Hoffmann, Kennedy, & Xie, 2020; Murray & Reiter, 2016; Van Buuren & Groothuis-Oudshoorn, 2010). In other words, suppose the vector of independent variables for observation t is $X_T = [x_{1t}, x_{2t}, \cdots]$. Suppose x_{1t} is known for some t but x_{2t} is missing. The algorithm uses a Gibbs sampler to draw a value from x_{2t} using the empirical distribution of x_2 conditional on x_{1t}. Now suppose x_{3t} is also missing for t. The algorithm uses both the observed value x_{1t} and the imputed x_{2t} to draw a value of the x_3 distribution conditional on both x_1 and x_2. Ultimately, in our primary specification, we impute 11% of the observation-variables for the agricultural analysis and none of the observation-variables for the retail analysis. We have re-estimated our models excluding imputed data and found similar results.

We next turn to the issue of variable selection. The number of potential control units (i.e. states other than Washington and Colorado) is less than the number of potential covariates. Instead of manually choosing covariates based on some prior hypothesis, which may be considered "cherry picking" (Ferman, Pinto, & Possebom, 2020), we use the LASSO algorithm to select appropriate covariates (Duncan, Ross, & Mikesell, 2019; Tibshirani, 1996). For each outcome variable, we fit prediction models to the pre-legalization data (i.e. data from 2000-2012) using the `glmnet` method of Friedman, Hastie, and Tibshirani (2010) and select the covariates with the highest frequency for each of the outcome variables.

The final covariate matrix X for our agricultural analysis includes "Barley for grain (acres)"; "Land in orchards (acres)"; "Snap beans harvested for sale (acres)"; "Cherries (acres)"; "Pears (acres)"; "Commercial fertilizer, lime, & soil conditioners (acres treated)"; "2000 Resident population 65 years & over, percent"; "2000 Savings institutions (FDIC-insured)-total deposits"; "2000 Civilian labor force unemployment rate"; "Federal Government expenditure-grants FY 2000"; "Federal Government insurance FY 2000"; "2000 Resident population: Black alone, percent"; "2000 Resident population: Two or more races, percent"; "2000 Resident population: Hispanic or Latino Origin, percent"; "2000 Resident population: total females, percent"; "Social security: retired workers-benefit recipients (Dec.) 2000"; "Corn grain production"; "Farm operations"; "Hay production"; "Labor hired wage rate ($ per hour)"; "Rent cash cropland expense ($ per acre)"; "Vegetable total production"; and "Wheat production". For our retail analyses, the covariate matrix includes "College Graduation Rate (percent)"; "High School Graduation Rate (percent)"; "Population Density (people per square mile)"; "Unemployment Rate (percent)"; and "GDP per capita". We also include the relevant outcome for stores in NAICS 453991 (Tobacco stores).

The agricultural census data is collected every five years – the last collection was in 2017. At the time of the last collection, only four states – Alaska, Colorado, Oregon, and Washington – had legalization cannabis for recreational use, and within those states, Colorado and Washington legalized earliest (voting in 2012, markets opening in 2014). To focus on the longest post-legalization period possible, we follow Hansen, Miller, and Weber (2020a) and focus on Colorado and Washington as the treated states. We further note that both Oregon and Alaska experienced significant supply issues in months immediately post market-opening (Andrews, 2017; Sacirbey, 2016) and thus any impact on agricultural labor is potentially more difficult to observe and/or interpret from the short post-legalization period available.

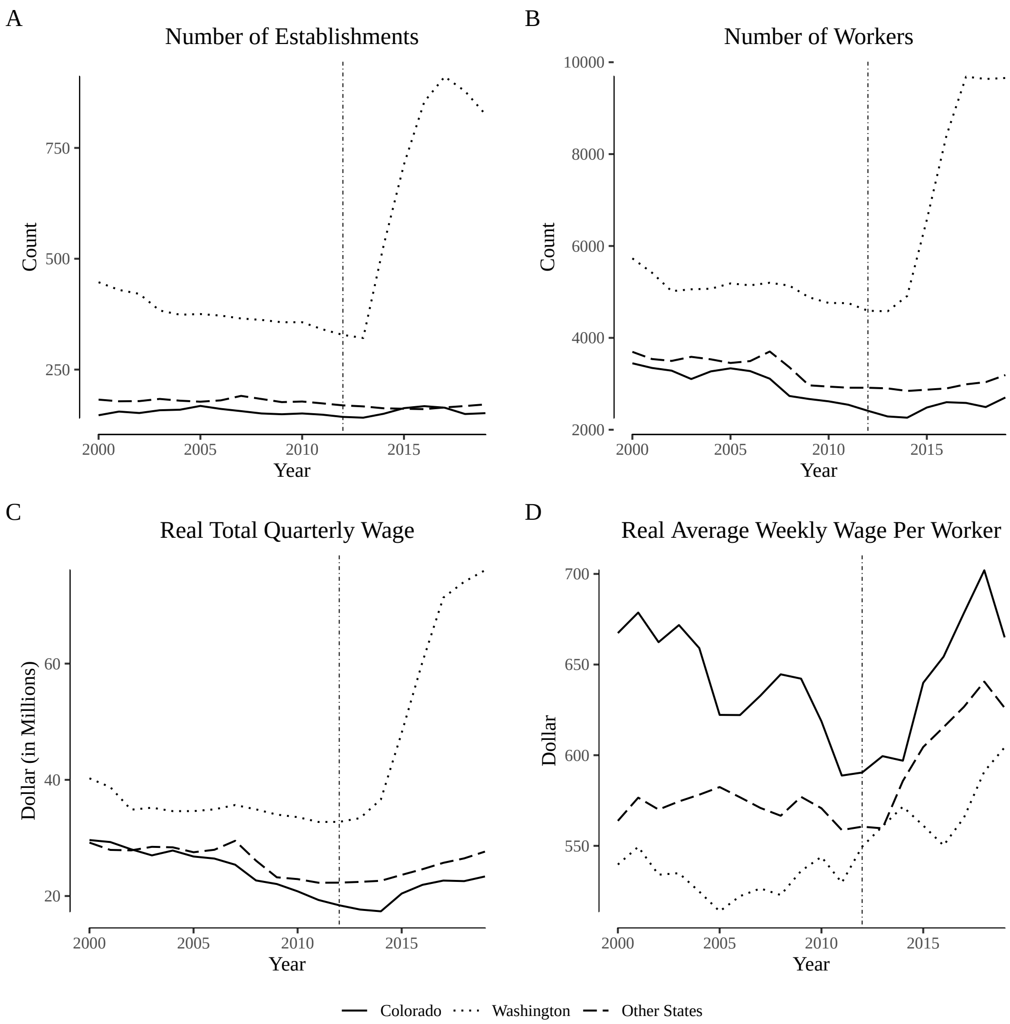
Figure 3 plots outcomes by year for Colorado, Washington, and the average of other states for the "greenhouse, nursery, and floriculture production" category

(NAICS 1114, the category containing cannabis production firms). Notably, the establishment count for Washington increased by roughly 500 between legalization and a peak in late 2015, which is similar to the count of cannabis production licenses issued by the state around the same time period as reported by Hansen et al. (2017). Washington experienced a similarly-shaped increase in the number of workers in the sector and the total wages paid, but those outcomes in Colorado and other states remained largely constant. Despite the increase in labor quantity observed in Washington, the real average weekly wage per week increased after legalization relatively uniformly everywhere.

Figure 4 reports analogous outcomes in the "store retailers not specified elsewhere" category (NAICS 453998, the category containing cannabis retailers). As with the agricultural sector, the establishment count in Washington increased by several hundred immediately post-legalization corresponding to descriptive statistics found in the literature (Thomas, 2018). Colorado also experienced an increase of roughly 200 establishments over the same time period. Increases of similar magnitude occured for worker counts and total wages paid in conjunction with the opening of these establishments. As in the agricultural sector, however, there are no clear patterns in the average weekly wage per worker; while the mean post-reform wage in Colorado is above the mean pre-reform wage, wages had begun increasing in the years prior to the passage of the ballot measure.

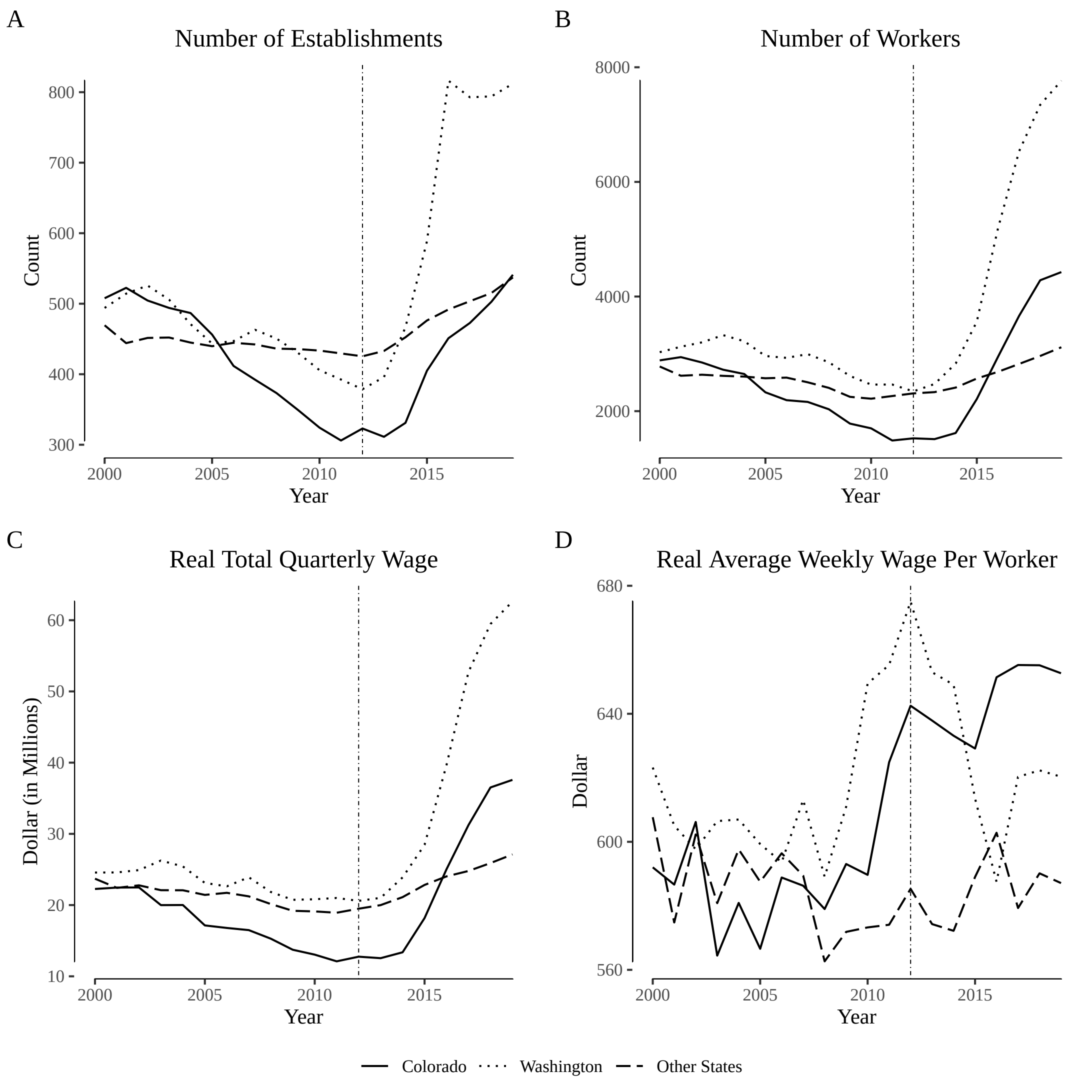
While the raw data suggest that the legalization of cannabis led to significant changes in employment in each state corresponding to their different regulatory structures, it is not clear that cannabis legalization caused these changes. Estimating a causal effect requires identifying an appropriate set of control units. While neighboring states might seem like a natural control group,

Figure 3. **Employment and wages for "narrowly defined" agricultural firms**



Notes: Data come from the Quarterly Census of Employment and Wages. We define "narrowly defined" agricultural firms as those within North American Industry Classification System category 1114 ("Greenhouse and Nursery Production"), which includes cannabis production firms.

Figure 4. **Employment and wages for "narrowly defined" retail firms**



Notes: Data come from the Quarterly Census of Employment and Wages. We define "narrowly defined" retail firms as those within North American Industry Classification System category 453998 ("Store retailers not specified elsewhere"), which includes cannabis retailers.

Hansen, Miller, and Weber (2020b) find evidence of substantial inter-state cannabis demand, and it is reasonable to believe that laborers may also move across state lines in response to cannabis legalization, particularly if cannabis producers are indeed offering higher wages. This is a particular concern for Washington, where many retailers are located close to the Oregon and Idaho borders.

To address this concern, we apply the synthetic control approach of Abadie, Diamond, and Hainmueller (2010, 2015); Abadie and Gardeazabal (2003). We construct synthetic control units separately for Washington and Colorado based on pre-legalization data (i.e. the covariates listed above plus the lagged value of the outcome variable) and then estimate the effect of cannabis legalization on our outcomes of interest by calculating the post-legalization difference between the outcomes for our treated states and for our synthetic controls. Our synthetic control units are convex combinations of non-treated states selected in such a way to match the pre-legalization outcomes. In addition to previous work on cannabis legalization and traffic fatalities (Hansen, Miller, & Weber, 2020a), the synthetic control approach has been used to analyze the effects of policy changes across a variety of domains, including economic liberalization (Billmeier & Nannicini, 2013), pediatric health (Bauhoff, 2014), tropical deforestation (Sills et al., 2015), foreign exchange rates (Chamon, Garcia, & Souza, 2017), tobacco policies (Chelwa, van Walbeek, & Blecher, 2017), and the effects of medical cannabis laws on labor market outcomes (Sabia & Nguyen, 2018) among many others.

We first select a “donor pool” of control units (i.e. states) which may be used to construct the synthetic control units. We start with all U.S. states and exclude any states which legalized cannabis and opened adult-use markets after 2012. We include Michigan as its first dispensary opened in December 2019, and thus any labor market effects are unlikely to be observed in annualized 2019 data. We also exclude states which are adjacent to the treated states to avoid spillover effects. While we present results using a donor pool which includes both states with and without legal medical cannabis markets, we have estimated separate models using only states with or states without these markets and found similar results.

For each treated unit $s \in \{\text{Washington, Colorado}\}$, we then select weights w_j for each of the control units j (with $0 \leq w_j \leq 1$ and $\sum w_j = 1$) to minimize the weighted difference between the synthetic control and the treated unit on the pre-treatment covariates identified above. The weight matrix V used to form the distance measure is chosen such that the mean square prediction error is minimized for the pre-intervention period following Abadie et al. (2010). We report the weights W^* chosen for each treated unit and outcome variable in Appendix A. Tables of covariate balance are available in Appendix B. We then obtain point estimates of the effect of recreational cannabis legalization with a standard differences-in-differences estimating equation. For outcome y for unit s (either a treated state or the synthetic control for that state) in year t, we estimate the parameters of

$$y_{st} = \beta_0 + \beta_1 * Legal_t + \beta_2 * Treated_t + \beta_3 * Legal_t * Treated_t + \epsilon_{st}. \tag{3.1}$$

To perform hypothesis testing, we use the "in-space" placebo tests described in Abadie et al. (2015). In particular, we apply the synthetic control model to each of our potential control units and interpret the results as placebos. We remove a small number of control states with particularly poor pre-treatment fit, though this does not affect our qualitative results. Plots of these placebos are available in the Appendix. For each outcome Y (and corresponding sequence of state-year outcome observations Y_{jt}), we then calculate the empirical distribution of the *ratio of the* *mean squared prediction errors* (RMSPE) where

$$\text{RMSPE} = \left(\frac{1}{T_0} \sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \right)^2 \right)^{1/2} \tag{3.2}$$

and T_0 is the positive number of pre-intervention periods. The p-value is then simply the fraction of placebo effect estimates which are greater than or equal to the effect estimated for the treated unit (Firpo & Possebom, 2016):

$$p := \frac{\sum_{j=1}^{J+1} \mathbb{1}\left[\text{RMSPE}_j \geqslant \text{RMSPE}_1\right]}{J+1}$$

Finally, it is plausible that, from the perspective of workers, jobs in the cannabis industry are substitutes for jobs beyond the narrowly-defined NAICS categories described above. We repeat this analysis for a broader set of categories taking advantage of the hierarchical nature of the NAICS inclusive of cannabis firms; for agriculture, we use "agriculture, forestry, fishing, and hunting" (NAICS 11) and for retail, we aggregate the "health and personal care stores" (NAICS 446), "general merchandise stores" (NAICS 452) and "miscellaneous store retailers" (NAICS 453) categories.

3.4 Results

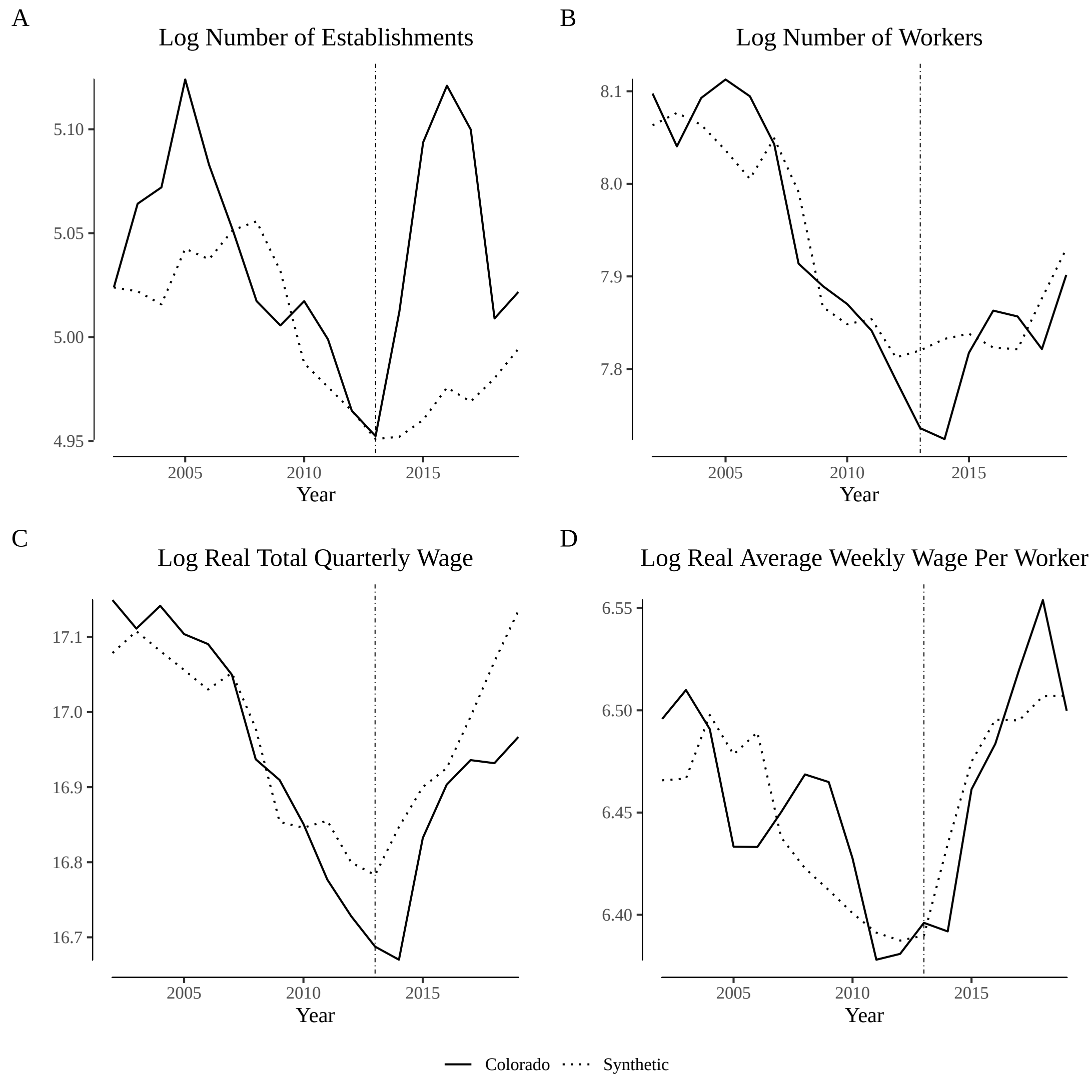
3.4.1 Narrowly-defined industries. Figure 5 illustrates agricultural labor market outcome measures in Colorado and its synthetic control unit (control weights are reported in Table B.2.1) for the "greenhouse, nursery, and floriculture production" NAICS category. Following Figure 3, Panel (a) illustrates the log of the number of establishments, Panel (b) illustrates the log of the number of worker, Panel (c) illustrates the log of the real total quarterly wage, and Panel (d) illustrates the log of the real average weekly wage. In general, the synthetic control closely follows both the trends and the level of Colorado's outcomes over the pre-legalization period. In the post-legalization period, the number of establishments temporarily grows relative to its synthetic control, but the number of workers tracks closely with its synthetic control, as do wages.

Figure 6 illustrates the analogous comparisons for Washington. As in Colorado, the synthetic control tracks closely with the Washington data in the pre-legalization period. However, the number of establishments increases significantly immediately after legalization, as does the number of works and (as a consequence), the total quarterly wages paid. Though the average weekly wage in Washington does increase post-legalization, the increase is also seen in the synthetic control.

Figures 7 and 8 repeat the exercise for outcomes for the "store retailers not specified elsewhere" NAICS category in Colorado and Washington, respectively. For Colorado, the synthetic control approach struggles to match the full volatility of the pre-reform data for the number of establishments and the number of workers. However, the method performs better (in a mean-squared-error sense) when matching per-reform average weekly wages per worker. Across outcomes, the synthetic control generally moves in the same direction as the Colorado data post reform, suggesting that other trends in Colorado contributed to the increase in establishments and workers seen in Figure 4. The synthetic control approach performs better for Washington, where pre-trends are closely matched for most outcomes.

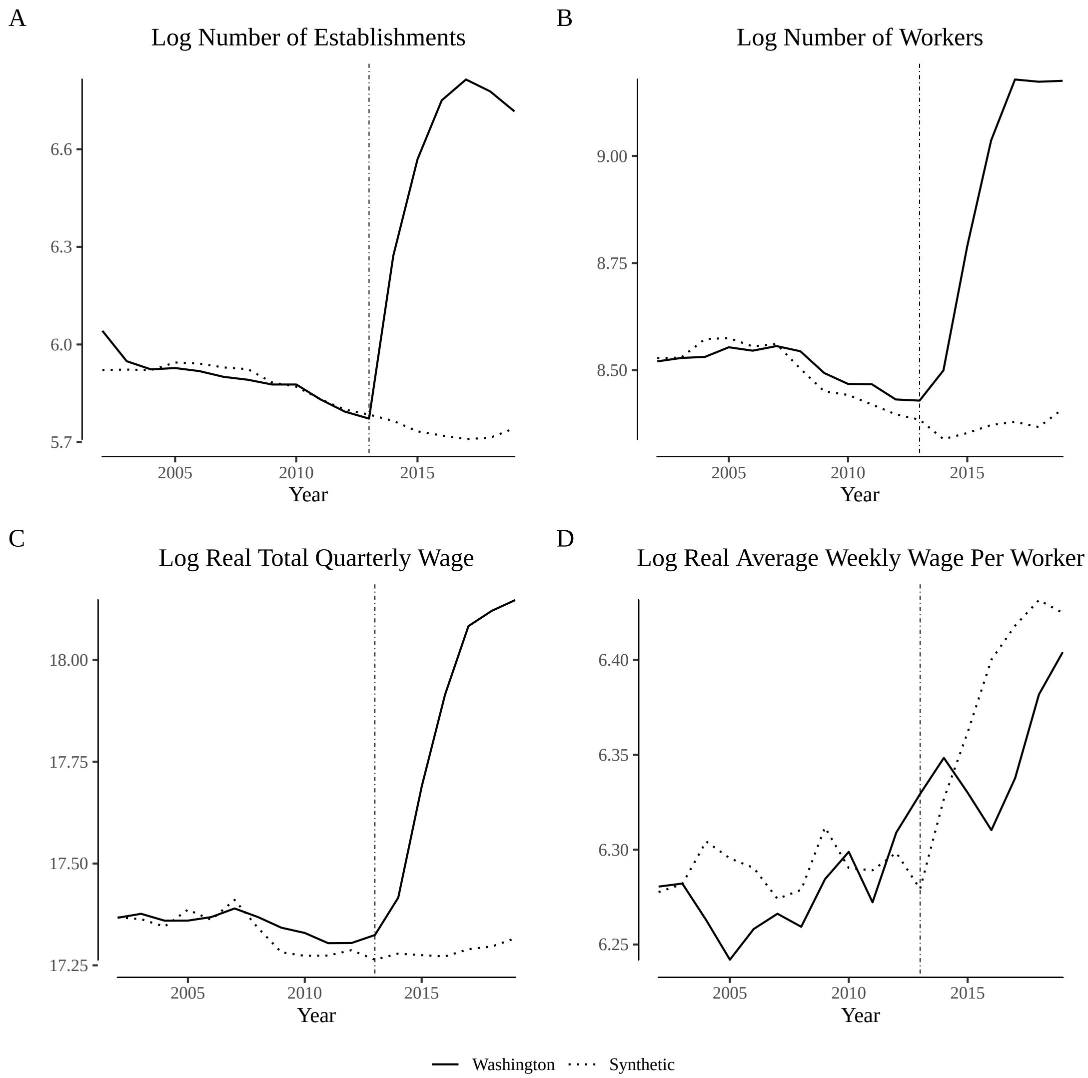
Point estimates of the effects seen in these Figures (i.e. estimates of β_3 in Equation (3.1)) are reported in Table 6. Several of the changes in the number of establishments, employees, and total wages are significant according to our placebo test at the 10% and 5% levels. However, the change in average weekly wage is either imprecisely estimated or negative for both sectors in both states.

Figure 5. **Comparing "narrowly defined" agricultural labor market outcomes in Colorado and its synthetic control**



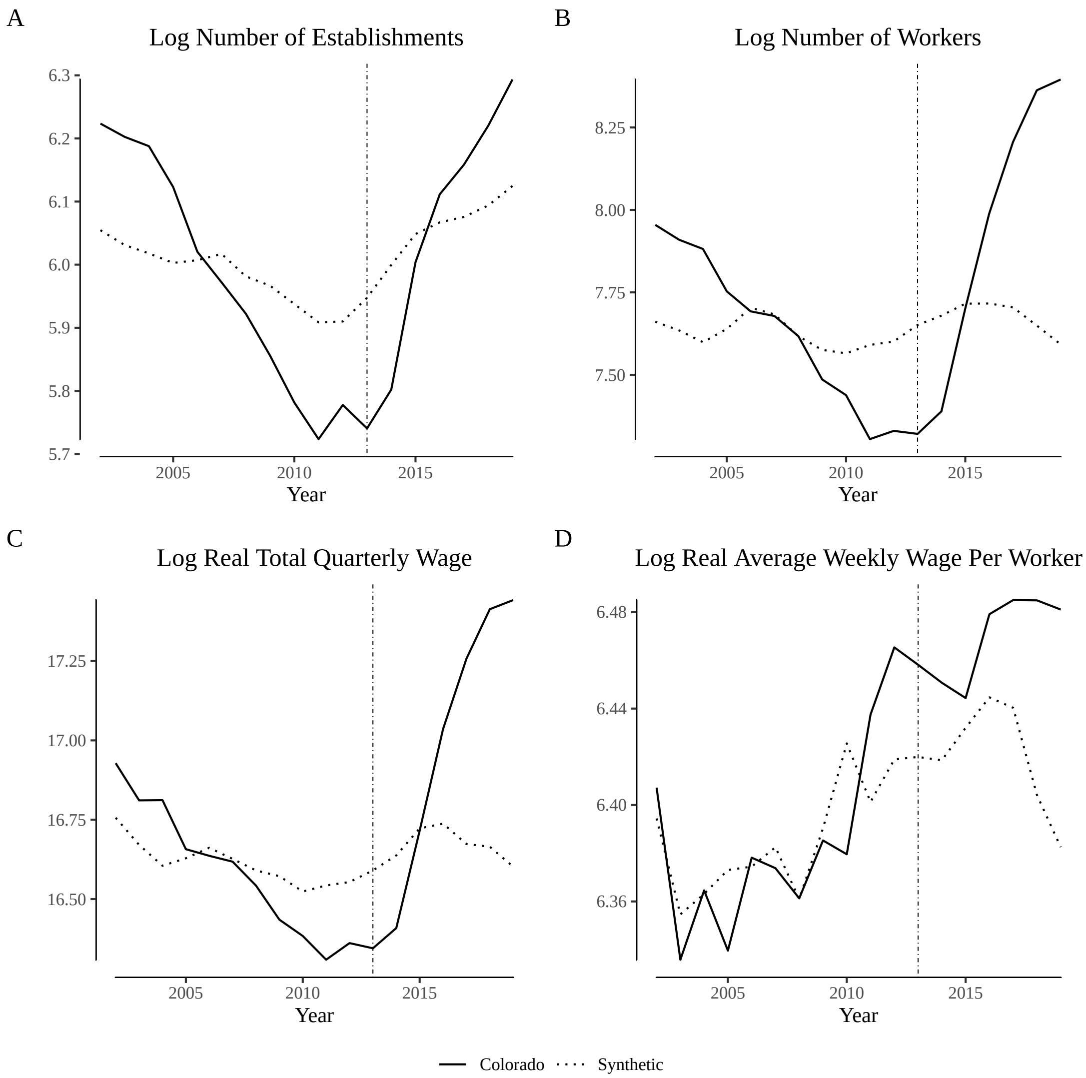
Notes: This figure depicts wage and employment outcomes for "narrowly defined" agricultural firms for Colorado and its synthetic control. We define "narrowly defined" agricultural firms as those within North American Industry Classification System category 1114 ("Greenhouse and Nursery Production"), which includes cannabis production firms.

***Figure 6.* Comparing “narrowly defined” agriculture labor market outcomes in Washington and its synthetic control**



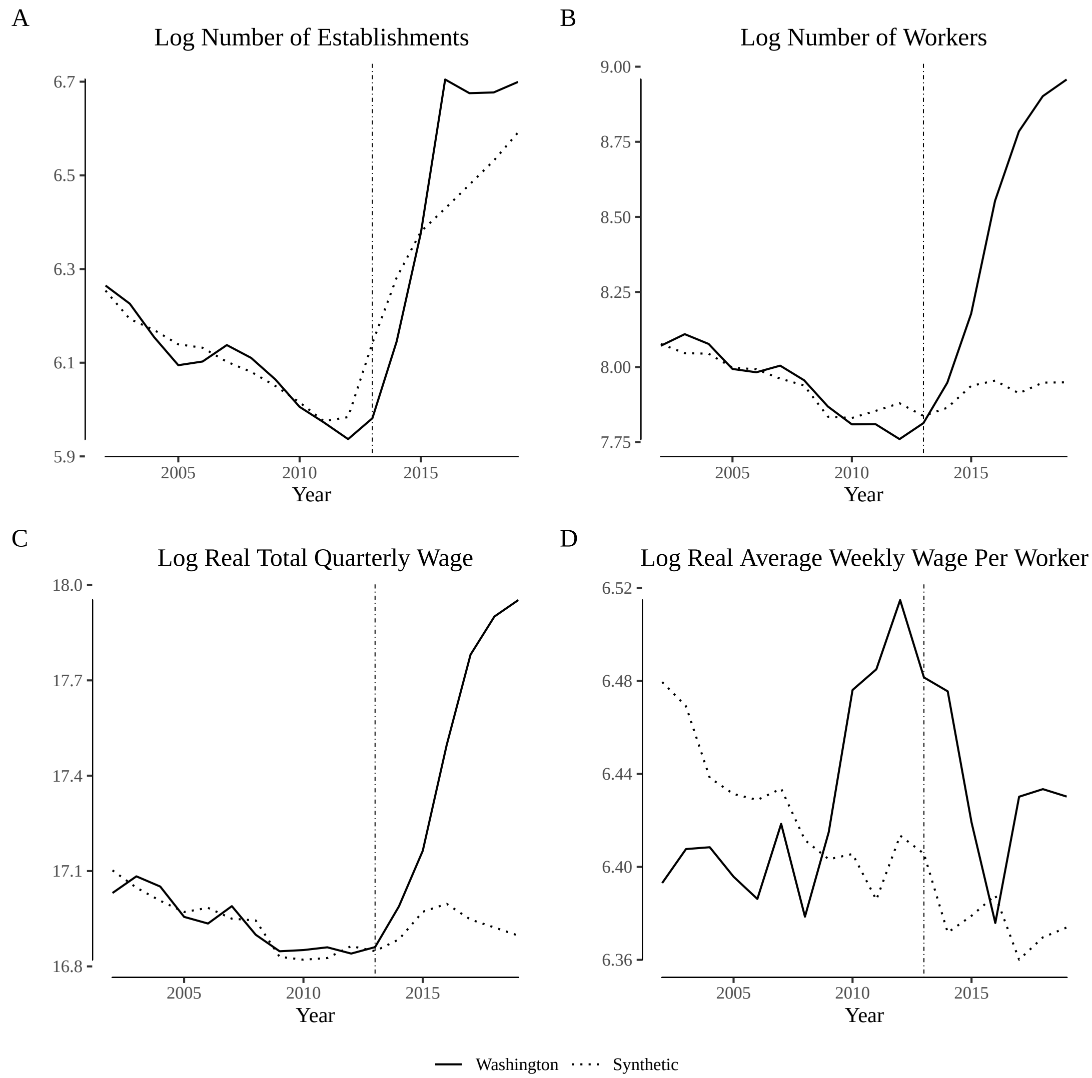
Notes: This figure depicts wage and employment outcomes for “narrowly defined” agricultural firms for Washington and its synthetic control. We define “narrowly defined” agricultural firms as those within North American Industry Classification System category 1114 (“Greenhouse and Nursery Production”), which includes cannabis production firms.

Figure 7. **Comparing “narrowly defined” retail labor market outcomes in Colorado and its synthetic control**



Notes: This figure depicts wage and employment outcomes for “narrowly defined” retail firms for Colorado and its synthetic control. We define “narrowly defined” retail firms as those within North American Industry Classification System category 453998 (“Store retailers not specified elsewhere”), which includes cannabis retailers.

Figure 8. **Comparing "narrowly defined" retail labor market outcomes in Washington and its synthetic control**



Notes: This figure depicts wage and employment outcomes for "narrowly defined" retail firms for Washington and its synthetic control. We define "narrowly defined" retail firms as those within North American Industry Classification System category 453998 ("Store retailers not specified elsewhere"), which includes cannabis retailers.

Table 6. Synthetic control estimates of the effect of recreational cannabis legalization on narrowly-defined labor market outcomes

	Colorado			
	Log number establishments	Log number of employees	Log total quarterly wages	Log weekly wage
Narrowly-defined Agriculture				
RCL	0.056**	-0.042	-0.113	-0.007
P-value	[0.030]	[0.303]	[0.303]	[0.576]
Narrowly-defined Retail				
RCL	0.000	0.220	0.306*	0.050
P-value	[0.818]	[0.152]	[0.091]	[0.212]
	Washington			
	Log number establishments	Log number of employees	Log total quarterly wages	Log weekly wage
Narrowly-defined Agriculture				
RCL	0.783*	0.516*	0.513**	-0.013*
P-value	[0.061]	[0.061]	[0.030]	[0.091]
Narrowly-defined Retail				
RCL	0.063	0.535**	0.525**	0.059
P-value	[0.152]	[0.030]	[0.030]	[0.606]

Notes: This table reports difference-in-difference estimates of the effect of recreational cannabis legalization (RCL) on labor market outcomes using synthetic controls for the treated units. *Agriculture* is the "Greenhouse and Nursery Production" (NAICS 1114) industry. *Retail* is the "Store retailers not specified elsewhere" category (NAICS 453998). P-values are calculated via a placebo test. Stars indicate standard significance levels: *10%, **5%, ***1%.

3.4.2 Broadly-defined industries. While the above results verify that the legalization of cannabis led to changes in the number of establishments and employees working in the categories which contain cannabis firms, they provide no evidence that legalization led to wage spillovers. Indeed, there is little evidence that legalization affected the Colorado labor market at all. One possibility is that although cannabis production facilities are coded as members of the green house and nursery sector, cannabis facilities do not compete with other members of that sector for labor. To explore this possibility, we first repeat the analysis for NAICS 11, which includes all "agriculture, forestry, fishing, and hunting" firms.

The results are reported in Table 7 under the headings for "Broadly-defined Agriculture" – the relevant Figures are available in the Appendix. It is important to note that the pre-treatment fit for Washington is generally poor. Ferman (2021) shows that the synthetic control model can be asymptotically unbiased even when the pre-treatment fit is imperfect. Relative to Table 6, the estimates for Washington are generally attenuated and more noisily estimated. For Colorado, the estimates indicate small and marginally significant increases in employees and total wages, though once again for both states there is no increase in average weekly wages.

The difference in results between Colorado and Washington is potentially driven by the vertical integration requirement in Colorado and the vertical dis-integration requirement in Washington. In particular, firms in Colorado may classify themselves completely as cannabis retailers, as opposed to cannabis producers. While it is unlikely that these firms would compete with other agriculture firms for labor (and indeed even if firms are classified in this way, we see no effect on agricultural wages in Tables 6 and 7), it is possible that firms organized

Table 7. Synthetic control estimates of the effect of recreational cannabis legalization on broadly-defined labor market outcomes

	Colorado			
	Log number establishments	Log number of employees	Log total quarterly wages	Log weekly wage
Broadly-defined Agriculture				
RCL	0.008*	0.108**	0.064**	0.007
P-value	[0.091]	[0.030]	[0.029]	[0.242]
Broadly-defined Retail				
RCL	-0.044	0.035**	0.055**	0.015*
P-value	[0.576]	[0.030]	[0.030]	[0.061]
	Washington			
	Log number establishments	Log number of employees	Log total quarterly wages	Log weekly wage
Broadly-defined Agriculture				
RCL	-0.021	0.312	0.369	-0.154
P-value	[0.091]	[0.333]	[0.242]	[0.667]
Broadly-defined Retail				
RCL	0.013	0.112	0.147*	0.014
P-value	[0.121]	[0.121]	[0.061]	[0.333]

Notes: This table reports difference-in-difference estimates of the effect of recreational cannabis legalization (RCL) on labor market outcomes using synthetic controls for the treated units. *Broadly-defined Agriculture* is the "Agriculture, Forestry, Fishing, and Hunting" (NAICS 11) industry. *Broadly-defined Retail* is the combination of 'NAICS 446 Health and personal care stores', 'NAICS 452 General merchandise stores', and 'NAICS 453 Miscellaneous store retailers'. P-values are calculated via a placebo test. Stars indicate standard significance levels: *10%, **5%, ***1%.

in this way have an effect on wages paid in the retail sector. We thus repeat the analysis once more for firms in related NAICS retail categories 446, 452, and 453. The results are reported in Table 7 under the heading "Broadly-defined Retail." As expected, the estimates are attenuated from the more narrowly defined category. We find limited evidence to support the hypothesis that weekly per-worker wages increased in Colorado (the point estimate of a 1.5% increase is significant at the 10% level) and no evidence to support such a hypothesis in Washington.

3.4.3 Robustness. In Table 8 we explore three alternative specifications, focusing on our primary outcome of average weekly wages per worker. In Column (2), we include only states with medical cannabis systems in our donor pool. In Column (3), we include only states with full prohibition of cannabis throughout our study period in our donor pool; the small number of states in this category limits the available inference. In Column (4) we follow the suggestion of Ferman and Pinto (2021) and repeat the analysis in levels while demeaning the outcomes. We do find potential evidence of a small increase in wages per worker in Washington in the broad retail category, though in context of the remainder of our estimates this is likely spurious.

3.5 Conclusion

Over the past decade, U.S. voters have undergone a rapid shift towards supporting the legalization of cannabis in some form and policy has changed to follow this support. These changes, however, have not come without frictions generated by broad society-wide concerns about (among other issues) public health and safety (Hall & Lynskey, 2016; Kilmer, 2019), educational outcomes (van Ours & Williams, 2015), and interactions with other substances (K. Miller & Seo, 2021). Other frictions have been caused by more immediate financial concerns:

Table 8. Results from alternative specifications of weekly wages per worker

	(1) Baseline	(2) Med. cannabis controls only	(3) Illegal controls only	(4) In levels, demeaned
Colorado				
Narrow Agriculture	-0.007	0.094*	0.077	14.22*
	[0.576]	[0.069]	[0.2]	[0.091]
Narrow Retail	0.050	-0.094	-0.044	-42.62
	[0.212]	[0.897]	[0.8]	[0.879]
Broad Agriculture	0.007	-0.010	0.055	-27.22
	[0.242]	[0.548]	[0.4]	[0.697]
Broad Retail	0.015*	-0.033	-0.063	-291.63
	[0.061]	[0.586]	[0.4]	[0.818]
Washington				
Narrow Agriculture	-0.013*	0.748*	0.695	397.8*
	[0.091]	[0.069]	[0.4]	[0.091]
Narrow Retail	0.059	0.158	0.198	-178.2
	[0.606]	[0.103]	[0.2]	[0.121]
Broad Agriculture	-0.154	-0.050	-0.020	-1200.23
	[0.667]	[0.419]	[0.6]	[0.697]
Broad Retail	0.014	0.025	0.012	52.10**
	[0.333]	[0.103]	[0.4]	[0.030]

Notes: Narrow agriculture is NAICS 1114, narrow retail is NAICS 453998, broad agriculture is NAICS 11, broad retail is NAICS 446, 452, and 453. P-values in brackets are calculated via a placebo test. Column (1) repeats results from Tables 6 and 7. In Column (2) we restrict the set of potential donor states to those with medical cannabis regimes. In Column (3) we restrict the set of potential donor states to those with full prohibition of cannabis throughout our study period. In Column (4) we use the level of average wages per worker per week (as opposed to the log wage) and demean the outcomes. Stars indicate significance levels: *10%, **5%, ***1 agricultural firms in areas with legal cannabis production have expressed concerns about upward wage pressures leading to reduced international competitiveness and domestic agricultural output. Indeed, Bampasidou and Salassi (2019) identify a number of instances of labor shortages in particular U.S. agricultural industries and regions around the time of the first successful cannabis legalization campaigns. At the same time, supporters of legalization have pointed to substantial employment within the nascent industry as a sign of success. Taken together, it is natural to suggest that cannabis legalization may be contributing to a highly competitive labor market from the perspective of agricultural employers.

We investigate the relationship between cannabis legalization and labor market outcomes across both the agricultural and retail sectors. Using a synthetic control approach paired with machine learning techniques including LASSO to select appropriate covariates on which to generate synthetic control units and CART for chained imputation of missing values, we ask whether equilibrium wages increased after legalization in Washington and Colorado, the first states to legalize. We find limited evidence to support this assertion; while the number of workers in the relevant sectors increased following the entry of cannabis producers and retailers, the wage per worker remained effectively constant.

Our results indicate that cannabis is not likely to be responsible for the broader changes in the agricultural or retail labor markets experienced during our study period. Indeed, others have pointed to changes in immigration policy including an increase in the intensity of enforcement (Escalante & Luo, 2017) and frictions in the H-2A guest worker program (Luckstead & Devadoss, 2019) as key contributing factors to changes in agricultural labor markets. On the retail side, aggregation in brick-and-mortar retailers (Neumark et al., 2008) and the increase in online shopping (Bram & Gorton, 2017) have been identified as key drivers of changes in retail employment outcomes. Relative to these broader labor market trends, cannabis legalization may well be the proverbial "drop in the bucket". At the same time, results from studies of MCLs suggest that increasing cannabis access may increase labor supply, though results from RCLs to this point have been mixed. If RCLs do increase labor supply, our null result could be explained by offsetting changes on the demand and supply side of the labor market. It is also possible that our results could be explained by the conversion of illegal production to legal production with minimal changes in the labor force (i.e. those who were engaged in illegal production became those employed by legal producers). More generally, if cannabis employment is particularly attractive to individuals who were not previously engaged in the labor market (including those who were unemployed or self-employed), our null result may well be expected.

These results are subject to a number of limitations which may be addressed by future work. While we have focused on the labor market motivated by anecdotal reports and popular press accounts, it is possible that the entry of adult-use cannabis firms may affect incumbent firms in the agricultural and retail sectors through other channels, such as competition for desirable real estate or within the product market. Our work is necessarily limited to a relatively short post-legalization period, and as cannabis production continues to grow, it is possible that other agricultural and retail firms may face competition from cannabis firms that differs from past experience. While many states have adopted regulatory frameworks similar to either Colorado's or Washington's, the details vary widely across dimensions including the number of licensed establishments, tax rates and licensing fees, quantity and potency limits, and out-of-state investment rules, amongst others (Hansen et al., 2021a). These differences may affect the cannabis industry's aggregate demand for labor across states and therefore the experience of agricultural and retail incumbents. Indeed, both Colorado and Washington allow counties and municipalities to ban entry by cannabis firms, and so there may be within-state heterogeneity. Finally, both Colorado and Washington had existing medical cannabis systems before opening their recreational markets. Our results therefore speak only to the incremental effect of recreational legalization; a state moving from full prohibition to a fully-legal regime may experience a larger effect.

Our study may give policymakers currently considering cannabis liberalization some indication that such a policy change is unlikely to significantly increase wage bills for existing retailers and agricultural firms in the short term. Indeed, legalization is likely to improve labor market outcomes for job-seekers, if only by slightly increasing demand for labor—though long-term cannabis use may affect labor market outcomes at the individual level (Sabia & Nguyen, 2018).

CHAPTER IV

ARE FIRMS' COST PREDICTIONS ACCURATE? EVIDENCE FROM MEDICARE ADVANTAGE

The synthetic procedure described in this chapter was a collaborative work where I did the data collection, curation, analysis, and visualization and Dr. Miller and I wrote and edited the manuscript together.

4.1 Introduction

How well can firms predict their costs? As costs play a key role in decisions about prices, benefits, and the adoption of innovative technologies (Baicker & Goldman, 2011; Dranove & Satterthwaite, 2000; Jena & Philipson, 2008), firms which consistently forecast their own costs inaccurately may make decisions that are consistently suboptimal *ex post* for both themselves and for their customers. While it is common knowledge that firms use the data they collect on their customers and specific market environment to make forecasts, it is difficult to evaluate performance on this dimension as forecasts and resulting performance evaluations are generally considered highly valuable trade secrets. As a consequence, previous empirical work on firm learning has largely focused on macroeconomic conditions. For instance, Coibion, Gorodnichenko, and Kumar (2018) use survey data from New Zealand and find that firms share a pervasive inattention to macroeconomics and have a wide set of beliefs especially regarding inflation; competition provides incentives to collect more accurate macroeconomic information.

We examine forecasts made by insurers participating in the Medicare Advantage (MA) market in the United States. MA is a "managed competition" (K. S. Miller, Petrin, Town, & Chernew, 2021) environment in which the government provides subsidies to firms who then compete with each other to provide health insurance plans to seniors that replace traditional fee-for-service Medicare benefits (TM) with (generally speaking) privately-administered managed care benefits. MA is substantial in its own right; the federal government spends more than $125 billion annually on payments to MA insurers in addition to the premiums paid by enrollees. As part of the regulatory "bidding" process detailed in Section 4.2, firms report detailed data on realized past medical costs and forecasted future costs across a number of categories. Firms then choose cost-sharing product characteristics relevant to those categories. Crucially, firms are incentivized to report accurate forecasts (i.e. the best forecasts they have) by the bidding system, as the forecasts constrain the set of possible products the firm may choose to offer. If a firm reports that it expects costs to be higher than it truly believes, it may be forced to offer an inferior product (i.e. higher copays and fewer benefits) than is optimal given its true beliefs.

We begin by combining the reported costs and forecasts of future costs to construct forecast errors. We construct the error both in total medical expenses and in the three largest individual categories: inpatient facility, surgery, and professional services. We show that forecast errors in these categories are smaller than the overall error but are positively correlated with each other. Thus, for a given plan-year, firms tend to "miss" in the same direction across service categories. We then document changes in the average forecast error as firms gain experience in the market and explore heterogeneity across firm experience (as measured by the number of enrollees in the firm's plans) and competitive environment. Using simple models of the prediction error, we find that the magnitude of the forecast error decreases as the firm gains experience and competitors, consistent with theories of learning.

Finally, we examine the relationship between these forecast errors and product characteristics. We show that, after accounting for time- and location-specific factors, firms which under-predict their costs offer plans with greater patient cost-sharing—i.e. products that are 'worse' from the consumer's perspective.

We contribute to an extensive literature studying firm learning and the evolution of equilibria over time (Börgers & Sarin, 1997; Fudenberg & Kreps, 1993; Hart & Mas-Colell, 2000; Milgrom & Roberts, 1991). This literature has branched in several directions, including using lab experiments to test and verify theoretical models (see e.g. Erev & Roth, 1998) and increasing the sophistication of the modelled learning process (see e.g. Camerer, Ho, & Chong, 2002; Crawford, 2003). Such refinements include adding a belief elicitation process into the model; Rutström and Wilcox (2009) find that eliciting beliefs can change games and that there is a significant difference between the strong-elicitation treatment, but not between the no-elicitation treatment and the less intrusive weak-elicitation procedure. In addition to the empirical work citied above, others have also studied firm learning and adaptation. Graham, Harvey, and Rajgopal (2005) find that upper managers main focus lies on meeting or beating earning benchmarks; specifically the quarterly earnings for the same quarter last year and the analyst consensus estimate. Most managers prefer to have smooth earnings and to maintain the stability of future predictions. Their information disclosure coincides with the effort to stability where they strive for clarity but also select news in a strategical manner. Doraszelski and Markovich (2007) use frequency response data from the new UK electricity system to see how firms compete and respond with prices and how the price converge to its equilibrium using an adaptive learning model. They find that especially during the middle phase, the best-fitting models are those in which firms more heavily weight recent rival behavior in forming beliefs about their rivals' bids and adaptively learn about the price elasticity of demand.

This work also builds upon a literature examining the behavior of firms and seniors within the MA system. Much of the work on the bidding system has focused on how changes in the subsidy rates offered by the government passes through to benefits (see e.g. Cabral, Geruso, & Mahoney, 2018; Duggan, Starc, & Vabson, 2016; Song, Landrum, & Chernew, 2013). We complement that work by investigating other potential drivers of firm behavior. Other work has focused on patient outcomes with a particular focus on comparisons between MA and the traditional Medicare system (see e.g. Figueroa et al., 2020; Meyers, Trivedi, Wilson, Mor, & Rahman, 2021; Park, White, Fishman, Larson, & Coe, 2020). We shed additional light on the way in which MA firms play a role in seniors' health.

The remainder of this paper is organized as follows: Section 2 contains some general background of the MA market and its bidding system; Section 3 describes the data and methodology of our practice; Section 4 presents the results in two different settings and then discusses our findings; Section 5 concludes.

4.2 Bidding in Medicare Advantage

Medicare Advantage (MA) was developed as a response to rising costs in the Traditional Medicare (TM) system.[1]Under TM, the government pays service providers according to a fee-for-service (FFS) schedule. Under MA, the government pays insurers a per-enrollee subsidy that is adjusted for observable risk factors (i.e. demographic characteristics and diagnoses) but that does not vary by realized medical expenditures. The enrollee experience is substantially different: TM enrollees may generally choose any provider for any service, whereas MA enrollees generally face restricted provider networks and referral requirements for specialists.

Firms offer plans on an annual basis. The bidding process for plans which will cover medical expenses[2] in year t begins in the spring of the previous year. First, the Centers for Medicare and Medicaid Services (CMS) releases 'benchmark' subsidy rates—the per-enrollee subsidy that CMS will pay firms for an 'average'

[1]One of the biggest goals to implement MA is to increase efficiencies of managed care and to save money: specifically to minimize the inefficiencies induced by the inevitable errors in TM's administered price system, by allowing the health plans and providers to negotiate prices or, in some cases, to integrate the finance and delivery functions (McGuire, Newhouse, & Sinaiko, 2011).

[2]These bids are based on the plans' estimates of the cost of providing required Medicare Part A, which covers most medically necessary hospital; skilled nursing facility; home health and hospice care; and Part B services, which cover medically necessary services by providers and other services deemed medically necessary, to cover an average beneficiary (Berenson, Sunshine, Helms, & Lawton, 2015).

risk enrollee—that vary at the county level. Firms then submit detailed proposals to provide MA plans. These 'bids' begin with detailed information about the firm's realized medical expenses per-member-per-month in year $t-1$ across more than 15 categories. Firms then report their projected expenses for each category for year t using the 'Bid Pricing Tool' released by CMS.[3] These estimated costs are to provide the minimal number of required benefits for an average mix of risks against the county-level benchmark. If a plan's bid was greater than the benchmark, it was required to collect the difference from its enrollees through a monthly premium. If it was lower, 75 percent of the difference was to be returned to enrollees in the form of supplemental coverage or lower premiums, and in an effort to reduce Medicare's obligations, the remaining 25 percent of the savings was to be returned to the Medicare program (McGuire et al., 2011). Thus, it is important for plan providers to have a reasonable estimate to plan ahead and hedge against any uncertainties for MA enrollees.

Figure C.1 illustrates an example of these submissions for a plan offered for the 2015 benefit year as entered into the Bid Pricing Tool in 2014. The left-hand columns represent the plan's realized costs for the 2013 plan year, and the right-hand columns represent the firm's projections of their costs for the 2015 plan year. This plan added dental coverage between 2013 and 2015, and so the firm did not have previous data on dental costs for this particular plan with which to form projections. Ultimately, this firm projected that the 'per-member per-month' cost of this plan would increase by $39.65 from 2013 to 2015, or 7.4% of the realized costs in 2013. In dollars, this increase is driven mostly by projected increases in inpatient and skilled nursing facility costs, as well as professional service costs.

A key question with respect to any reported firm projections, particularly those which the firm knows will eventually be released to the public, concerns the extent to which these reports reflect the actual beliefs of the firm i.e. the beliefs

[3] While firms are required to provide projections based on their past experience, these projections may be manually overridden.

which underlie the firm's actions in the market (see e.g. Coibion et al., 2018; Graham et al., 2005; Rutström & Wilcox, 2009). In this environment, incentives for accurate reporting come from the payment system. The projected costs are use to form a final 'bid amount' which is then compared to the benchmark subsidy rates. Plans with a bid amount that is higher than the benchmark must charge premiums to enrollees; supplemental premiums may also be charged if plans include benefits beyond those offered by TM. Firms that bid below the benchmark receive a portion of the difference as a 'rebate' that must be passed on the consumers through plan benefits. The rebate payment varies across firms and over time based on the CMS 'star rating' measure of insurer quality, which is a summary of multiple measures of service quality such as the fraction of members receiving influenza vaccinations, the 30-day hospital readmittance rate, and enrollee qualitative assessments of care quality. After taking into account risk adjustment, we can write the rebate payment as a function of the bid b_{jt} and the plan-level benchmark $B_{jt} = B_{mt} \times \phi_{ft}$ with

$$reb(b_{jt}; B_{jt}, \lambda_{ft}) = \begin{cases} \lambda_{ft}(B_{jt} - b_{jt}) & \text{if } b_{jt} < B_{jt} \\ 0 & \text{if } b_{jt} \geq B_{jt} \end{cases}, \tag{4.1}$$

where λ_{ft} is the rebate percentage.

Crucially, any reductions in patient cost-sharing, relative to TM, *must* be 'paid for' with rebate funds. Cost-sharing plan features such as copays and deductibles have previously been shown to be determinants of demand for MA plans (see e.g. Curto, Einav, Levin, & Bhattacharya, 2021; K. S. Miller et al., 2021). As a consequence, the projections made by firms directly constrain the set of possible product characteristics offered to consumers along dimensions that are relevant for their (the firms') success in the market. If a firm reports that it expects to have much higher costs than it truly believes it will face, then it may be constrained to offer an less-than-optimally generous plan (i.e. higher copays/deductibles and fewer additional benefit categories) and therefore lose market share. If a firm reports that it expects to have lower costs than it truly believes it will have, it may not receive enough revenue (in the form of the government subsidy plus premium payments from enrollees) to pay for the care that it covers under the plan.

One of the major differences between MA and TM is that MA plans must provide all of the mandated insurance benefits of TM in exchange for a capitated monthly payment (Abaluck, Caceres Bravo, Hull, & Starc, 2021). In general, MA plans typically offer more generous benefits and lower cost-sharing than TM, whereas MA plans tend to have limited physician networks and require higher cost-sharing for costly services, perhaps because of these risk adjustment practices (Meyers & Trivedi, 2021). According to the Medicare Payment Advisory Commission (MedPAC), MA payers plans to continue to increase enrollment by offering extra benefits that beneficiaries find attractive. According to their 2021 report, bids slightly decreased to 87 percent of TM, a record low (Commission et al., 2021).

4.3 Data and Methods

We collect data on all bids and plans offered from 2007 to 2015 from public CMS files, including data generated by submissions made through the Bid Pricing Tool and (separately) detailed data on the provision of benefits. Following K. S. Miller et al. (2021), we focus on the market for individual insurance. We drop plans sponsored by employers and plans designed for individuals who are "dual-eligible" for Medicare and Medicaid.

For each plan, we collect enrollment, the average per-capita payment, the star rating, the deductible, the out of pocket limit, and copays for primary care visits, specialist visits, and 7-day hospital stays.[4] We construct the forecast error for each plan j offered in year t by comparing the projected costs reported at $t-1$ to the actual costs reported at $t+1$: $Err_{jt} = actual_{j,t+1} - projected_{j,t-1}$. In other

[4]A small fraction of plans use coinsurance cost-sharing mechanisms. We convert these to copayments using the Medicare Physician Fee Schedule and the American Hospital Association Annual Survey. Details and code are available upon request.

words, positive prediction errors indicate that firms underestimated their costs, while negative prediction errors indicate that firms overestimated their costs.

We construct the overall forecast error (i.e. the forecast error in the Total Medical Expenses line in Figure C.1), as well as forecast errors in the Inpatient Facility, Professional, and OP Facility - Surgery service categories. We choose these three categories as the largest components of plans' medical expenses. We also construct the percent forecast error as $Err_{jt}/projected_{j,t-1}$.

We construct four additional covariates from these data. First, we identify the state in which the plan has the most enrollees. Second, we construct the total previous enrollment at the contract-level to capture variation in the experience of the firm submitting the bid.[5] Finally, we construct two variables designed to capture the extent of competition: the share-weighted average number of contracts and plans offered by competitors in each plan's service area. That is, we weight the number of competitor contracts/plans offered in each county in which the plan offered by the share of the plan in that county.

Table C.1 provides summary statistics on our $11,670$ plan-year observations offered from 2008-2013 (i.e. after calculating forecast errors). On average, firms over-estimate their costs by \$270 per-member-per-month (PMPM), though the interquartile range include firms who underestimate their costs. The average prediction error in inpatient facility costs is \$103 PMPM, 38% of the average overall prediction error. Professional and outpatient surgery facility costs are overestimated on average by \$77 and \$16 PMPM, respectively. Over 75% of the plan-years in our data feature zero deductible and a limit on out-of-pocket expenses. A few plans feature zero copays; most feature positive copays for primary

[5]Firms sign contracts with CMS to offer potentially several plans; plans under the same contract generally offer similar provider networks and are available in similar geographies. Large firms (e.g. Aetna, Blue Cross Blue Shield, etc.) may have multiple contracts with CMS to offer plans in different areas; we calculate enrollment at the contract-level instead of the firm level to capture the possibility that the firm's experience may vary locally in ways that affect their ability to accurately forecast costs.

care visits, specialist visits, and hospital stays. Roughly one quarter of plan-years did not receive a star rating from CMS, either because they were too new or because not enough data was available for CMS to evaluate the plans. The median plan is in a competitive environment featuring more than 10 firms offering approximately two plans each.

Table C.2 reports correlations between the overall forecast error and the three components discussed above. All correlations are positive and greater than 0.7, indicating that errors in the prediction process tend to compound across cost categories, rather than offsetting.

Table C.3 details the distribution of the overall forecast error across quartiles of our measure of experience in the top panel and our contract measure of competition in the bottom panel. With respect to experience, the biggest difference is between the first (lowest experience) quartile and other quartiles; the difference between (for example) the third and fourth (highest experience) quartile is minimal. This suggests that while there may be some benefit to experience, the benefit is quickly realized. This is perhaps reasonable if firms are using regression techniques to estimate future costs, as the power of such techniques does not scale linearly. With respect to competition, variation across quartiles is minimal – if anything, this cut of the data suggests firms' prediction errors increase in magnitude as the number of competitors increases. While one might expect firms to invest more in accurate predictions when the competitive environment is more saturated, the presence of other firms may increase the variance in costs as insurers compete to add providers to their network (Gaynor, Ho, & Town, 2015).

Figure C.2 illustrates the evolution of the distribution of forecast errors over time. While the interquartile range shrinks dramatically from 2008 to 2011, along with a decrease in the absolute value of the mean and median, these trends reverse in 2012 and 2013. We note that several provisions of the Affordable Care

Act affecting MA went into effect in around this time, which may have increased the aggregate uncertainty experienced by MA firms.

Our analyses broadly fall into two categories. First, we study the relationship between the prediction error and covariates of interest including our measures of experience and competition, taking into account the possibility that prediction errors may be serially correlated. We model prediction error outcomes y_{jt} as

$$y_{jt} = \alpha_y y_{j,t-1} + X'_{jt}\alpha_x + FX_t + FX_f + FX_s + \epsilon_{y,jt}, \tag{4.2}$$

where α_x is the vector of coefficients of interest (where X_{jt} includes a constant), FX are fixed effects (t indicates time, f indicates the insurer sponsoring plan J, and s indicates the primary state in which j is offered), and $\epsilon_{y,jt}$ are unobservable factors influencing the forecast error.

Second, we study the relationship between the last-period prediction error and the product characteristics chosen by the firm. We model product characteristic x_{jt} as

$$x_{jt} = \beta_y y_{j,t-1} + Z'_{jt}\beta_z + FX_t + FX_f + FX_s + \epsilon_{x,jt}, \tag{4.3}$$

where β_y is the coefficient of interest, Z'_{jt} includes the star rating and the average per-enrollee payment from the government, FX are the fixed effects described above, and $\epsilon_{x,jt}$ are unobservable factors influencing product characteristics.

We estimate the parameters of Equations (4.2) and (4.3) using OLS and report heteroskedasticity-robust standard errors.

4.4 Results

In this table we report the results of our analyses. We begin by exploring the relationship between prediction errors and plan-level observables and then continue by examining the relationship between product characteristics and prior prediction errors.

4.4.1 The relationship between prediction errors and plan-level observables. Table C.4 reports estimates of the parameters of Equation (4.2) when the prediction error is measured in levels. Column (1) reports estimates for the overall prediction error. The error is increasing in the previous forecast error, the number of competitors, and the age of the plan (measured as the number of years the plan has been in the market since 2007, the plan year that Medicare Advantage implemented risk adjustment). The error is decreasing in the number of competing plans, though this is estimated with slightly more noise. These patterns generally continue across the major cost components reported in Columns (2)-(4), though the sign on the estimated relationship between the prediction errors and the log of the total previous contract enrollment varies.

In isolation, these results suggest that prediction errors may worsen over time, due to the positive coefficients on both the past prediction error and the age of the plan. However, we note that the mean prediction error is significantly negative; thus these results may instead indicate that firms in our data are becoming more accurate. To investigate this possibility, we re-estimate Equation (4.2) where y_{jt} is the absolute value of the prediction error. Table C.5 reports the results. We estimate in Column (1) that last period's error, previous enrollment, the number of competitors, and the age of the plan are all associated with a decrease in the magnitude of the overall prediction error. This pattern is largely consistent across the major cost components in Columns (2)-(4). While the number of competing plans enters positively for several regressions, the effect size is smaller than that of the number of competitors, indicating that, overall, increased competition is associated with smaller forecast errors (in magnitude).

4.4.2 The relationship between prediction errors and product characteristics. Table C.6 reports estimates of the parameters of Equation (4.3) when the prediction error is measured in levels. Each column represents a distinct product characteristic: the deductible, the out-of-pocket spending limit, and copays for primary care, specialist, and hospital visits. We include the star rating (and relevant indicators) and risk-adjusted payments as covariates to control for plan quality and observable differences in mean costs across geographies (K. S. Miller et al., 2021). At the point estimates, increases in the last-period prediction error (i.e. as firms actual costs are more and more above their prediction costs), the deductible, out-of-pocket expenditure limit, specialist visit copay, and hospital visit copay all increase (though the hospital visit copay is estimated imprecisely). In other words, as firms realize costs are higher than they believed, they seek to pass-through a portion of those costs onto consumers (see e.g. Butters, Sacks, & Seo, 2020; Kim, 2021). Indeed, the negative coefficient on primary care copays (indicating that first realizing higher costs than expected decrease the patient-facing cost of primary care visits) is consistent with this hypothesis, as increasing primary care service utilization is thought to reduce health care expenses overall (see e.g. Starfield, Shi, & Macinko, 2005).

As above, while these results are informative about the *direction* of the relationship, they are less information about the impact of the magnitude of the prediction error. We therefore re-estimate the parameters of Equation (4.3) with the absolute overall prediction error as an independent variable. The results are reported in Table C.7. Across product characteristics, the estimated coefficients on the absolute prediction error are the opposite sign of those reported in Table C.6. This is perhaps best interpreted in light of the results reported in Table C.4. Firms with large prediction errors are likely to more accurately predict their costs in the future. If uncertainty had pushed firms to offer plans with high levels of cost sharing (i.e. large deductibles, etc.), a reduction in that uncertainty should push firms to offer more desirable plans (see e.g. De Vany & Saving, 1977).

4.5 Discussion and conclusion

The idea that firms can learn about and adjust to changes in market environments is at the heart of many common models of equilibrium (see e.g. Berry, Levinsohn, & Pakes, 1995; Ifrach & Weintraub, 2016). While there is an extensive literature developing the theory of firm learning, the empirical literature has largely been limited by a lack of revealed preference data on firm beliefs. We fill that gap by studying the MA market, in which health insurance firms are incentivized to accurately disclose their best estimates of their own future medical costs as part of a mandatory regulatory process. These disclosures are then released publicly some years after the plan year.

We document the relationship between these estimates along two directions. First, we consider the "input" side: what factors influence the accuracy of the forecasts? We find that as the experience of the firm increases, as the age of the plan increases, and as the competitive environment intensifies, the magnitude of forecast errors decreases. In other words, those factors which we might reasonably expect to positively affect firms' ability to forecast their own costs do so in the direction consistent with reasonable priors.

We then turn to the "output" side: How do accurate and inaccurate forecasts influence product characteristic decisions. We document two stylized facts. First, the more firms underestimate their own costs, the more they seek to pass their costs on to consumers in the future through cost-sharing benefit structures. Second, the larger the magnitude of past forecast errors (in either direction), the more firms seek to improve their plans by offering reduced cost sharing.

Taken together, our results paint a picture of firms behaving as is often implicitly assumed in dynamic models of imperfect competition. Even though the firms in our data consistently overestimate costs on average, at the plan level the predictions increase in accuracy, allowing firms to offer better products (i.e. insurance plans with lower out-of-pocket costs) to consumers.

We conclude by pointing out that significant gaps in our understanding of firm expectations remain for future research to fill. For example, firms generally must predict not just supply (cost) conditions but also demand conditions. These predictions may behave differently than cost predictions, particularly as they are made in the context of competitors who are also making similar predictions.

APPENDIX A

CHAPTER 2 APPENDIX

Table A.1. Basic Model Results After 2002

	(1) exit	(2) exit	(3) exit	(4) exit
LOG_Cap	-0.522^{***}	-0.525^{***}	-0.592^{***}	-0.598^{***}
	(0.0545)	(0.0569)	(0.0609)	(0.0640)
Elev_Ownership	0.155	0.208	0.229	0.263
	(0.210)	(0.214)	(0.228)	(0.232)
Entrant	-0.795^{***}	-0.632^{***}	-0.947^{***}	-0.799^{***}
	(0.124)	(0.133)	(0.135)	(0.146)
Time Fixed Effect	✗	✓	✗	✓
Subdivision Fixed Effect	✗	✗	✓	✓
Constant	2.397^{***}	3.118^{***}	2.740^{***}	3.469^{***}
	(0.489)	(0.517)	(0.752)	(0.780)
N	5214	4869	5153	4816
Log Likelihood	−1221.539	−1163.7757	−1180.4597	−1125.3187

standard errors in parentheses

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

APPENDIX B

CHAPTER 3 APPENDIX

B.1 Abbreviations

Table B.1. List of Abbreviations

Abbrev.	Meaning
BLS	Bureau of Labor Statistics
CART	classification and regression trees
CBD	cannabidiol
GDP	gross domestic product
LASSO	least absolute shrinkage and selection operator
NAICS	North American Industry Classification System
RMSPE	ratio of the mean squared prediction errors
THC	tetrahydrocannbinol
US	United States

B.2 Additional tables and figures

Table B.2.1. Synthetic control weights assigned to each state for narrowly-defined agriculture labor market outcomes

	Log Number of Establishments	Log Number of Workers	Log Real Total Quarterly Wage	Log Real Average Weekly Wage Per Worker
Colorado				
Arizona	0.00	0.00	0.00	0.04
Georgia	0.19	0.00	0.38	0.00
Hawaii	0.00	0.25	0.00	0.00
Maryland	0.00	0.00	0.00	0.33
Minnesota	0.00	0.00	0.00	0.25
Montana	0.00	0.22	0.20	0.11
New Hampshire	0.19	0.00	0.00	0.00
South Carolina	0.00	0.00	0.00	0.23
Texas	0.45	0.53	0.41	0.04
Vermont	0.17	0.00	0.00	0.00
Washington				
Arizona	0.00	0.03	0.11	0.00
Connecticut	0.00	0.02	0.00	0.07
Florida	0.10	0.04	0.00	0.00
Georgia	0.00	0.00	0.00	0.37
Hawaii	0.00	0.05	0.01	0.07
Illinois	0.08	0.00	0.00	0.00
Kentucky	0.00	0.00	0.08	0.00
Michigan	0.54	0.40	0.05	0.00
Minnesota	0.28	0.37	0.00	0.00
Montana	0.00	0.00	0.00	0.14
South Dakota	0.00	0.00	0.00	0.15
Texas	0.00	0.08	0.66	0.19
West Virginia	0.00	0.00	0.09	0.00

Notes: The table provides the weights assigned to states for the synthetic controls used to estimate the "narrowly-defined agriculture" models in Table 6. All states except those which legalized cannabis during our study period and those bordering either Washington or Colorado were included in the pool of potential control units. Only states which received positive weight for at least one outcome are included in the table.

Table B.2.2. Synthetic control weights assigned to each state for narrowly-defined retail labor market outcomes

	Log Number of Establishments	Log Number of Workers	Log Real Total Quarterly Wage	Log Real Average Weekly Wage Per Worker
Colorado				
Georgia	0.07	0.11	0.19	0.00
Iowa	0.41	0.26	0.18	0.00
Kentucky	0.10	0.00	0.01	0.33
Louisiana	0.20	0.28	0.21	0.00
Minnesota	0.00	0.28	0.18	0.00
Mississippi	0.00	0.00	0.00	0.11
Missouri	0.00	0.00	0.01	0.00
New Hampshire	0.00	0.00	0.00	0.14
Pennsylvania	0.00	0.00	0.00	0.22
South Dakota	0.00	0.00	0.00	0.19
Texas	0.22	0.06	0.05	0.00
Wisconsin	0.00	0.00	0.18	0.00
Washington				
Connecticut	0.00	0.27	0.31	0.00
Illinois	0.00	0.38	0.36	0.00
Iowa	0.35	0.00	0.00	0.21
Michigan	0.33	0.04	0.05	0.07
Mississippi	0.05	0.00	0.00	0.00
New York	0.00	0.00	0.02	0.00
North Carolina	0.12	0.18	0.00	0.53
Pennsylvania	0.15	0.00	0.00	0.08
South Carolina	0.00	0.13	0.26	0.11

Notes: The table provides the weights assigned to states for the synthetic controls used to estimate the "narrowly-defined retail" models in Table 6. All states except those which legalized cannabis during our study period and those bordering either Washington or Colorado were included in the pool of potential control units. Only states which received positive weight for at least one outcome are included in the table.

Table B.2.3. Synthetic control weights assigned to each state for broadly-defined agriculture labor market outcomes

	Log Number of Establishments	Log Number of Workers	Log Real Total Quarterly Wage	Log Real Average Weekly Wage Per Worker
Colorado				
Arizona	0.19	0.26	0.12	0.00
Georgia	0.52	0.00	0.10	0.02
Hawaii	0.01	0.02	0.01	0.00
Kentucky	0.00	0.00	0.00	0.16
Minnesota	0.00	0.24	0.10	0.43
Montana	0.07	0.27	0.08	0.00
New Hampshire	0.00	0.05	0.22	0.00
South Dakota	0.21	0.00	0.00	0.00
Texas	0.00	0.15	0.29	0.02
Virginia	0.00	0.00	0.08	0.37
Washington				
Connecticut	0.00	0.00	0.00	0.07
Florida	0.00	0.86	0.74	0.18
Michigan	0.00	0.00	0.00	0.69
Minnesota	0.00	0.00	0.00	0.05
Montana	0.07	0.00	0.00	0.00
Texas	0.93	0.12	0.24	0.00

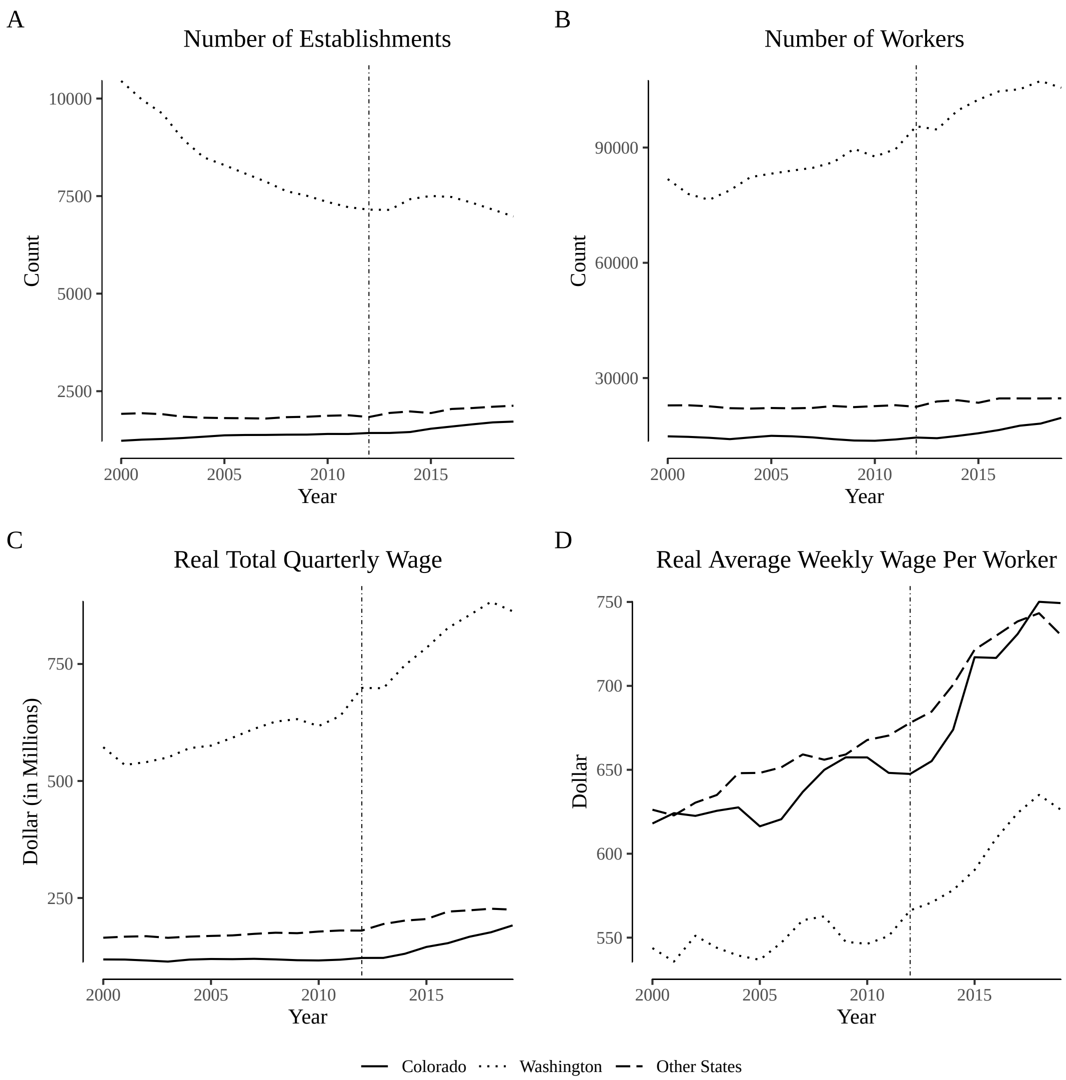
Notes: The table provides the weights assigned to states for the synthetic controls used to estimate the "broadly-defined agriculture" models in Table 7. All states except those which legalized cannabis during our study period and those bordering either Washington or Colorado were included in the pool of potential control units; Alaska and California were added to the pool for Washington due to the similarity in their agriculture, forestry, and fishing industries. Only states which received positive weight for at least one outcome are included in the table.

Table B.2.4. Synthetic control weights assigned to each state for broadly-defined retail labor market outcomes

	Log Number of Establishments	Log Number of Workers	Log Real Total Quarterly Wage	Log Real Average Weekly Wage Per Worker
Colorado				
Alabama	0.02	0.01	0.02	0.02
Arizona	0.02	0.18	0.11	0.02
Arkansas	0.02	0.01	0.03	0.03
Connecticut	0.23	0.00	0.00	0.01
Florida	0.00	0.00	0.00	0.01
Georgia	0.02	0.00	0.01	0.02
Hawaii	0.04	0.00	0.01	0.00
Illinois	0.01	0.00	0.01	0.03
Indiana	0.02	0.00	0.01	0.03
Iowa	0.02	0.01	0.02	0.05
Kentucky	0.02	0.01	0.01	0.02
Louisiana	0.03	0.01	0.02	0.05
Maryland	0.02	0.00	0.00	0.01
Michigan	0.01	0.00	0.01	0.02
Minnesota	0.02	0.01	0.03	0.07
Mississippi	0.02	0.01	0.03	0.03
Missouri	0.02	0.01	0.02	0.03
Montana	0.03	0.35	0.29	0.03
New Hampshire	0.03	0.00	0.01	0.03
New Jersey	0.07	0.00	0.00	0.01
New York	0.12	0.00	0.00	0.00
North Carolina	0.01	0.00	0.01	0.02
Ohio	0.01	0.00	0.00	0.02
Pennsylvania	0.01	0.00	0.00	0.02
South Carolina	0.02	0.00	0.01	0.02
South Dakota	0.03	0.04	0.04	0.06
Tennessee	0.02	0.00	0.01	0.01
Texas	0.01	0.34	0.28	0.19
Vermont	0.03	0.00	0.00	0.02
Virginia	0.02	0.00	0.01	0.04
West Virginia	0.02	0.01	0.01	0.03
Wisconsin	0.02	0.01	0.01	0.04
Washington				
Alabama	0.00	0.00	0.00	0.01
Connecticut	0.38	0.00	0.00	0.00
Georgia	0.00	0.00	0.00	0.01
Hawaii	0.00	0.00	0.00	0.09
Illinois	0.28	0.00	0.00	0.01
Iowa	0.00	0.24	0.01	0.00
Michigan	0.00	0.00	0.00	0.01
Minnesota	0.00	0.00	0.00	0.01
Missouri	0.00	0.00	0.00	0.01
New York	0.02	0.01	0.00	0.49
North Carolina	0.01	0.31	0.01	0.01
Pennsylvania	0.01	0.00	0.00	0.00
South Carolina	0.27	0.22	0.23	0.00
Texas	0.00	0.00	0.00	0.31
Vermont	0.00	0.00	0.00	0.01
Virginia	0.00	0.21	0.74	0.00

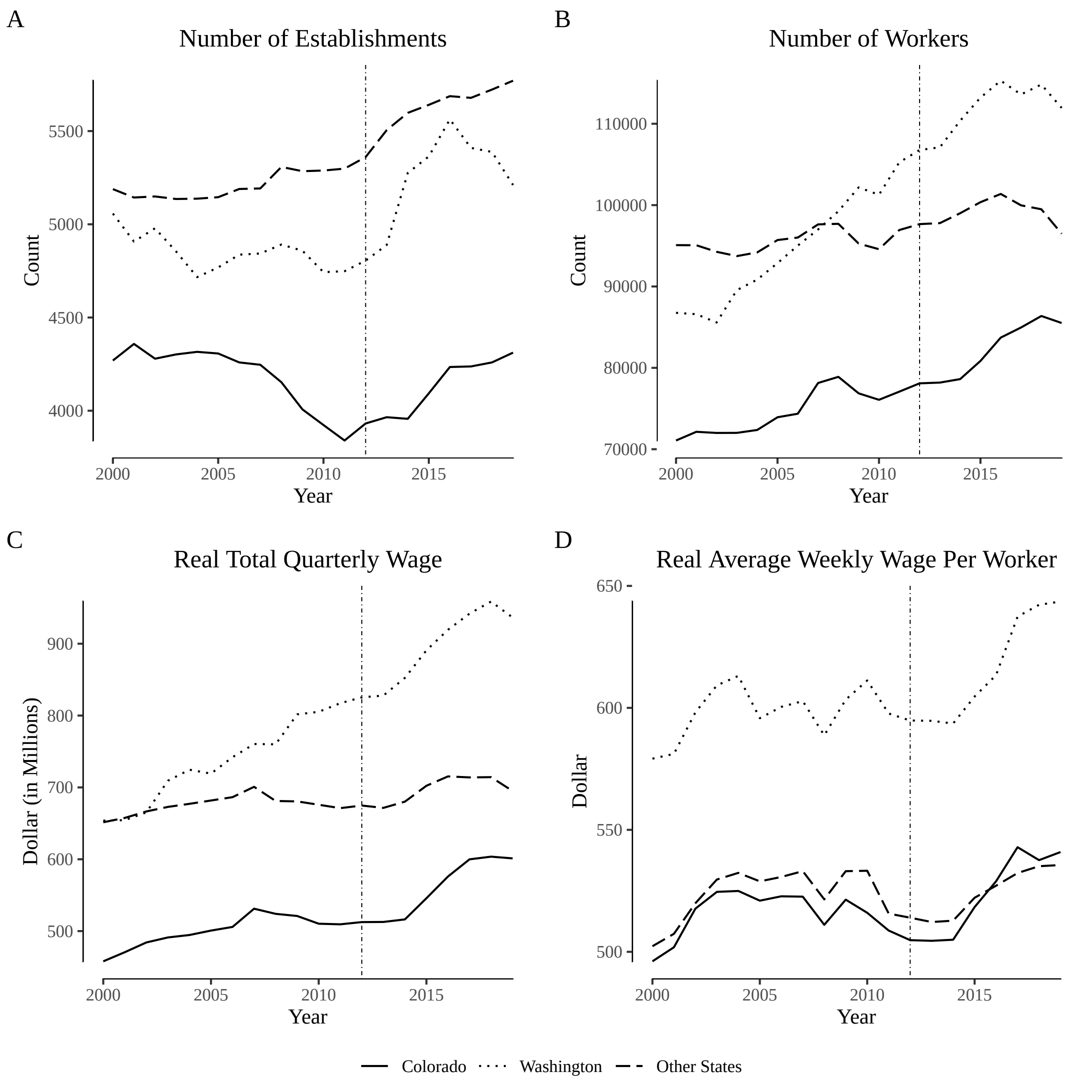
Notes: The table provides the weights assigned to states for the synthetic controls used to estimate the “broadly-defined retail” models in Table 7. All states except those which legalized cannabis during our study period and those bordering either Washington or Colorado were included in the pool of potential control units. Only states which received positive weight for at least one outcome are included in the table.

Figure B.2.1. **Employment and wages for "broadly defined" agricultural firms**



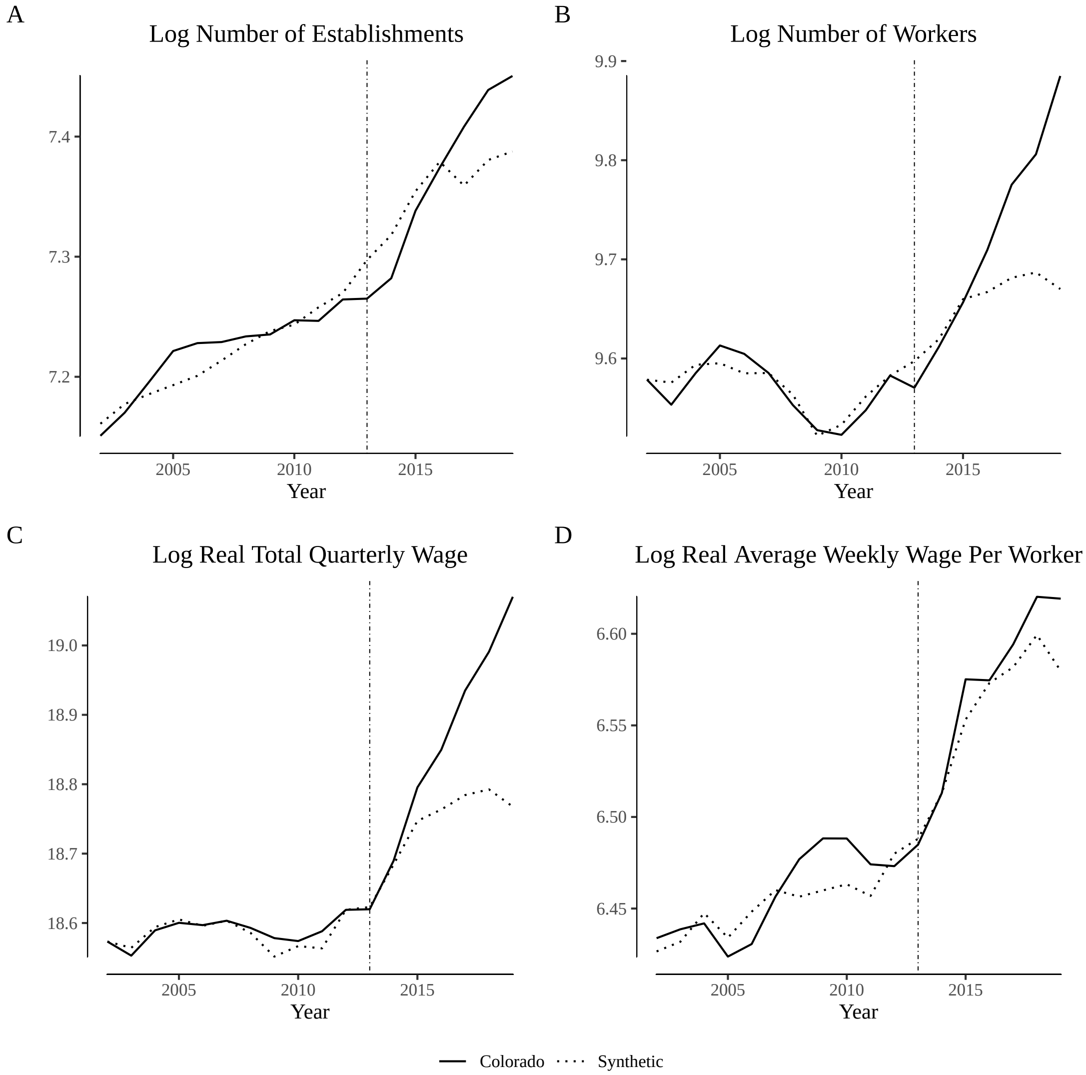
Notes: Data come from the Quarterly Census of Employment and Wages. We define "broadly defined" agricultural firms as those within NAICS 11 ("Agriculture, Forestry, Fishing, and Hunting").

Figure B.2.2. **Employment and wages for "broadly defined" retail firms**



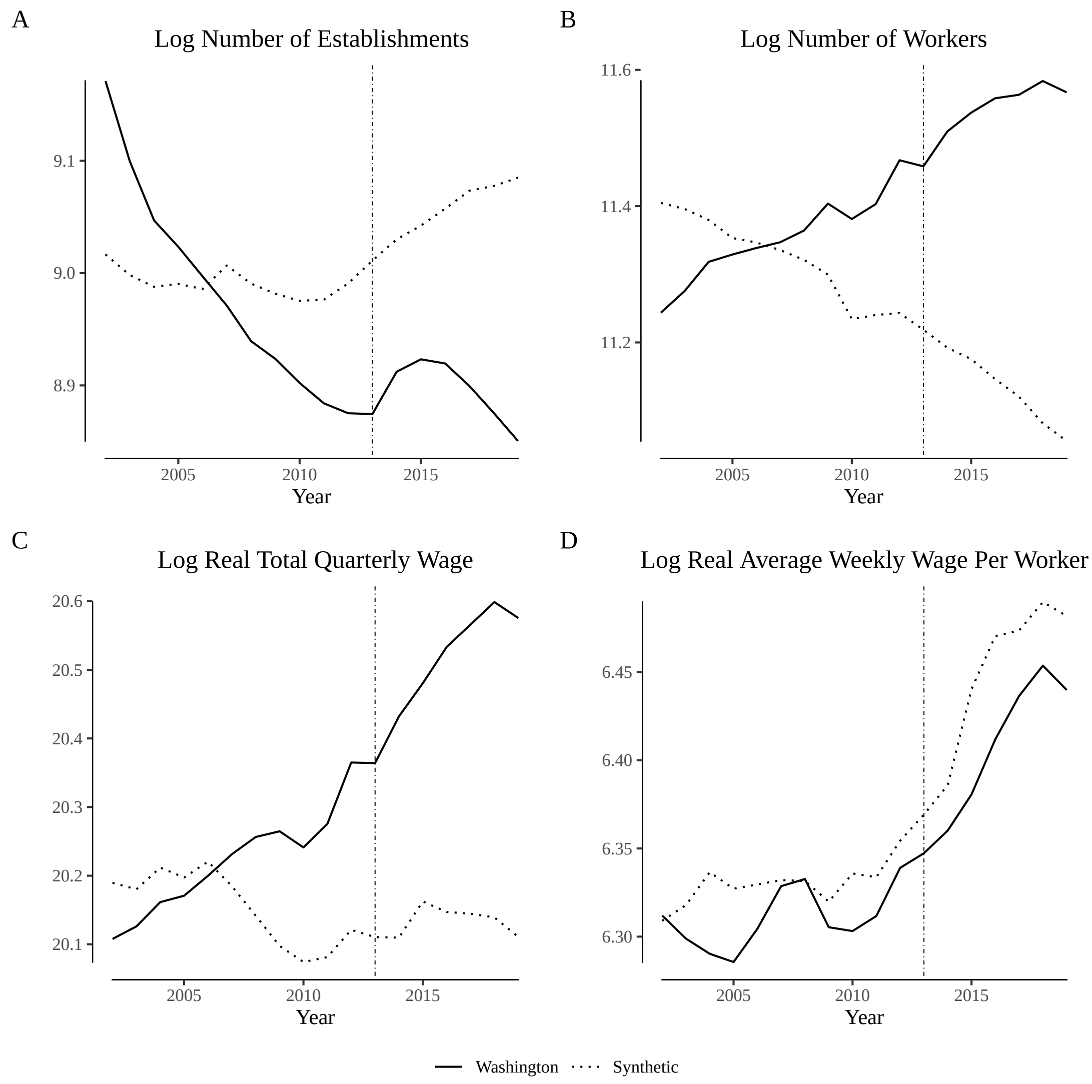
Notes: Data come from the Quarterly Census of Employment and Wages. We define "broadly defined" retail firms as those within NAICS 446, 452, and 453 ("Health and personal care stores", "General merchandise stores", and "Miscellaneous stores", respectively).

Figure B.2.3. **Comparing broadly-defined agriculture labor market outcomes in Colorado and its synthetic control**



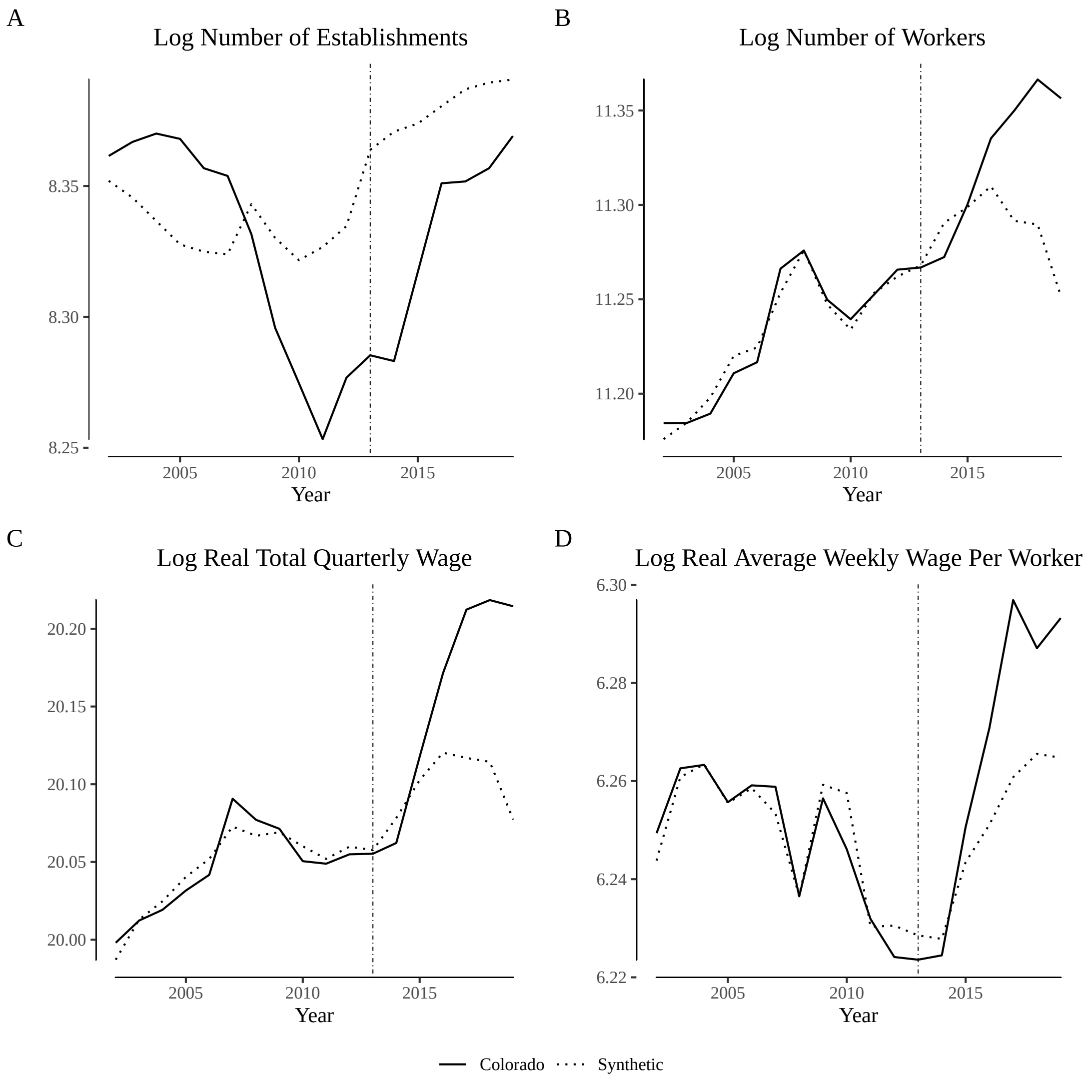
Notes: This figure depicts wage and employment outcomes for "broadly defined" agricultural firms for Colorado and its synthetic control. We define "broadly defined" agricultural firms as those within NAICS 11 ("Agriculture, Forestry, Fishing, and Hunting").

Figure B.2.4. **Comparing broadly-defined agriculture labor market outcomes in Washington and its synthetic control**



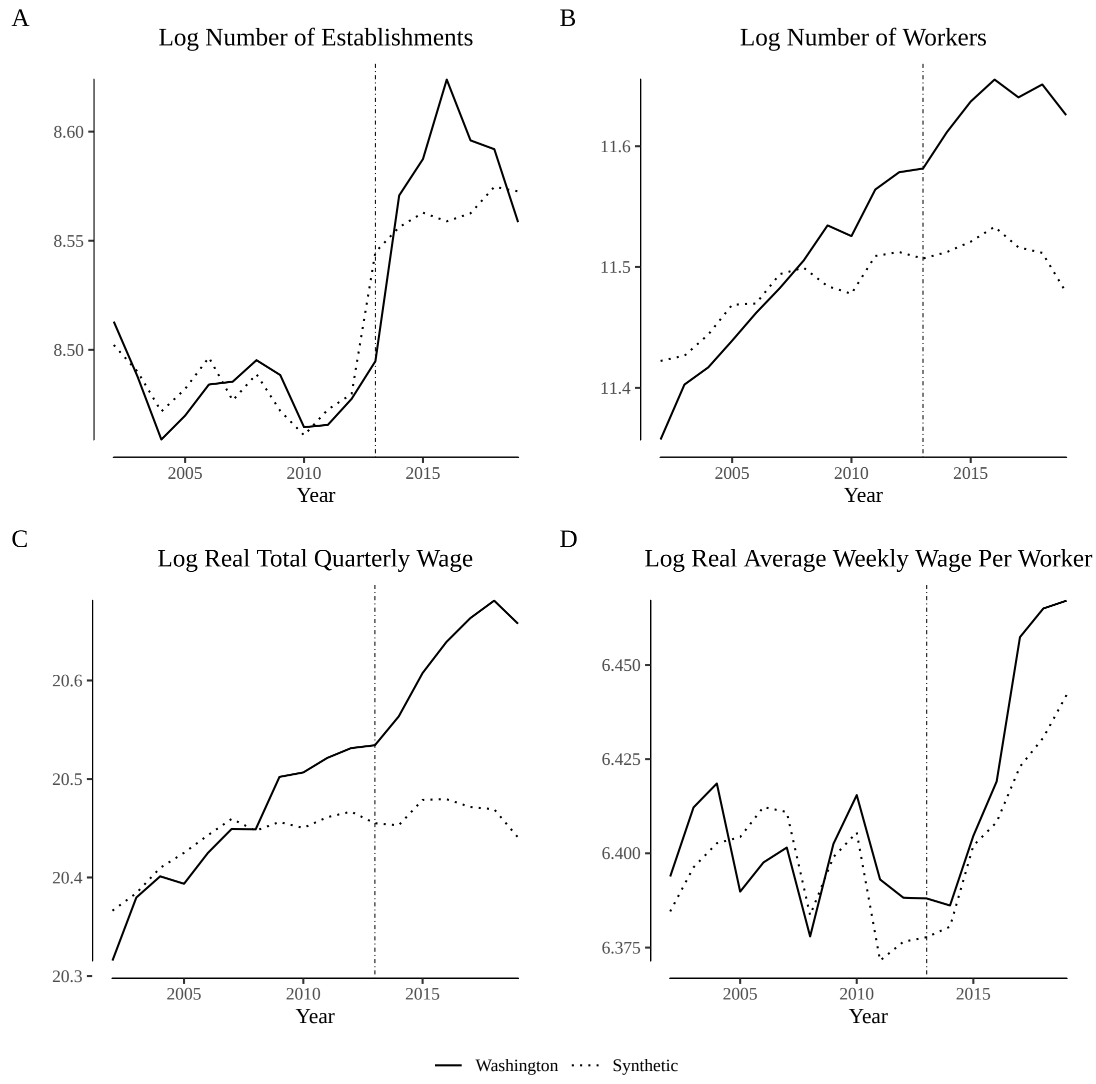
Notes: This figure depicts wage and employment outcomes for "broadly defined" agricultural firms for Washington and its synthetic control. We define "broadly defined" agricultural firms as those within NAICS 11 ("Agriculture, Forestry, Fishing, and Hunting").

Figure B.2.5. **Comparing broadly-defined retailer labor market outcomes in Colorado and its synthetic control**



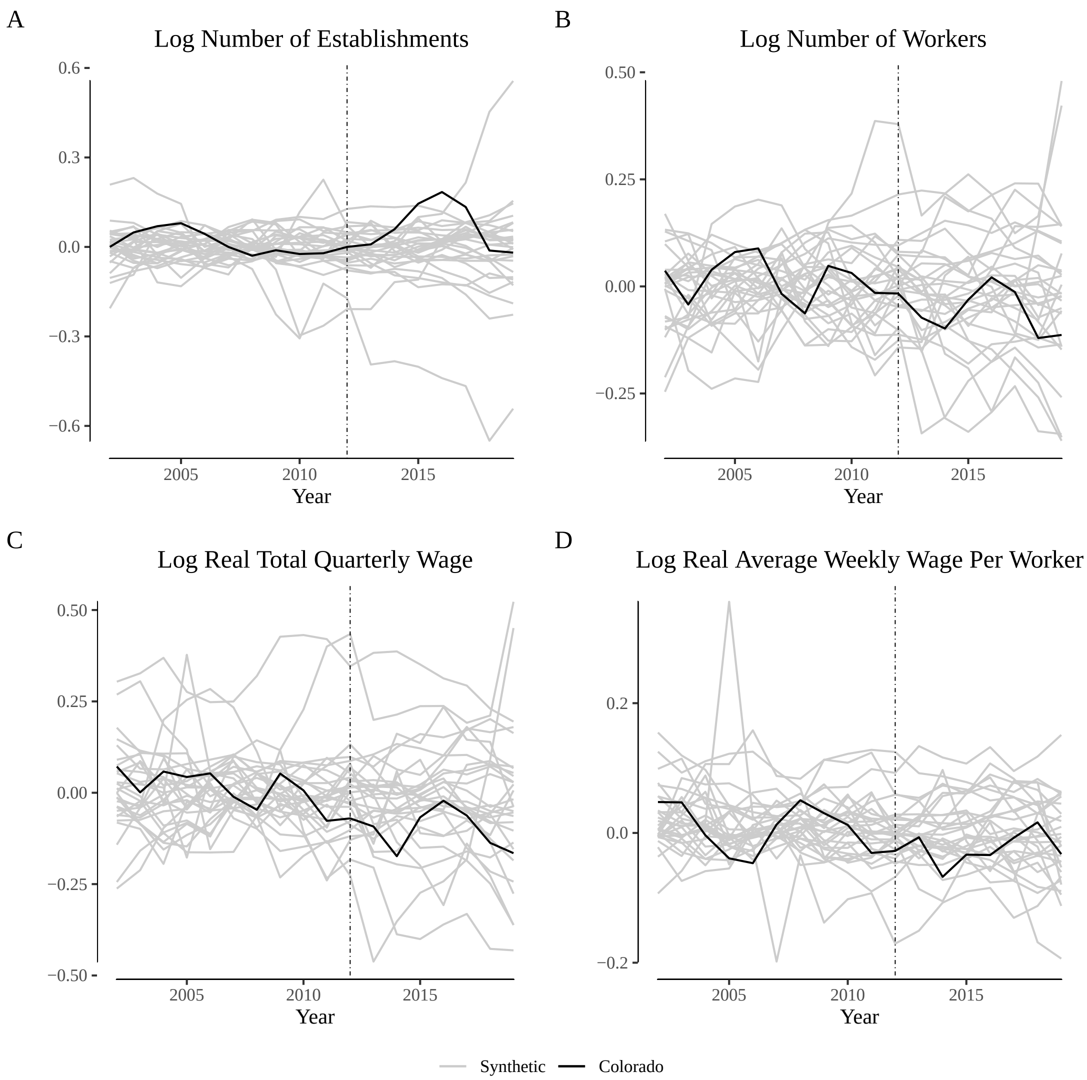
Notes: This figure depicts wage and employment outcomes for "broadly defined" retail firms for Colorado and its synthetic control. We define "broadly defined" retail firms as those within NAICS 446, 452, and 453 ("Health and personal care stores", "General merchandise stores", and "Miscellaneous stores", respectively).

Figure B.2.6. **Comparing broadly-defined retailer labor market outcomes in Washington and its synthetic control**



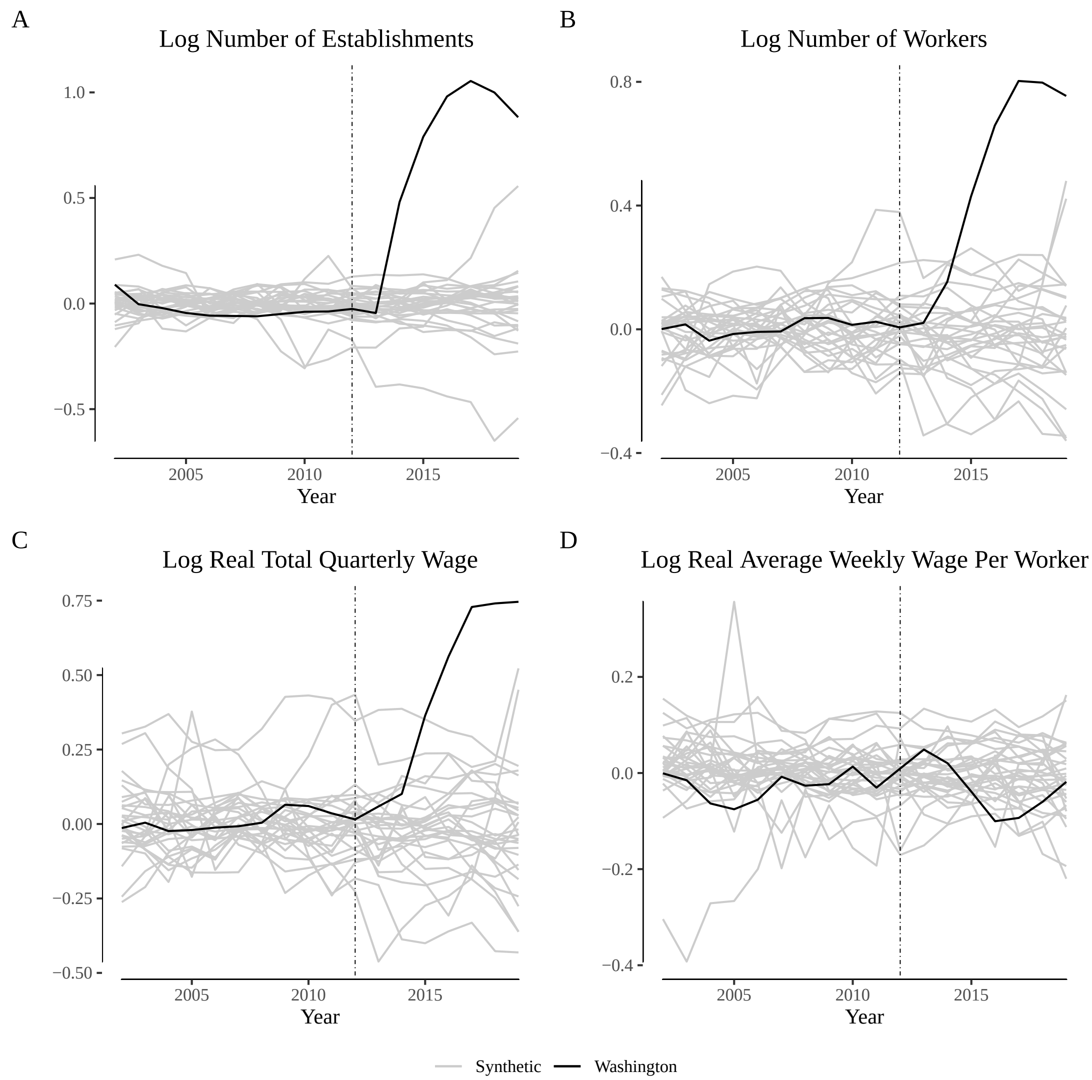
Notes: This figure depicts wage and employment outcomes for "broadly defined" retail firms for Washington and its synthetic control. We define "broadly defined" retail firms as those within NAICS 446, 452, and 453 ("Health and personal care stores", "General merchandise stores", and "Miscellaneous stores", respectively).

Figure B.2.7. **Placebo tests for narrowly-defined agriculture labor market outcomes in Colorado**



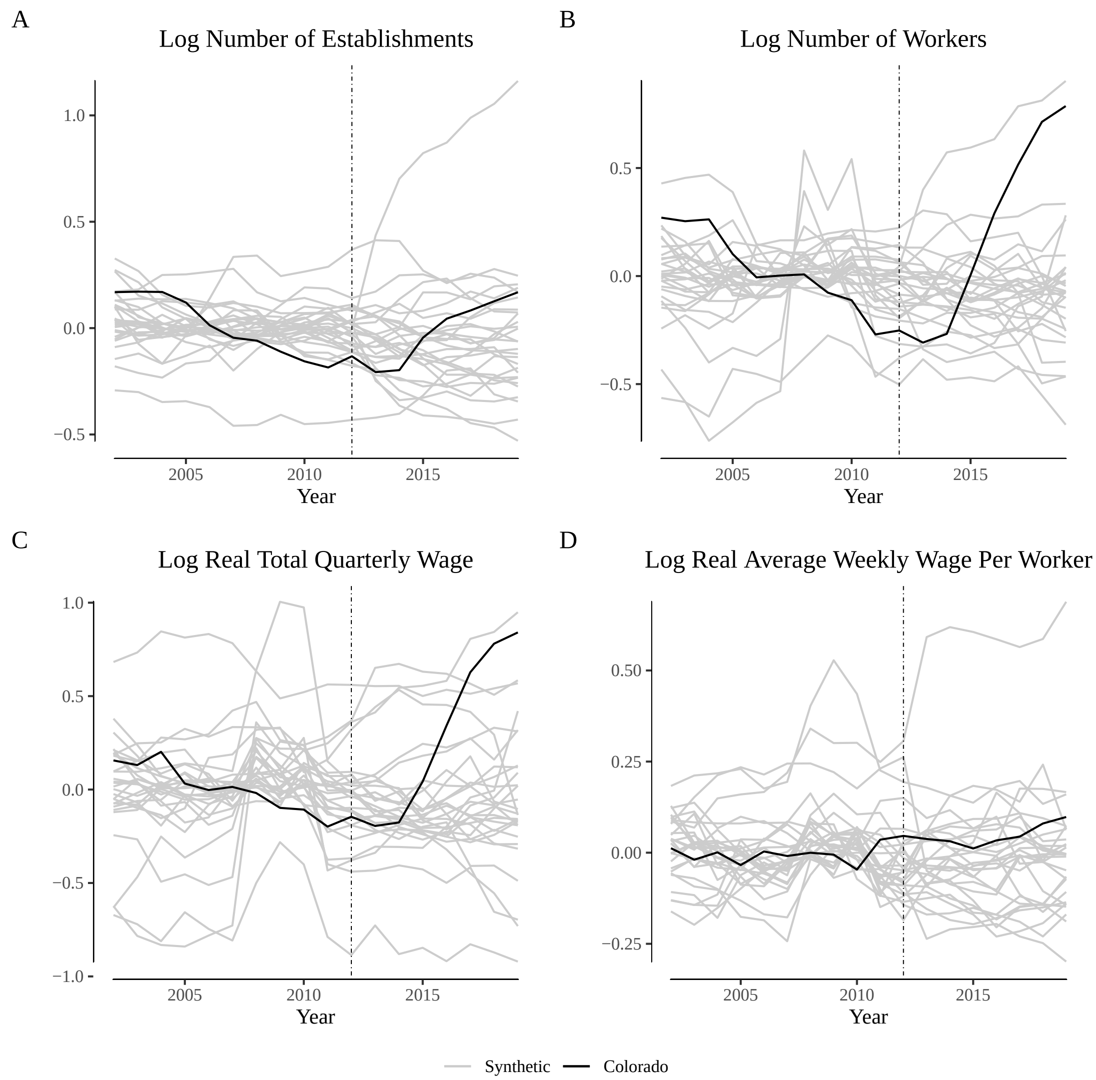
Notes: This figure depicts the placebo tests for "narrowly defined" agricultural firms for Colorado. We define "narrowly defined" agricultural firms as those within the "Greenhouse and Nursery Production" (NAICS 1114) industry.

Figure B.2.8. **Placebo tests for narrowly-defined agriculture labor market outcomes in Washington**



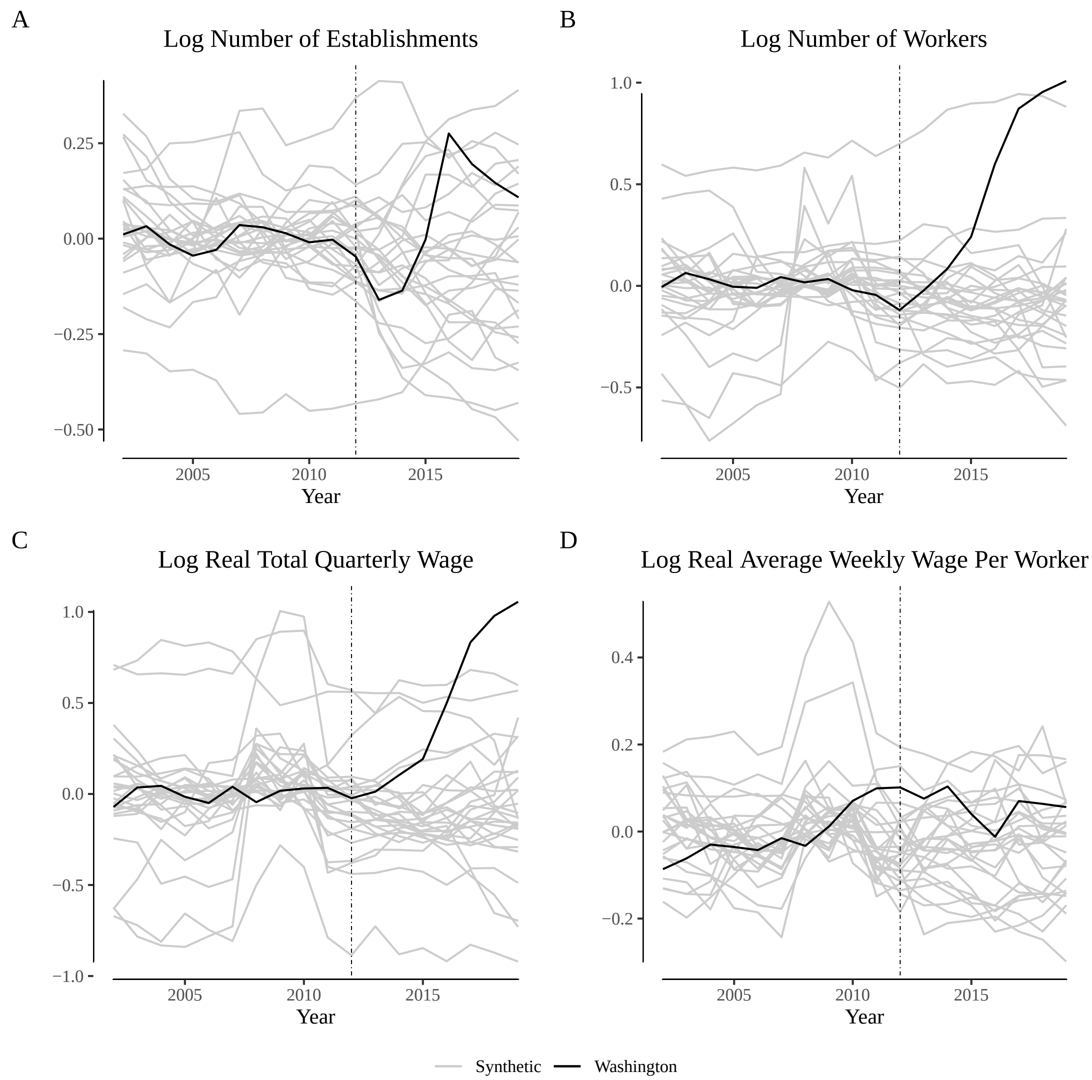
Notes: This figure depicts the placebo tests for "narrowly defined" agricultural firms for Washington. We define "narrowly defined" agricultural firms as those within the "Greenhouse and Nursery Production" (NAICS 1114) industry.

Figure B.2.9. **Placebo tests for narrowly-defined retailer labor market outcomes in Colorado**



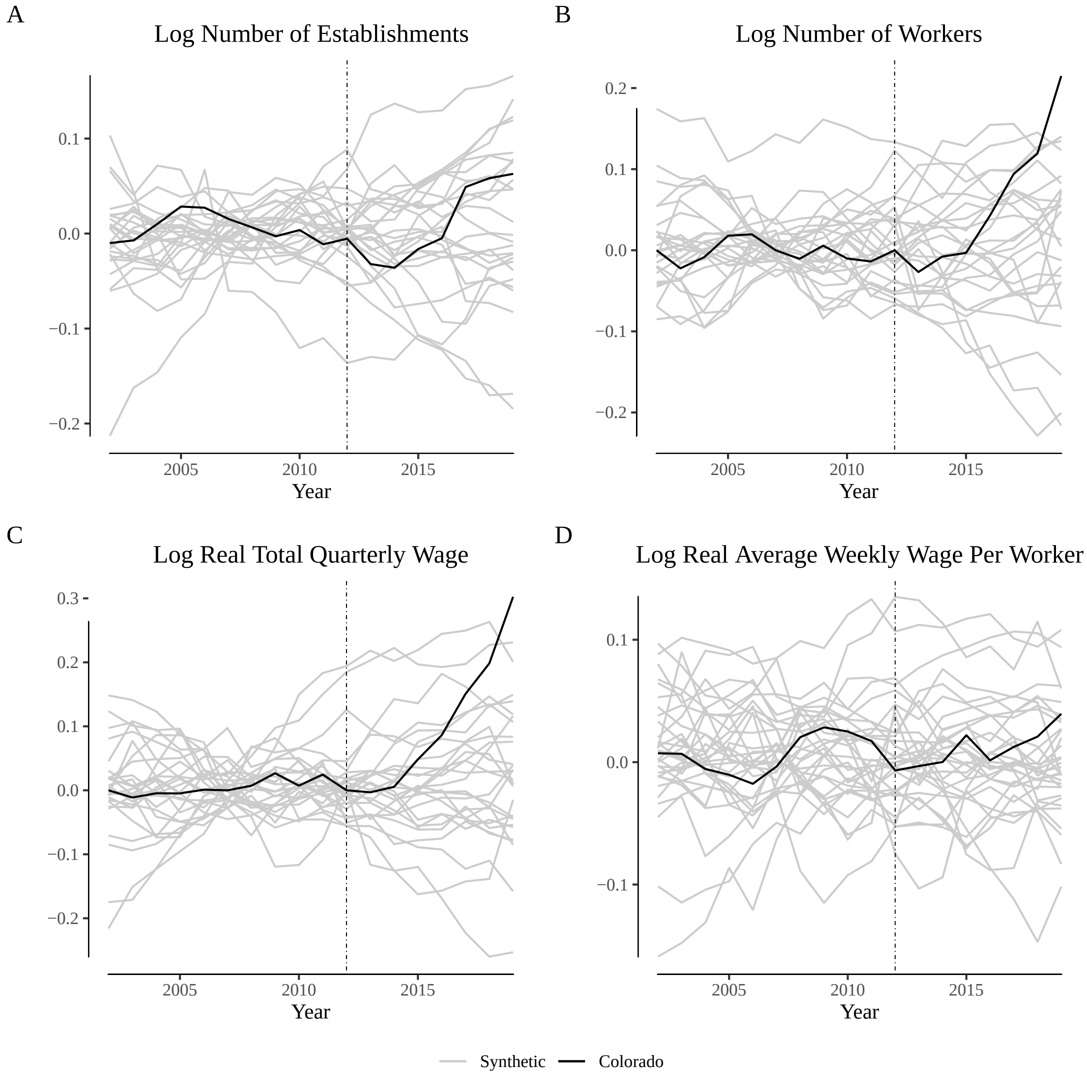
Notes: This figure depicts the placebo tests for wage and employment outcomes for "narrowly defined" retail firms for Colorado. We define "narrowly defined" retail firms as those within the "Store retailers not specified elsewhere" category (NAICS 453998).

Figure B.2.10. **Placebo tests for narrowly-defined retailer labor market outcomes in Washington**



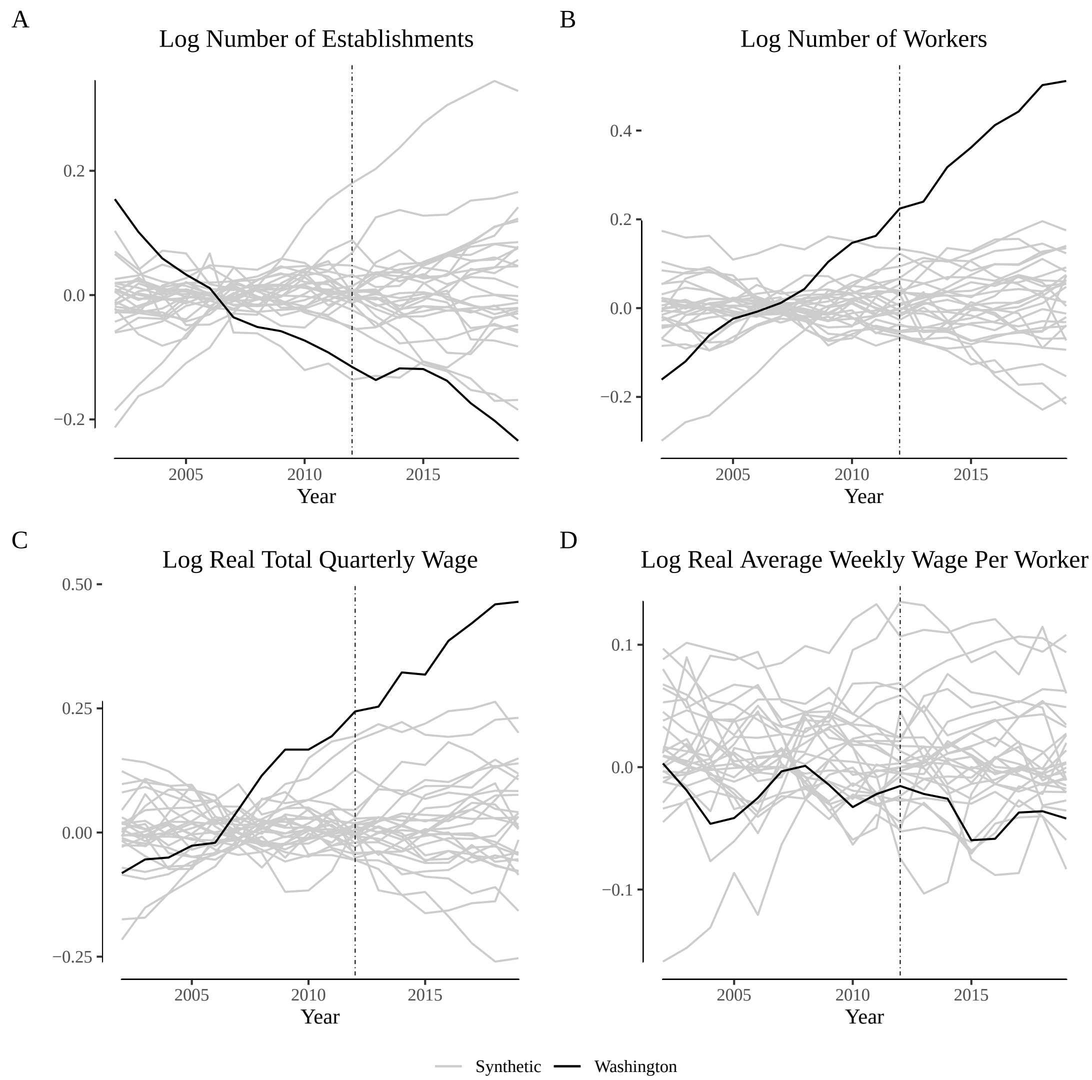
Notes: This figure depicts the placebo tests for wage and employment outcomes for “narrowly defined” retail firms for Washington. We define ”narrowly defined” retail firms as those within the “Store retailers not specified elsewhere” category (NAICS 453998).

Figure B.2.11. **Placebo tests for broadly-defined agriculture labor market outcomes in Colorado**



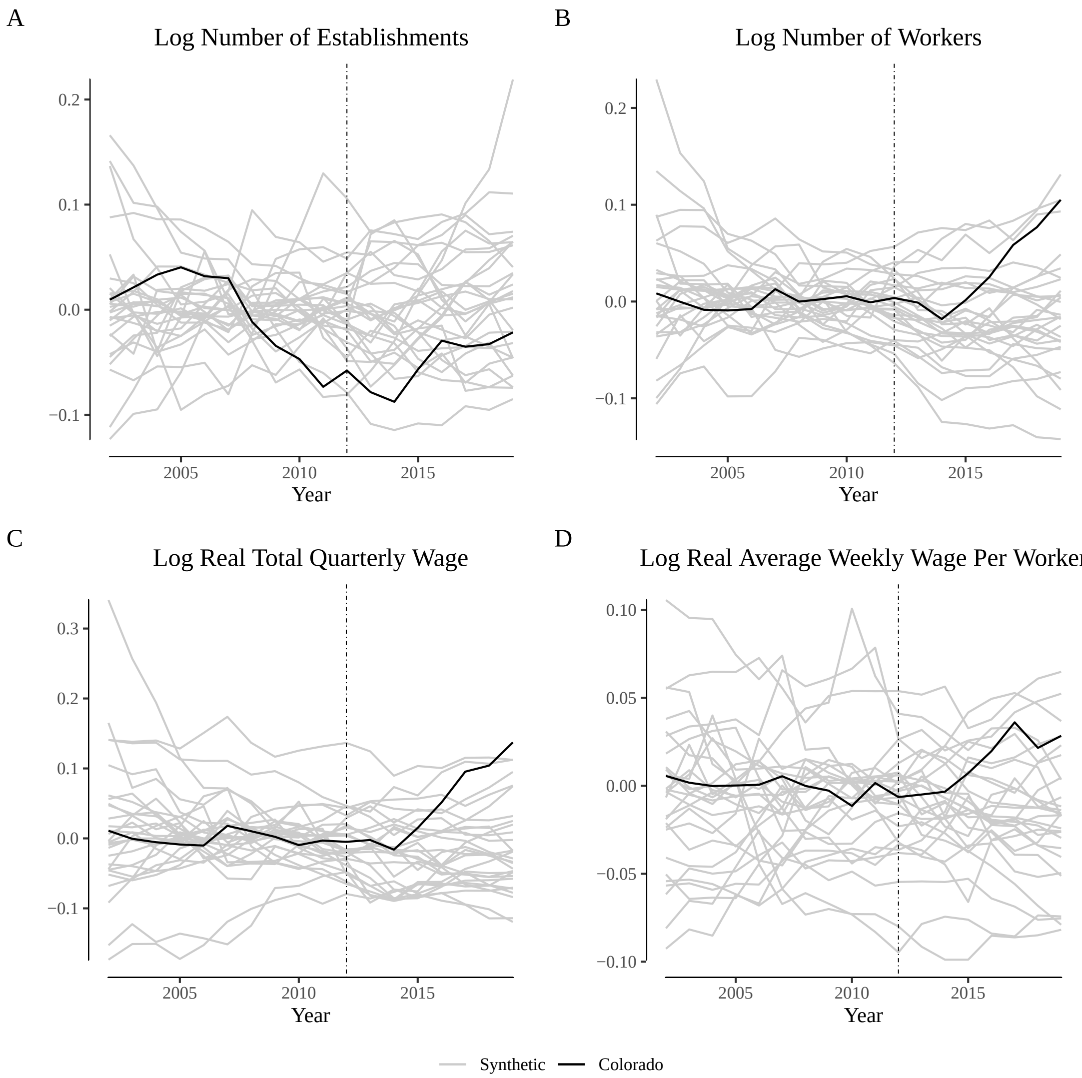
Notes: This figure depicts the placebo tests for "broadly defined" agricultural firms for Colorado. We define "broadly defined" agricultural firms as those within NAICS 11 ("Agriculture, Forestry, Fishing, and Hunting").

Figure B.2.12. **Placebo tests for broadly-defined agriculture labor market outcomes in Washington**



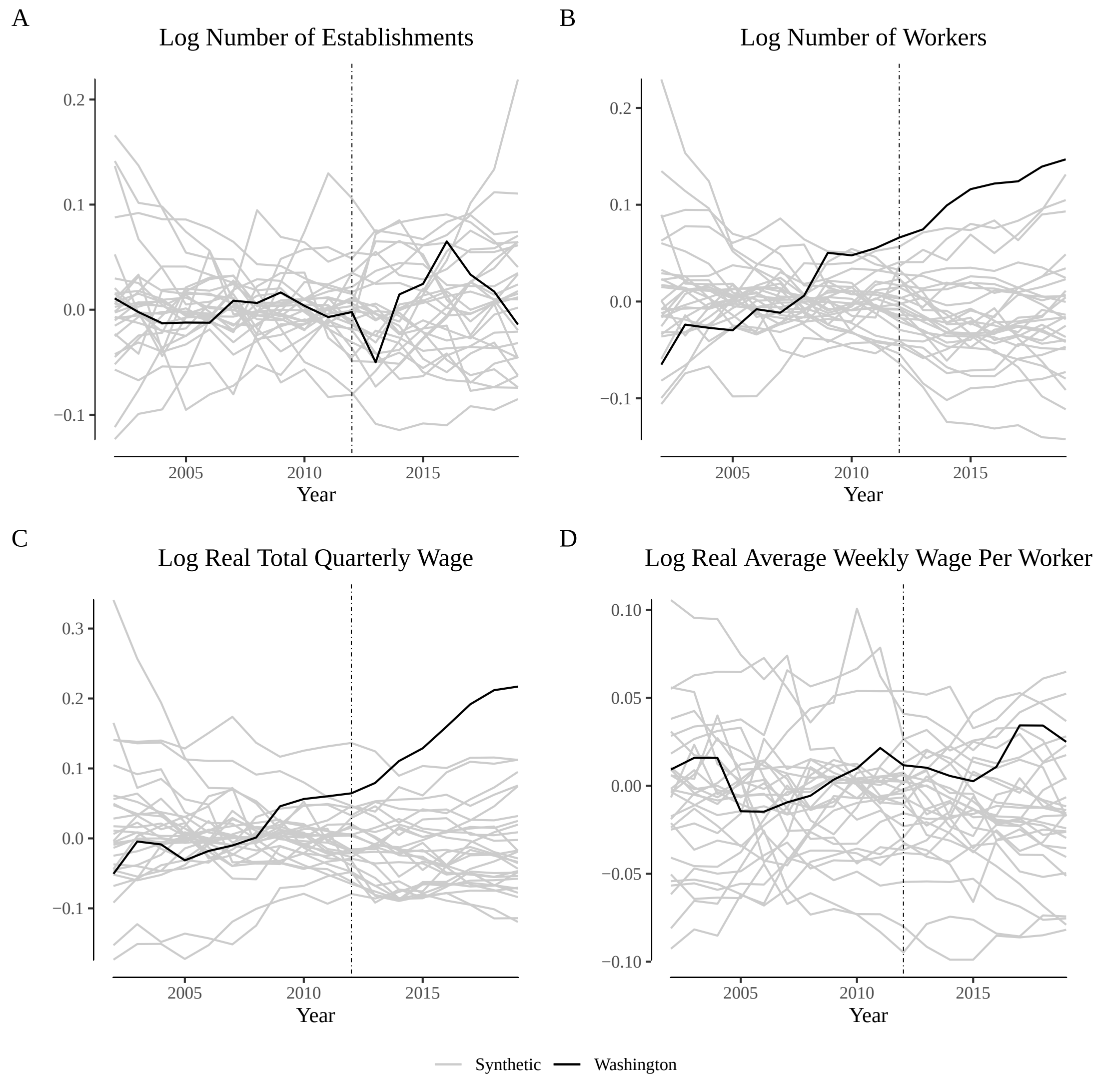
Notes: This figure depicts the placebo tests for "broadly defined" agricultural firms for Washington. We define "broadly defined" agricultural firms as those within NAICS 11 ("Agriculture, Forestry, Fishing, and Hunting").

Figure B.2.13. **Placebo tests for broadly-defined retailer labor market outcomes in Colorado**



Notes: This figure depicts the placebo tests for wage and employment outcomes for "broadly defined" retail firms for Colorado. We define "broadly defined" retail firms as those within NAICS 446, 452, and 453 ("Health and personal care stores", "General merchandise stores", and "Miscellaneous stores", respectively).

Figure B.2.14. **Placebo tests for broadly-defined retailer labor market outcomes in Washington**



Notes: This figure depicts the placebo tests for wage and employment outcomes for “broadly defined” retail firms for Washington. We define ”broadly defined” retail firms as those within NAICS 446, 452, and 453 (“Health and personal care stores”, “General merchandise stores”, and “Miscellaneous stores”, respectively).

B.3 Tables of covariate balance

Table B.3.1. CO broadly-defined agriculture average weekly wage per worker

	Treated	Synthetic	Sample Mean
Lagged outcome	6.46	6.45	6.45
Barley for grain (acres)	65547.33	64022.64	41592.84
Land in orchards (acres)	6444.00	18069.70	51262.40
Snap beans harvested for sale, harvested (acres)	590.67	4442.38	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	159.67	28.15	2208.25
Fruits & nuts, pears, all, total acres (acres)	313.67	128.15	227.10
Comm. soil conds. (thousands of treated acres)	4130.86	8059.39	5357.77
Resident population 65 years & over (percent)	10.29	11.95	13.03
Savings institutions - total deposits (thousands)	1210.38	1377.93	2692.87
Civilian labor force unemployment rate (percent)	5.36	4.93	5.53
Federal Government expenditure-grants (millions)	6.04	8.5	10.89
Federal Government insurance (millions)	3.89	12.04	25.58
Resident population: Black alone (percent)	4.25	11.00	13.53
Resident population: Two or more races (percent)	1.79	1.42	1.71
Resident population: Hispanic or Latino Origin (percent)	17.58	5.92	7.41
Resident population: total females (percent)	49.70	50.64	50.93
Social security: retired workers-benefit recipients (thousands)	386.55	612.67	716.18
Corn Grain Production (dollar, millions)	492.99	1678.67	1045.22
Hay production (dollar, millions)	493.11	412.94	260.35
Farm operations (acres, millions)	62.65	39.43	34.27
Labor hired wage (per hour)	8.50	8.64	11.20
Rent cash cropland expense (acres)	60.00	77.05	75.20
Vegetable totals (dollars, millions)	110.31	56.15	140.21
Wheat production (dollars, millions)	352.67	233.34	138.08

Table B.3.2. CO broadly-defined agriculture total quarterly wages

	Treated	Synthetic	Sample Mean
Lagged outcome	18.59	18.58	18.37
Barley for grain (acres)	65547.33	87647.33	41592.84
Land in orchards (acres)	6444.00	86064.13	51262.40
Snap beans harvested for sale, harvested (acres)	590.67	4755.89	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	159.67	105.70	2208.25
Fruits & nuts, pears, all, total acres (acres)	313.67	247.04	227.10
Comm. soil conds. (thousands of treated acres)	4130.86	8048.08	5357.77
Resident population 65 years & over (percent)	10.29	11.79	13.03
Savings institutions - total deposits (thousands)	1210.38	2651.68	2692.87
Civilian labor force unemployment rate (percent)	5.36	5.10	5.53
Federal Government expenditure-grants (millions)	6.04	13.57	10.89
Federal Government insurance (millions)	3.89	36.76	25.58
Resident population: Black alone (percent)	4.25	9.76	13.53
Resident population: Two or more races (percent)	1.79	1.44	1.71
Resident population: Hispanic or Latino Origin (percent)	17.58	15.23	7.41
Resident population: total females (percent)	49.70	50.44	50.93
Social security: retired workers-benefit recipients (thousands)	386.55	825.32	716.18
Corn Grain Production (dollar, millions)	492.99	728.86	1045.22
Hay production (dollar, millions)	493.11	439.95	260.35
Farm operations (acres, millions)	62.65	99.11	34.27
Labor hired wage (per hour)	8.50	10.15	11.20
Rent cash cropland expense (acres)	60.00	65.61	75.20
Vegetable totals (dollars, millions)	110.31	194.72	140.21
Wheat production (dollars, millions)	352.67	216.10	138.08

Table B.3.3. CO broadly-defined agriculture average employment

	Treated	Synthetic	Sample Mean
Lagged outcome	9.57	9.57	9.36
Barley for grain (acres)	65547.33	258337.20	41592.84
Land in orchards (acres)	6444.00	47636.03	51262.40
Snap beans harvested for sale, harvested (acres)	590.67	2589.34	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	159.67	89.21	2208.25
Fruits & nuts, pears, all, total acres (acres)	313.67	118.67	227.10
Comm. soil conds. (thousands of treated acres)	4130.86	8445.85	5357.77
Resident population 65 years & over (percent)	10.29	12.61	13.03
Savings institutions - total deposits (thousands)	1210.38	1339.26	2692.87
Civilian labor force unemployment rate (percent)	5.36	5.15	5.53
Federal Government expenditure-grants (millions)	6.04	9.72	10.89
Federal Government insurance (millions)	3.89	19.85	25.58
Resident population: Black alone (percent)	4.25	5.00	13.53
Resident population: Two or more races (percent)	1.79	1.70	1.71
Resident population: Hispanic or Latino Origin (percent)	17.58	14.14	7.41
Resident population: total females (percent)	49.70	50.23	50.93
Social security: retired workers-benefit recipients (thousands)	386.55	600.74	716.18
Corn Grain Production (dollar, millions)	492.99	1036.97	1045.22
Hay production (dollar, millions)	493.11	451.59	260.35
Farm operations (acres, millions)	62.65	92.69	34.27
Labor hired wage (per hour)	8.50	12.82	11.20
Rent cash cropland expense (acres)	60.00	71.83	75.20
Vegetable totals (dollars, millions)	110.31	246.95	140.21
Wheat production (dollars, millions)	352.67	396.12	138.08

Table B.3.4. CO broadly-defined agriculture number of establishments

	Treated	Synthetic	Sample Mean
Lagged outcome	7.22	7.21	7.15
Barley for grain (acres)	65547.33	72538.36	41592.84
Land in orchards (acres)	6444.00	81145.72	51262.40
Snap beans harvested for sale, harvested (acres)	590.67	7659.43	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	159.67	13.79	2208.25
Fruits & nuts, pears, all, total acres (acres)	313.67	161.75	227.10
Comm. soil conds. (thousands of treated acres)	4130.86	4692.31	5357.77
Resident population 65 years & over (percent)	10.29	11.78	13.03
Savings institutions - total deposits (thousands)	1210.38	576.47	2692.87
Civilian labor force unemployment rate (percent)	5.36	5.16	5.53
Federal Government expenditure-grants (millions)	6.04	8.54	10.89
Federal Government insurance (millions)	3.89	10.31	25.58
Resident population: Black alone (percent)	4.25	16.11	13.53
Resident population: Two or more races (percent)	1.79	1.47	1.71
Resident population: Hispanic or Latino Origin (percent)	17.58	9.66	7.41
Resident population: total females (percent)	49.70	50.52	50.93
Social security: retired workers-benefit recipients (thousands)	386.55	533.48	716.18
Corn Grain Production (dollar, millions)	492.99	427.44	1045.22
Hay production (dollar, millions)	493.11	253.37	260.35
Farm operations (acres, millions)	62.65	42.90	34.27
Labor hired wage (per hour)	8.50	11.43	11.20
Rent cash cropland expense (acres)	60.00	70.24	75.20
Vegetable totals (dollars, millions)	110.31	282.20	140.21
Wheat production (dollars, millions)	352.67	214.24	138.08

Table B.3.5. WA broadly-defined agriculture average weekly wage per worker

	Treated	Synthetic	Sample Mean
Lagged outcome	6.31	6.33	6.45
Barley for grain (acres)	245385.00	17117.25	41592.84
Land in orchards (acres)	308608.00	209756.76	51262.40
Snap beans harvested for sale, harvested (acres)	3418.67	19386.12	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	1976.33	30113.96	2208.25
Fruits & nuts, pears, all, total acres (acres)	26240.67	751.24	227.10
Comm. soil conds. (thousands of treated acres)	3959.26	5087.90	5357.77
Resident population 65 years & over (percent)	11.55	13.50	13.03
Savings institutions - total deposits (thousands)	3693.15	4339.17	2692.87
Civilian labor force unemployment rate (percent)	6.50	6.98	5.53
Federal Government expenditure-grants (millions)	9.92	14.84	10.89
Federal Government insurance (millions)	7.33	69.71	25.58
Resident population: Black alone (percent)	4.45	13.37	13.53
Resident population: Two or more races (percent)	2.78	1.40	1.71
Resident population: Hispanic or Latino Origin (percent)	9.38	7.01	7.41
Resident population: total females (percent)	50.23	50.91	50.93
Social security: retired workers-benefit recipients (thousands)	622.15	1242.64	716.18
Corn Grain Production (dollar, millions)	78.87	911.96	1045.22
Hay production (dollar, millions)	445.95	247.26	260.35
Farm operations (acres, millions)	30.02	20.39	34.27
Labor hired wage (per hour)	9.50	14.85	11.20
Rent cash cropland expense (acres)	136.50	81.52	75.20
Vegetable totals (dollars, millions)	182.97	362.61	140.21
Wheat production (dollars, millions)	782.89	168.52	138.08

Table B.3.6. WA broadly-defined agriculture total quarterly wages

	Treated	Synthetic	Sample Mean
Lagged outcome	20.20	20.16	18.37
Barley for grain (acres)	245385.00	8933.69	41592.84
Land in orchards (acres)	308608.00	586488.64	51262.40
Snap beans harvested for sale, harvested (acres)	3418.67	28444.94	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	1976.33	20969.07	2208.25
Fruits & nuts, pears, all, total acres (acres)	26240.67	236.82	227.10
Comm. soil conds. (thousands of treated acres)	3959.26	6550.59	5357.77
Resident population 65 years & over (percent)	11.55	15.42	13.03
Savings institutions - total deposits (thousands)	3693.15	6468.79	2692.87
Civilian labor force unemployment rate (percent)	6.50	5.76	5.53
Federal Government expenditure-grants (millions)	9.92	23.07	10.89
Federal Government insurance (millions)	7.33	296.41	25.58
Resident population: Black alone (percent)	4.45	14.89	13.53
Resident population: Two or more races (percent)	2.78	1.28	1.71
Resident population: Hispanic or Latino Origin (percent)	9.38	21.37	7.41
Resident population: total females (percent)	50.23	50.86	50.93
Social security: retired workers-benefit recipients (thousands)	622.15	2187.51	716.18
Corn Grain Production (dollar, millions)	78.87	214.88	1045.22
Hay production (dollar, millions)	445.95	301.58	260.35
Farm operations (acres, millions)	30.02	78.23	34.27
Labor hired wage (per hour)	9.50	32.52	11.20
Rent cash cropland expense (acres)	136.50	77.57	75.20
Vegetable totals (dollars, millions)	182.97	1073.11	140.21
Wheat production (dollars, millions)	782.89	65.42	138.08

Table B.3.7. WA broadly-defined agriculture average employment

	Treated	Synthetic	Sample Mean
Lagged outcome	11.34	11.33	9.36
Barley for grain (acres)	245385.00	10088.92	41592.84
Land in orchards (acres)	308608.00	642996.75	51262.40
Snap beans harvested for sale, harvested (acres)	3418.67	31630.43	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	1976.33	24223.85	2208.25
Fruits & nuts, pears, all, total acres (acres)	26240.67	171.85	227.10
Comm. soil conds. (thousands of treated acres)	3959.26	4600.35	5357.77
Resident population 65 years & over (percent)	11.55	16.18	13.03
Savings institutions - total deposits (thousands)	3693.15	6420.39	2692.87
Civilian labor force unemployment rate (percent)	6.50	5.75	5.53
Federal Government expenditure-grants (millions)	9.92	21.60	10.89
Federal Government insurance (millions)	7.33	325.02	25.58
Resident population: Black alone (percent)	4.45	15.18	13.53
Resident population: Two or more races (percent)	2.78	1.29	1.71
Resident population: Hispanic or Latino Origin (percent)	9.38	19.48	7.41
Resident population: total females (percent)	50.23	50.94	50.93
Social security: retired workers-benefit recipients (thousands)	622.15	2241.95	716.18
Corn Grain Production (dollar, millions)	78.87	134.56	1045.22
Hay production (dollar, millions)	445.95	191.14	260.35
Farm operations (acres, millions)	30.02	47.49	34.27
Labor hired wage (per hour)	9.50	36.43	11.20
Rent cash cropland expense (acres)	136.50	85.01	75.20
Vegetable totals (dollars, millions)	182.97	1198.46	140.21
Wheat production (dollars, millions)	782.89	34.54	138.08

Table B.3.8. WA broadly-defined agriculture number of establishments

	Treated	Synthetic	Sample Mean
Lagged outcome	9.00	8.99	7.15
Barley for grain (acres)	245385.00	62094.10	41592.84
Land in orchards (acres)	308608.00	199298.35	51262.40
Snap beans harvested for sale, harvested (acres)	3418.67	7564.68	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	1976.33	292.81	2208.25
Fruits & nuts, pears, all, total acres (acres)	26240.67	586.85	227.10
Comm. soil conds. (thousands of treated acres)	3959.26	17524.58	5357.77
Resident population 65 years & over (percent)	11.55	10.91	13.03
Savings institutions - total deposits (thousands)	3693.15	6096.23	2692.87
Civilian labor force unemployment rate (percent)	6.50	5.72	5.53
Federal Government expenditure-grants (millions)	9.92	29.17	10.89
Federal Government insurance (millions)	7.33	100.50	25.58
Resident population: Black alone (percent)	4.45	12.36	13.53
Resident population: Two or more races (percent)	2.78	1.25	1.71
Resident population: Hispanic or Latino Origin (percent)	9.38	29.70	7.41
Resident population: total females (percent)	50.23	50.33	50.93
Social security: retired workers-benefit recipients (thousands)	622.15	1658.19	716.18
Corn Grain Production (dollar, millions)	78.87	718.13	1045.22
Hay production (dollar, millions)	445.95	926.67	260.35
Farm operations (acres, millions)	30.02	250.92	34.27
Labor hired wage (per hour)	9.50	7.59	11.20
Rent cash cropland expense (acres)	136.50	32.10	75.20
Vegetable totals (dollars, millions)	182.97	249.71	140.21
Wheat production (dollars, millions)	782.89	302.52	138.08

Table B.3.9. CO narrowly-defined agriculture average weekly wage per worker

	Treated	Synthetic	Sample Mean
Lagged outcome	6.46	6.45	6.34
Barley for grain (acres)	65547.33	120968.85	41592.84
Land in orchards (acres)	6444.00	19911.85	51262.40
Snap beans harvested for sale, harvested (acres)	590.67	4764.35	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	159.67	50.67	2208.25
Fruits & nuts, pears, all, total acres (acres)	313.67	91.63	227.10
Comm. soil conds. (thousands of treated acres)	4130.86	7547.31	5357.77
Resident population 65 years & over (percent)	10.29	11.98	13.03
Savings institutions - total deposits (thousands)	1210.38	1257.63	2692.87
Civilian labor force unemployment rate (percent)	5.36	4.97	5.53
Federal Government expenditure-grants (millions)	6.04	8.89	10.89
Federal Government insurance (millions)	3.89	9.88	25.58
Resident population: Black alone (percent)	4.25	13.61	13.53
Resident population: Two or more races (percent)	1.79	1.44	1.71
Resident population: Hispanic or Latino Origin (percent)	17.58	7.24	7.41
Resident population: total females (percent)	49.70	50.75	50.93
Social security: retired workers-benefit recipients (thousands)	386.55	542.95	716.18
Corn Grain Production (dollar, millions)	492.99	1575.22	1045.22
Hay production (dollar, millions)	493.11	311.07	260.35
Farm operations (acres, millions)	62.65	41.48	34.27
Labor hired wage (per hour)	8.50	10.21	11.20
Rent cash cropland expense (acres)	60.00	81.00	75.20
Vegetable totals (dollars, millions)	110.31	111.04	140.21
Wheat production (dollars, millions)	352.67	268.18	138.08

Table B.3.10. CO narrowly-defined agriculture total quarterly wages

	Treated	Synthetic	Sample Mean
Lagged outcome	17.01	17.00	16.39
Barley for grain (acres)	65547.33	155321.91	41592.84
Land in orchards (acres)	6444.00	140953.22	51262.40
Snap beans harvested for sale, harvested (acres)	590.67	9375.06	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	159.67	133.44	2208.25
Fruits & nuts, pears, all, total acres (acres)	313.67	370.37	227.10
Comm. soil conds. (thousands of treated acres)	4130.86	9932.14	5357.77
Resident population 65 years & over (percent)	10.29	11.01	13.03
Savings institutions - total deposits (thousands)	1210.38	2910.26	2692.87
Civilian labor force unemployment rate (percent)	5.36	5.52	5.53
Federal Government expenditure-grants (millions)	6.04	17.65	10.89
Federal Government insurance (millions)	3.89	48.94	25.58
Resident population: Black alone (percent)	4.25	17.28	13.53
Resident population: Two or more races (percent)	1.79	1.26	1.71
Resident population: Hispanic or Latino Origin (percent)	17.58	15.83	7.41
Resident population: total females (percent)	49.70	50.51	50.93
Social security: retired workers-benefit recipients (thousands)	386.55	1026.54	716.18
Corn Grain Production (dollar, millions)	492.99	384.11	1045.22
Hay production (dollar, millions)	493.11	499.44	260.35
Farm operations (acres, millions)	62.65	133.07	34.27
Labor hired wage (per hour)	8.50	7.91	11.20
Rent cash cropland expense (acres)	60.00	46.64	75.20
Vegetable totals (dollars, millions)	110.31	233.26	140.21
Wheat production (dollars, millions)	352.67	291.41	138.08

Table B.3.11. CO narrowly-defined agriculture average employment

	Treated	Synthetic	Sample Mean
Lagged outcome	8.00	7.98	7.48
Barley for grain (acres)	65547.33	201800.53	41592.84
Land in orchards (acres)	6444.00	132432.44	51262.40
Snap beans harvested for sale, harvested (acres)	590.67	7592.61	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	159.67	155.41	2208.25
Fruits & nuts, pears, all, total acres (acres)	313.67	358.83	227.10
Comm. soil conds. (thousands of treated acres)	4130.86	10667.28	5357.77
Resident population 65 years & over (percent)	10.29	11.41	13.03
Savings institutions - total deposits (thousands)	1210.38	3147.63	2692.87
Civilian labor force unemployment rate (percent)	5.36	5.39	5.53
Federal Government expenditure-grants (millions)	6.04	17.55	10.89
Federal Government insurance (millions)	3.89	52.10	25.58
Resident population: Black alone (percent)	4.25	14.02	13.53
Resident population: Two or more races (percent)	1.79	1.98	1.71
Resident population: Hispanic or Latino Origin (percent)	17.58	16.85	7.41
Resident population: total females (percent)	49.70	50.40	50.93
Social security: retired workers-benefit recipients (thousands)	386.55	1011.87	716.18
Corn Grain Production (dollar, millions)	492.99	435.23	1045.22
Hay production (dollar, millions)	493.11	553.07	260.35
Farm operations (acres, millions)	62.65	149.25	34.27
Labor hired wage (per hour)	8.50	9.84	11.20
Rent cash cropland expense (acres)	60.00	45.00	75.20
Vegetable totals (dollars, millions)	110.31	219.30	140.21
Wheat production (dollars, millions)	352.67	346.31	138.08

Table B.3.12. CO narrowly-defined agriculture number of establishments

	Treated	Synthetic	Sample Mean
Lagged outcome	5.05	5.03	4.82
Barley for grain (acres)	65547.33	197582.96	41592.84
Land in orchards (acres)	6444.00	63451.92	51262.40
Snap beans harvested for sale, harvested (acres)	590.67	5331.53	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	159.67	300.51	2208.25
Fruits & nuts, pears, all, total acres (acres)	313.67	194.82	227.10
Comm. soil conds. (thousands of treated acres)	4130.86	8550.48	5357.77
Resident population 65 years & over (percent)	10.29	11.99	13.03
Savings institutions - total deposits (thousands)	1210.38	1793.17	2692.87
Civilian labor force unemployment rate (percent)	5.36	4.97	5.53
Federal Government expenditure-grants (millions)	6.04	10.85	10.89
Federal Government insurance (millions)	3.89	24.73	25.58
Resident population: Black alone (percent)	4.25	10.88	13.53
Resident population: Two or more races (percent)	1.79	2.48	1.71
Resident population: Hispanic or Latino Origin (percent)	17.58	10.44	7.41
Resident population: total females (percent)	49.70	50.41	50.93
Social security: retired workers-benefit recipients (thousands)	386.55	686.91	716.18
Corn Grain Production (dollar, millions)	492.99	1068.64	1045.22
Hay production (dollar, millions)	493.11	425.24	260.35
Farm operations (acres, millions)	62.65	85.34	34.27
Labor hired wage (per hour)	8.50	12.01	11.20
Rent cash cropland expense (acres)	60.00	64.08	75.20
Vegetable totals (dollars, millions)	110.31	158.05	140.21
Wheat production (dollars, millions)	352.67	338.87	138.08

Table B.3.13. WA narrowly-defined agriculture average weekly wage per worker

	Treated	Synthetic	Sample Mean
Lagged outcome	6.27	6.30	6.34
Barley for grain (acres)	245385.00	143746.92	41592.84
Land in orchards (acres)	308608.00	167305.66	51262.40
Snap beans harvested for sale, harvested (acres)	3418.67	12799.15	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	1976.33	3479.40	2208.25
Fruits & nuts, pears, all, total acres (acres)	26240.67	183.71	227.10
Comm. soil conds. (thousands of treated acres)	3959.26	4957.35	5357.77
Resident population 65 years & over (percent)	11.55	12.11	13.03
Savings institutions - total deposits (thousands)	3693.15	1313.45	2692.87
Civilian labor force unemployment rate (percent)	6.50	5.21	5.53
Federal Government expenditure-grants (millions)	9.92	10.07	10.89
Federal Government insurance (millions)	7.33	54.63	25.58
Resident population: Black alone (percent)	4.45	18.64	13.53
Resident population: Two or more races (percent)	2.78	1.27	1.71
Resident population: Hispanic or Latino Origin (percent)	9.38	7.05	7.41
Resident population: total females (percent)	50.23	50.67	50.93
Social security: retired workers-benefit recipients (thousands)	622.15	748.42	716.18
Corn Grain Production (dollar, millions)	78.87	305.53	1045.22
Hay production (dollar, millions)	445.95	215.55	260.35
Farm operations (acres, millions)	30.02	46.90	34.27
Labor hired wage (per hour)	9.50	12.33	11.20
Rent cash cropland expense (acres)	136.50	62.55	75.20
Vegetable totals (dollars, millions)	182.97	350.80	140.21
Wheat production (dollars, millions)	782.89	255.36	138.08

Table B.3.14. WA narrowly-defined agriculture total quarterly wages

	Treated	Synthetic	Sample Mean
Lagged outcome	17.36	17.35	16.39
Barley for grain (acres)	245385.00	8869.55	41592.84
Land in orchards (acres)	308608.00	207300.29	51262.40
Snap beans harvested for sale, harvested (acres)	3418.67	8706.87	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	1976.33	3410.88	2208.25
Fruits & nuts, pears, all, total acres (acres)	26240.67	378.24	227.10
Comm. soil conds. (thousands of treated acres)	3959.26	10768.68	5357.77
Resident population 65 years & over (percent)	11.55	12.04	13.03
Savings institutions - total deposits (thousands)	3693.15	4829.62	2692.87
Civilian labor force unemployment rate (percent)	6.50	5.40	5.53
Federal Government expenditure-grants (millions)	9.92	21.16	10.89
Federal Government insurance (millions)	7.33	103.23	25.58
Resident population: Black alone (percent)	4.45	9.97	13.53
Resident population: Two or more races (percent)	2.78	1.19	1.71
Resident population: Hispanic or Latino Origin (percent)	9.38	22.69	7.41
Resident population: total females (percent)	50.23	50.47	50.93
Social security: retired workers-benefit recipients (thousands)	622.15	1354.89	716.18
Corn Grain Production (dollar, millions)	78.87	564.14	1045.22
Hay production (dollar, millions)	445.95	581.62	260.35
Farm operations (acres, millions)	30.02	152.02	34.27
Labor hired wage (per hour)	9.50	12.47	11.20
Rent cash cropland expense (acres)	136.50	59.11	75.20
Vegetable totals (dollars, millions)	182.97	345.46	140.21
Wheat production (dollars, millions)	782.89	161.37	138.08

Table B.3.15. WA narrowly-defined agriculture average employment

	Treated	Synthetic	Sample Mean
Lagged outcome	8.52	8.52	7.48
Barley for grain (acres)	245385.00	75321.52	41592.84
Land in orchards (acres)	308608.00	81773.67	51262.40
Snap beans harvested for sale, harvested (acres)	3418.67	8773.77	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	1976.33	8330.37	2208.25
Fruits & nuts, pears, all, total acres (acres)	26240.67	403.45	227.10
Comm. soil conds. (thousands of treated acres)	3959.26	13251.21	5357.77
Resident population 65 years & over (percent)	11.55	11.94	13.03
Savings institutions - total deposits (thousands)	3693.15	2691.01	2692.87
Civilian labor force unemployment rate (percent)	6.50	5.78	5.53
Federal Government expenditure-grants (millions)	9.92	15.09	10.89
Federal Government insurance (millions)	7.33	30.89	25.58
Resident population: Black alone (percent)	4.45	8.54	13.53
Resident population: Two or more races (percent)	2.78	1.37	1.71
Resident population: Hispanic or Latino Origin (percent)	9.38	10.79	7.41
Resident population: total females (percent)	50.23	50.45	50.93
Social security: retired workers-benefit recipients (thousands)	622.15	964.12	716.18
Corn Grain Production (dollar, millions)	78.87	2217.22	1045.22
Hay production (dollar, millions)	445.95	554.67	260.35
Farm operations (acres, millions)	30.02	98.75	34.27
Labor hired wage (per hour)	9.50	8.89	11.20
Rent cash cropland expense (acres)	136.50	79.54	75.20
Vegetable totals (dollars, millions)	182.97	135.63	140.21
Wheat production (dollars, millions)	782.89	339.72	138.08

Table B.3.16. WA narrowly-defined agriculture number of establishments

	Treated	Synthetic	Sample Mean
Lagged outcome	5.91	5.94	4.82
Barley for grain (acres)	245385.00	12036.46	41592.84
Land in orchards (acres)	308608.00	110787.49	51262.40
Snap beans harvested for sale, harvested (acres)	3418.67	17676.70	8048.64
Fruits & nuts, cherries, tart, total acres (acres)	1976.33	35032.51	2208.25
Fruits & nuts, pears, all, total acres (acres)	26240.67	1001.38	227.10
Comm. soil conds. (thousands of treated acres)	3959.26	5728.86	5357.77
Resident population 65 years & over (percent)	11.55	12.73	13.03
Savings institutions - total deposits (thousands)	3693.15	3733.34	2692.87
Civilian labor force unemployment rate (percent)	6.50	7.50	5.53
Federal Government expenditure-grants (millions)	9.92	14.36	10.89
Federal Government insurance (millions)	7.33	4.51	25.58
Resident population: Black alone (percent)	4.45	13.49	13.53
Resident population: Two or more races (percent)	2.78	1.43	1.71
Resident population: Hispanic or Latino Origin (percent)	9.38	3.78	7.41
Resident population: total females (percent)	50.23	50.86	50.93
Social security: retired workers-benefit recipients (thousands)	622.15	1062.24	716.18
Corn Grain Production (dollar, millions)	78.87	960.76	1045.22
Hay production (dollar, millions)	445.95	305.97	260.35
Farm operations (acres, millions)	30.02	22.56	34.27
Labor hired wage (per hour)	9.50	9.01	11.20
Rent cash cropland expense (acres)	136.50	74.39	75.20
Vegetable totals (dollars, millions)	182.97	161.09	140.21
Wheat production (dollars, millions)	782.89	219.83	138.08

Table B.3.17. CO narrow retail average weekly wage per worker

	Treated	Synthetic	Sample Mean
Lagged outcome	6.65	6.57	6.30
College Graduation Rate (percent)	52.48	52.41	53.63
High School Graduation Rate (percent)	76.25	76.77	75.35
Population Density (people per square mile)	45.97	127.83	204.23
State Unemployment Rate (percent)	5.67	5.72	5.93
GDP per capita (dollars, thousands)	68.59	55.18	59.65
Tobacco Store log average weekly wage per worker	6.12	6.11	6.09

Table B.3.18. CO narrow retail total quarterly wages

	Treated	Synthetic	Sample Mean
log total quarterly wages 998 lag	17.11	17.05	16.16
College Graduation Rate (percent)	52.48	52.45	53.63
High School Graduation Rate (percent)	76.25	76.28	75.35
Population Density (people per square mile)	45.97	99.35	204.23
State Unemployment Rate (percent)	5.67	5.66	5.93
GDP per capita (dollars, thousands)	68.59	62.11	59.65
Tobacco Store log total quarterly wages	14.78	14.76	14.66

Table B.3.19. CO narrow retail average employment

	Treated	Synthetic	Sample Mean
log average employment 998 lag	7.90	7.85	7.31
College Graduation Rate (percent)	52.48	52.44	53.63
High School Graduation Rate (percent)	76.25	76.28	75.35
Population Density (people per square mile)	45.97	87.94	204.23
State Unemployment Rate (percent)	5.67	5.48	5.93
GDP per capita (dollars, thousands)	68.59	63.53	59.65
Tobacco Store log average employment	6.10	6.10	6.02

Table B.3.20. CO narrow retail number of establishments

	Treated	Synthetic	Sample Mean
log number of establishments 998 lag	6.28	6.27	5.73
College Graduation Rate (percent)	52.48	52.54	53.63
High School Graduation Rate (percent)	76.25	76.04	75.35
Population Density (people per square mile)	45.97	84.01	204.23
State Unemployment Rate (percent)	5.67	5.49	5.93
GDP per capita (dollars, thousands)	68.59	61.31	59.65
Tobacco Store log number of establishments	4.77	4.74	4.47

Table B.3.21. WA narrow retail average weekly wage per worker

	Treated	Synthetic	Sample Mean
Lagged outcome	6.32	6.33	6.30
College Graduation Rate (percent)	63.07	59.27	53.63
High School Graduation Rate (percent)	73.58	73.90	75.35
Population Density (people per square mile)	95.95	158.82	204.23
State Unemployment Rate (percent)	6.90	6.43	5.93
GDP per capita (dollars, thousands)	70.29	57.73	59.65
Tobacco Store log average weekly wage per worker	6.03	6.20	6.09

Table B.3.22. WA narrow retail total quarterly wages

	Treated	Synthetic	Sample Mean
log total quarterly wages 998 lag	16.09	16.07	16.16
College Graduation Rate (percent)	63.07	59.18	53.63
High School Graduation Rate (percent)	73.58	73.89	75.35
Population Density (people per square mile)	95.95	362.27	204.23
State Unemployment Rate (percent)	6.90	6.69	5.93
GDP per capita (dollars, thousands)	70.29	69.04	59.65
Tobacco Store log total quarterly wages	14.52	14.71	14.66

Table B.3.23. WA narrow retail average employment

	Treated	Synthetic	Sample Mean
log average employment 998 lag	7.20	7.24	7.31
College Graduation Rate (percent)	63.07	59.31	53.63
High School Graduation Rate (percent)	73.58	75.02	75.35
Population Density (people per square mile)	95.95	342.54	204.23
State Unemployment Rate (percent)	6.90	6.68	5.93
GDP per capita (dollars, thousands)	70.29	68.95	59.65
Tobacco Store log average employment	5.93	6.11	6.02

Table B.3.24. WA narrow retail number of establishments

	Treated	Synthetic	Sample Mean
log number of establishments 998 lag	5.63	6.14	5.73
College Graduation Rate (percent)	63.07	59.52	53.63
High School Graduation Rate (percent)	73.58	78.49	75.35
Population Density (people per square mile)	95.95	144.57	204.23
State Unemployment Rate (percent)	6.90	6.26	5.93
GDP per capita (dollars, thousands)	70.29	57.48	59.65
Tobacco Store log number of establishments	4.97	4.94	4.47

Table B.3.25. CO broad retail average weekly wage per worker

	Treated	Synthetic	Sample Mean
Lagged outcome	6.25	6.25	6.26
College Graduation Rate (percent)	52.48	52.72	53.63
High School Graduation Rate (percent)	76.25	76.39	75.35
Population Density (people per square mile)	45.97	126.41	204.23
State Unemployment Rate (percent)	5.67	5.76	5.93
GDP per capita (dollars, thousands)	68.59	59.97	59.65

Table B.3.26. CO broad retail total quarterly wages

	Treated	Synthetic	Sample Mean
Lagged outcome	20.04	20.04	20.14
College Graduation Rate (percent)	52.48	48.54	53.63
High School Graduation Rate (percent)	76.25	76.28	75.35
Population Density (people per square mile)	45.97	68.04	204.23
State Unemployment Rate (percent)	5.67	5.73	5.93
GDP per capita (dollars, thousands)	68.59	56.62	59.65

Table B.3.27. CO broad retail average employment

	Treated	Synthetic	Sample Mean
Lagged outcome	11.23	11.23	11.31
College Graduation Rate (percent)	52.48	47.95	53.63
High School Graduation Rate (percent)	76.25	76.70	75.35
Population Density (people per square mile)	45.97	55.57	204.23
State Unemployment Rate (percent)	5.67	5.67	5.93
GDP per capita (dollars, thousands)	68.59	56.81	59.65

Table B.3.28. CO broad retail number of establishments

	Treated	Synthetic	Sample Mean
Lagged outcome	8.33	8.33	8.43
College Graduation Rate (percent)	52.48	56.25	53.63
High School Graduation Rate (percent)	76.25	76.25	75.35
Population Density (people per square mile)	45.97	381.19	204.23
State Unemployment Rate (percent)	5.67	5.79	5.93
GDP per capita (dollars, thousands)	68.59	68.52	59.65

Table B.3.29. WA broad retail average weekly wage per worker

	Treated	Synthetic	Sample Mean
Lagged outcome	6.40	6.40	6.26
College Graduation Rate (percent)	63.07	53.43	53.63
High School Graduation Rate (percent)	73.58	70.26	75.35
Population Density (people per square mile)	95.95	264.90	204.23
State Unemployment Rate (percent)	6.90	5.98	5.93
GDP per capita (dollars, thousands)	70.29	70.85	59.65

Table B.3.30. WA broad retail total quarterly wages

	Treated	Synthetic	Sample Mean
Lagged outcome	20.43	20.43	20.14
College Graduation Rate (percent)	63.07	60.78	53.63
High School Graduation Rate (percent)	73.58	73.94	75.35
Population Density (people per square mile)	95.95	180.77	204.23
State Unemployment Rate (percent)	6.90	5.17	5.93
GDP per capita (dollars, thousands)	70.29	63.21	59.65

Table B.3.31. WA broad retail average employment

	Treated	Synthetic	Sample Mean
Lagged outcome	11.47	11.47	11.31
College Graduation Rate (percent)	63.07	59.88	53.63
High School Graduation Rate (percent)	73.58	73.84	75.35
Population Density (people per square mile)	95.95	146.39	204.23
State Unemployment Rate (percent)	6.90	5.87	5.93
GDP per capita (dollars, thousands)	70.29	58.93	59.65

Table B.3.32. WA broad retail number of establishments

	Treated	Synthetic	Sample Mean
Lagged outcome	8.48	8.48	8.43
College Graduation Rate (percent)	63.07	59.59	53.63
High School Graduation Rate (percent)	73.58	73.87	75.35
Population Density (people per square mile)	95.95	396.66	204.23
State Unemployment Rate (percent)	6.90	6.50	5.93
GDP per capita (dollars, thousands)	70.29	70.13	59.65

APPENDIX C

CHAPTER 4 APPENDIX

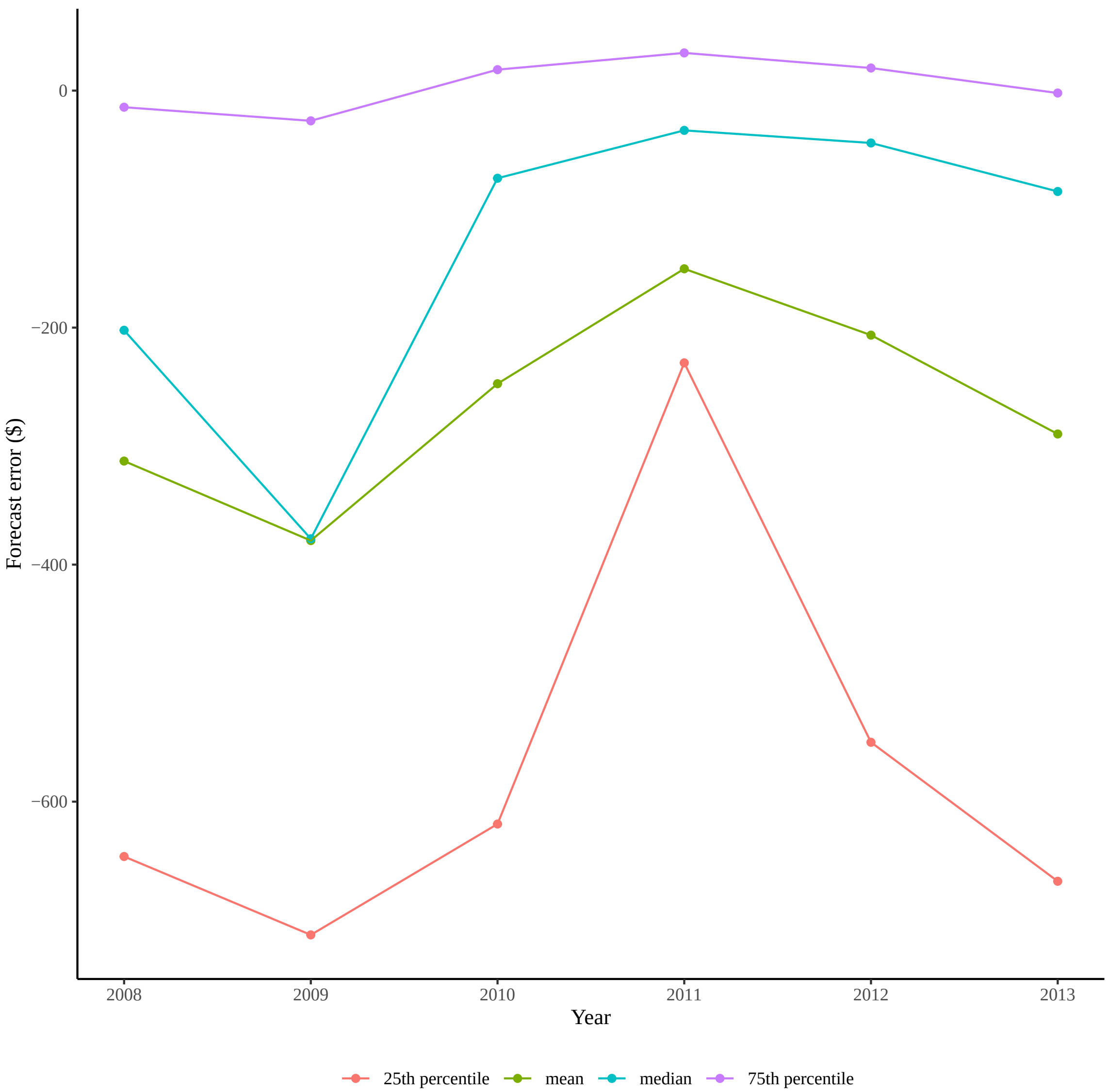
C.1 Tables and Figures

Figure C.1. Example bid cost reporting

	Actual, 1/1/2013-12/31/2013			Projected, 1/1/2015-12/31/2015		
Service Category	**Annualized Util/1000**	**Avg Cost**	**Allowed PMPM**	**Annual Util/1000**	**Avg Cost**	**Total Allowed PMPM**
a. Inpatient Facility	802	$2,358.54	$157.71	813	$2,486.41	$168.47
b. Skilled Nursing Facility	672	427.90	23.97	800	473.02	31.54
c. Home Health	580	192.78	9.32	674	211.30	11.86
d. Ambulance	107	438.99	3.92	125	480.31	4.99
e. DME/Prosthetics/Supplies	9,312	14.40	11.18	10,850	15.78	14.27
f. OP Facility - Emergency	305	747.00	19.00	323	778.66	20.94
g. OP Facility - Surgery	444	2,133.11	78.91	462	2,196.53	84.57
h. OP Facility - Other	5,106	108.59	46.20	5,384	112.78	50.60
i. Professional	14,225	122.25	144.92	15,009	120.85	151.14
j. Part B Rx	1,765	216.40	31.83	1,861	201.07	31.18
k. Other Medicare Part B	0	0.00	0.00	0	0.00	0.00
l. Transportation (Non-Covered)	0	0.00	0.00	0	0.00	0.00
m. Dental (Non-Covered)	0	0.00	0.00	195	88.10	1.43
n. Vision (Non-Covered)	0	0.00	0.00	0	0.00	0.00
o. Hearing (Non-Covered)	0	0.00	0.00	0	0.00	0.00
p. Health & Education (Non-Covered)	19,435	4.92	7.96	6,791	5.53	3.13
q. Other Non-Covered	5,347	1.17	0.52	7,256	1.64	0.99
r. COB/Subrg. (outside claim system)						0.00
s. Total Medical Expenses			**$535.47**			**$575.12**

Notes: This figure is reconstructed from the 2015 bid data for plan H0028-01-0, “Humana Gold Plus”, offered in Cedar Rapids, Iowa, and surrounding counties, using the CMS public bid data and the 2015 Bid Pricing Tool (BPT). The ‘Actual’ columns are excerpted from the “MA Base Period Experience” portion of the BPT. The ‘Projected’ columns are excerpted from the “MA Projected Allowed Costs” portion.

Figure C.2. **The distribution of the overall prediction error over time**



Notes: An observation is a plan-year. Statistics are unweighted. The overall prediction error is calculated per-member-per-month (PMPM). Negative prediction errors indicate firms overpredicted their costs.

Table C.1. Summary statistics

	Min	25th	Mean	Median	75th	Max
Prediction errors:						
Overall	-1515.93	-638.00	-270.20	-89.29	5.89	2228.70
Inpatient Facility	-638.93	-243.72	-103.09	-43.10	1.80	778.10
Professional	-596.44	-167.62	-77.03	-37.73	1.03	2168.77
Outpatient Facility - Surgery	-244.07	-38.12	-15.61	-8.42	3.27	158.45
Firm-chosen product characteristics:						
Deductible	0	0	96.15	0	0	5000
Out-of-pocket limit	0	2000	3437.44	3400	5000	12000
Primary care copay	0	5	11.83	10	15	47.5
Specialist copay	0	20	26.09	30	35	60
Hospital copay	0	400	893.79	875	1400	2800
CMS-determined quality rating variables:						
Star rating	0.0	0.0	2.48	3.0	3.5	5.0
'Too new' indicator	0	0	0.08	0	0	1
'No data' indicator	0	0	0.18	0	0	1
Other covariates:						
Risk-adjusted payment	312.75	683.11	731.49	726.17	775.25	1134.19
Number of competitors	1.06	7.41	11.72	10.21	15.02	38
Number of competing plans	0	12.38	23.42	20.10	29.90	89
Log previous enrollment	2.48	8.34	9.47	9.69	10.73	13.30
Observations			11,670			

Notes: An observation is a plan-year. Statistics are unweighted. Prediction errors and risk-adjusted payments are per-member-per-month (PMPM). Negative prediction errors indicate firms overpredicted their costs. Hospital copays are for seven day stays. The number of competitors is calculated at the contract level, weighted by the number of contract enrollees in each county covered by the contract. The number of competing plans is calculated at the plan level, weighted by the number of plan enrollees in each county covered by the contract.

Table C.2. Correlations between the overall forecast error and its major components

	Overall	Inpatient Facility	Professional	Outpatient Facility Surgery
Overall	1			
Inpatient Facility	0.97	1		
Professional	0.84	0.82	1	
Outpatient Facility - Surgery	0.94	0.86	0.73	1

Notes: This table presents correlations between the overall forecast error and its three largest (on average) components. An observation is a plan-year; there are 11,670 observations total. Statistics are unweighted.

Table C.3. The distribution of the overall prediction error across quartiles of measures of experience and competition

	Experience			
	First quartile	Second quartile	Third quartile	Last quartile
Outcome: Overall prediction error				
Min	-1515.93	-1358.61	-1247.43	-1308.29
25th percentile	-667.15	-632.43	-570.92	-649.47
Mean	-308.36	-257.80	-227.50	-287.14
Median	-191.04	-79.68	-58.62	-86.17
75th percentile	0	12.96	9.73	0.55
Max	1276.81	924.99	792.77	2228.70
	Competition			
	First quartile	Second quartile	Third quartile	Last quartile
Outcome: Overall prediction error				
Min	-1494.49	-1309.37	-1358.61	-1515.93
25th percentile	-590.81	-639.82	-685.87	-614.90
Mean	-230.48	-287.67	-307.58	-255.08
Median	-56.11	-124.33	-131.33	-81.24
75th percentile	13.27	7.53	-2.87	7.77
Max	1276.81	2278.70	792.77	1023.31

Notes: An observation is a plan-year. Statistics are unweighted. The overall prediction error is calculated per-member-per-month (PMPM). Negative prediction errors indicate firms overpredicted their costs. Our measure of experience is the log of the total number of previous enrollees at the contract level. Our measure of competition is the number of competitors calculated at the contract level, weighted by the number of contract enrollees in each county covered by the contract.

Table C.4. The association between prediction errors and plan observables

	Dependent variable: Prediction errors			
	Overall	Inpatient facility	OP Facility - Surgery	Professional
	(1)	(2)	(3)	(4)
Overall prediction error$_{t-1}$	0.002***	0.001***	0.0001***	0.001***
	(0.0001)	(0.00003)	(0.00001)	(0.00002)
Log previous enrollment	2.945	−1.907*	−0.572***	3.122***
	(2.461)	(1.018)	(0.205)	(0.756)
Number of competitors	6.413***	3.199***	0.267**	1.661***
	(1.403)	(0.580)	(0.117)	(0.431)
Number of competing plans	−0.902**	−0.430***	−0.059*	−0.229**
	(0.367)	(0.152)	(0.031)	(0.113)
Plan age	46.284***	17.676***	4.218***	11.924***
	(9.357)	(3.869)	(0.779)	(2.872)
Constant	−811.990***	−295.974***	−27.565	−292.505***
	(206.648)	(85.458)	(17.196)	(63.439)
Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	11,530	11,530	11,530	11,530
R^2	0.257	0.238	0.239	0.265
Adjusted R^2	0.232	0.213	0.214	0.241

Notes: An observation is a plan-year. Estimates are obtained using unweighted OLS. Stars indicate p-values: *p<0.1; **p<0.05; ***p<0.01.

Table C.5. The association between absolute prediction errors and plan observables

	Dependent variable: Prediction errors			
	|Overall|	|Inpatient facility |	|OP Facility - Surgery |	|Professional |
	(1)	(2)	(3)	(4)
Overall prediction error$_{t-1}$	-0.001^{***}	-0.001^{***}	-0.0001^{***}	-0.0005^{***}
	(0.0001)	(0.00003)	(0.00001)	(0.00002)
Log previous enrollment	-6.270^{***}	-1.580^{*}	-0.083	-1.883^{***}
	(2.170)	(0.855)	(0.158)	(0.640)
Number of competitors	-5.345^{***}	-2.363^{***}	-0.281^{***}	-0.956^{***}
	(1.236)	(0.487)	(0.090)	(0.365)
Number of competing plans	0.634^{*}	0.336^{***}	-0.023	0.316^{***}
	(0.324)	(0.128)	(0.024)	(0.096)
Plan age	-43.944^{***}	-15.107^{***}	-4.044^{***}	-7.898^{***}
	(8.248)	(3.252)	(0.602)	(2.434)
Constant	888.563^{***}	314.236^{***}	68.322^{***}	246.858^{***}
	(182.162)	(71.815)	(13.303)	(53.755)
Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	11,530	11,530	11,530	11,530
R^2	0.261	0.259	0.218	0.282
Adjusted R^2	0.236	0.234	0.191	0.258

Notes: An observation is a plan-year. Estimates are obtained using unweighted OLS. Stars indicate p-values: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table C.6. The association between plan benefits and the overall prediction error

	Dependent variable:				
	Deductible	OOP limit	Primary care copay	Specialist copay	Hospital copay
	(1)	(2)	(3)	(4)	(5)
Overall prediction error$_{t-1}$	0.012*	0.373***	−0.001***	0.001***	0.008
	(0.007)	(0.052)	(0.0002)	(0.0002)	(0.013)
Star rating	−11.060	−1.900	0.319*	−0.882***	−37.499***
	(6.746)	(52.015)	(0.182)	(0.246)	(12.696)
Star 'too new' indicator	−43.882*	784.931***	0.545	−3.279***	−67.160
	(22.888)	(176.471)	(0.616)	(0.835)	(43.072)
Star 'no data' indicator	−56.314**	388.196**	1.477**	−3.496***	−150.724***
	(21.984)	(169.503)	(0.592)	(0.802)	(41.372)
Risk-adjusted payment	0.369***	0.588*	0.009***	−0.006***	0.010
	(0.041)	(0.314)	(0.001)	(0.001)	(0.077)
Constant	−194.165	1,169.898	−1.320	29.584***	983.983***
	(141.780)	(1,093.145)	(3.816)	(5.174)	(266.812)
Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Mean Dependent Variable	96.14722	3437.44	11.83048	26.0924	893.788
Observations	11,670	11,670	11,670	11,670	11,670
R^2	0.566	0.514	0.476	0.475	0.447
Adjusted R^2	0.551	0.498	0.459	0.458	0.429

Notes: An observation is a plan-year. The overall prediction error is defined as actual costs for the year minus the projected costs at the time of plan submission. Estimates are obtained using unweighted OLS. Stars indicate p-values: *p<0.1; **p<0.05; ***p<0.01.

Table C.7. The association between plan benefits and the absolute overall prediction error

	Dependent variable:				
	Deductible	OOP Limit	Primary care copay	Specialist copay	Hospital copay
	(1)	(2)	(3)	(4)	(5)
$\lvert\text{Prediction Error}\rvert_{t-1}$	0.002	-0.483^{***}	0.0004^{*}	-0.002^{***}	-0.053^{***}
	(0.008)	(0.059)	(0.0002)	(0.0003)	(0.014)
Star rating	−10.551	−5.256	0.306^{*}	-0.917^{***}	-39.096^{***}
	(6.748)	(51.987)	(0.182)	(0.246)	(12.690)
Star 'too new' indicator	-43.005^{*}	780.005^{***}	0.521	-3.335^{***}	−69.828
	(22.892)	(176.355)	(0.616)	(0.834)	(43.048)
Star 'no data' indicator	-55.382^{**}	390.855^{**}	1.445^{**}	-3.528^{***}	-152.776^{***}
	(21.985)	(169.373)	(0.592)	(0.801)	(41.344)
Risk-adjusted payment	0.355^{***}	0.636^{**}	0.010^{***}	-0.005^{***}	0.050
	(0.041)	(0.314)	(0.001)	(0.001)	(0.077)
Constant	−191.285	1,215.398	−1.455	29.616^{***}	981.348^{***}
	(141.791)	(1,092.351)	(3.818)	(5.167)	(266.643)
Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Mean Dependent Variable	96.14722	3437.44	11.83048	26.0924	893.788
Observations	11,670	11,670	11,670	11,670	11,670
R^2	0.566	0.515	0.475	0.477	0.448
Adjusted R^2	0.551	0.499	0.458	0.459	0.429

Notes: An observation is a plan-year. The overall prediction error is defined as actual costs for the year minus the projected costs at the time of plan submission. Estimates are obtained using unweighted OLS. Stars indicate p-values: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

REFERENCES CITED

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, *105*(490), 493–505.

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, *59*(2), 495–510.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American economic review*, *93*(1), 113–132.

Abaluck, J., Caceres Bravo, M., Hull, P., & Starc, A. (2021). Mortality effects and choice across private health insurance plans. *The quarterly journal of economics*, *136*(3), 1557–1610.

Abouk, R., Ghimire, K. M., Maclean, J. C., & Powell, D. (2021). *Does marijuana legalization affect work capacity? evidence from workers' compensation benefits* (Tech. Rep.). National Bureau of Economic Research.

Aizpurua-Olaizola, O., Soydaner, U., Öztürk, E., Schibano, D., Simsir, Y., Navarro, P., ... Usobiaga, A. (2016). Evolution of the cannabinoid and terpene content during the growth of cannabis sativa plants from different chemotypes. *Journal of Natural Products*, *79*(2), 324-331. Retrieved from `https://doi.org/10.1021/acs.jnatprod.5b00949` (PMID: 26836472) doi: 10.1021/acs.jnatprod.5b00949

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, *23*(4), 589–609.

Altman, E. I. (1973). Predicting railroad bankruptcies in america. *The Bell Journal of Economics and Management Science*, 184–211.

Alvarez, R., & Görg, H. (2009, January). Multinationals and plant exit: Evidence from Chile. *International Review of Economics & Finance*, *18*(1), 45-51.

Andrews, L. (2017, Jan). Legal weed is hard to come by in alaska. *Anchorage Daily News*. Retrieved from `https://www.adn.com/alaska-marijuana/2017/01/04/legal-weed-is-hard-to-come-by-in-alaska/`

Audretsch, D. B. (1991). New-firm survival and the technological regime. *The Review of Economics and Statistics*, *73*(3), 441–450.

Audretsch, D. B. (1995). The propensity to exit and innovation. *Review of Industrial Organization*, *10*(5), 589–605.

Audretsch, D. B., Houweling, P., & Thurik, A. R. (2000). Firm survival in the netherlands. *Review of Industrial Organization*, *16*(1), 1–11.

Audretsch, D. B., & Mahmood, T. (1995). New firm survival: new results using a hazard function. *The Review of Economics and Statistics*, *77*(1), 97–103.

Aydelotte, J. D., Brown, L. H., Luftman, K. M., Mardock, A. L., Teixeira, P. G., Coopwood, B., & Brown, C. V. (2017). Crash fatality rates after recreational marijuana legalization in washington and colorado. *American journal of public health*, *107*(8), 1329–1331.

Baggio, M., Chong, A., & Kwon, S. (2018). Marijuana and alcohol: Evidence using border analysis and retail sales data. *Canadian Journal of Economics/Revue canadienne d'économique*.

Baicker, K., & Goldman, D. (2011). Patient cost-sharing and healthcare spending growth. *Journal of Economic Perspectives*, *25*(2), 47–68.

Bampasidou, M., & Salassi, M. E. (2019). Agricultural labor trends. *Journal of ASFMRA*, 6–13.

Bandick, R. (2007). *Multinationals and plant survival in Swedish manufacturing* (Discussion Papers No. 07/31). University of Nottingham, GEP.

Barcott, B., & Whitney, B. (2019). *Special report: cannabis jobs count.* Leafly.

Bauhoff, S. (2014). The effect of school district nutrition policies on dietary intake and overweight: a synthetic control approach. *Economics & Human Biology*, *12*, 45–55.

Berenson, R. A., Sunshine, J. H., Helms, D., & Lawton, E. (2015). Why medicare advantage plans pay hospitals traditional medicare prices. *Health Affairs*, *34*(8), 1289–1295.

Berger, E., & Seegert, N. (2020). *Half banked: The real effects of financial exclusion on firms* (Tech. Rep.). Working paper.

Bernard, A. B., & Jensen, J. B. (2007). Firm structure, multinationals, and manufacturing plant deaths. *The Review of Economics and Statistics*, *89*(2), 193–204.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, *63*(4), pp. 841-890.

Billmeier, A., & Nannicini, T. (2013). Assessing economic liberalization episodes: A synthetic control approach. *Review of Economics and Statistics*, *95*(3), 983–1001.

Blakely, E. J., & Leigh, N. G. (2013). *Planning local economic development*. Sage.

Blonigen, B. A., Liebman, B. H., & Wilson, W. W. (2007, December). *Trade Policy and Market Power: The Case of the US Steel Industry* (NBER Working Papers). National Bureau of Economic Research, Inc.

Börgers, T., & Sarin, R. (1997). Learning through reinforcement and replicator dynamics. *Journal of economic theory*, *77*(1), 1–14.

Bram, J., & Gorton, N. (2017). *How is online shopping affecting retail employment?* (Tech. Rep.). Federal Reserve Bank of New York.

Brewin, D. G., Schmitz, T. G., Nolan, J. F., & Gray, R. S. (2017). Grain transportation policy reform in western canada. *Canadian Journal of Agricultural Economics*, *65*(4), 643–665.

Brigham, E. F. (1975). Hurdle rates for screening capital expenditure proposals. *Financial Management*, 17–26.

Butters, R. A., Sacks, D. W., & Seo, B. (2020). How do national firms respond to local shocks? *Available at SSRN 3535206*.

Cabral, M., Geruso, M., & Mahoney, N. (2018). Does privatized medicare benefit patients or producers? evidence from the benefits improvement and protection act. *American Economic Review*.

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2002). Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *Journal of Economic theory*, *104*(1), 137–188.

Canadian Transportation Agency. (2007). Decision No.344-R-2007 *https://otc-cta.gc.ca/eng/ruling/344-r-2007*.

Cannabis Business Times. (2020). *State of the industry report* (Tech. Rep.). Cannabis Business Times. Retrieved from `https://www.cannabisbusinesstimes.com/article/2020-state-of-the-industry-report-nexus-greenhouse-systems/`

Caulkins, J. P. (2010). *Estimated cost of production for legalized cannabis* (Tech. Rep.). RAND Drug Policy Research Center.

Cervantes, J. (2006). *Marijuana horticulture: The indoor/outdoor medical grower's bible*. Van Patten Publishing.

Chakraborty, A., Doremus, J., & Stith, S. (2020). *The effects of recreational cannabis access on the labor market: Evidence from colorado* (Tech. Rep.).

Chamon, M., Garcia, M., & Souza, L. (2017). Fx interventions in brazil: a synthetic control approach. *Journal of International Economics*, *108*, 157–168.

Chan, N. W., Burkhardt, J., & Flyr, M. (2020). The effects of recreational marijuana legalization and dispensing on opioid mortality. *Economic Inquiry*, *58*(2), 589–606.

Chandra, S., Lata, H., Khan, I. A., & ElSohly, M. A. (2017). Cannabis sativa l.: Botany and horticulture. In S. Chandra, H. Lata, & M. A. ElSohly (Eds.), *Cannabis sativa l. - botany and biotechnology* (pp. 79–100). Cham: Springer International Publishing. Retrieved from https://doi.org/10.1007/978-3-319-54564-6_3 doi: 10.1007/978-3-319-54564-6_3

Chelwa, G., van Walbeek, C., & Blecher, E. (2017). Evaluating south africa's tobacco control policy using a synthetic control method. *Tobacco Control*, *26*(5), 509–517.

Chen, M.-Y. (2002). Survival duration of plants: Evidence from the us petroleum refining industry. *International Journal of Industrial Organization*, *20*(4), 517–555.

Coibion, O., Gorodnichenko, Y., & Kumar, S. (2018). How do firms form their expectations? new survey evidence. *American Economic Review*, *108*(9), 2671–2713.

Colombo, M. G., & Delmastro, M. (2000). A note on the relation between size, ownership status and plant's closure: sunk costs vs. strategic size liability. *Economics Letters*, *69*(3), 421–427.

Commission, M. P. A., et al. (2021). The medicare advantage program: status report. In *Report to the congress: Medicare payment policy.*

Crawford, V. P. (2003). Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions. *American Economic Review*, *93*(1), 133–149.

Curto, V., Einav, L., Levin, J., & Bhattacharya, J. (2021). Can health insurance competition work? evidence from medicare advantage. *Journal of Political Economy*, *129*(2), 570–606.

de Figueiredo, J. M., & Silverman, B. S. (2012). Firm survival and industry evolution in vertically related populations. *Management Science*, *58*(9), 1632–1650.

De Vany, A. S., & Saving, T. R. (1977). Product quality, uncertainty, and regulation: The trucking industry. *The American Economic Review*, *67*(4), 583–594.

Doraszelski, U., & Markovich, S. (2007). Advertising dynamics and competitive advantage. *The RAND Journal of Economics*, *38*(3), 557–592.

Dragone, D., Prarolo, G., Vanin, P., & Zanella, G. (2019). Crime and the legalization of recreational marijuana. *Journal of economic behavior & organization*, *159*, 488–501.

Dranove, D., & Satterthwaite, M. A. (2000). The industrial organization of health care markets. *Handbook of health economics*, *1*, 1093–1139.

Duggan, M., Starc, A., & Vabson, B. (2016). Who benefits when the government pays more? pass-through in the medicare advantage program. *Journal of Public Economics*, *141*, 50 - 67. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0047272716300767` doi: https://doi.org/10.1016/j.jpubeco.2016.07.003

Duncan, D., Ross, J. M., & Mikesell, J. L. (2019). Can the federal budget process promote fiscal sustainability? evidence from synthetic control using lasso. *Evidence from Synthetic Control Using Lasso (December 4, 2019)*.

Dunne, P., & Hughes, A. (1994, June). Age, Size, Growth and Survival: UK Companies in the 1980s. *Journal of Industrial Economics*, *42*(2), 115-140.

Dunne, T., Klimek, S. D., & Roberts, M. J. (2005). Exit from regional manufacturing markets: The role of entrant experience. *International Journal of Industrial Organization*, *23*(5-6), 399–421.

Dunne, T., Roberts, M. J., & Samuelson, L. (1988, Winter). Patterns of Firm Entry and Exit in U.S. Manufacturing Industries. *RAND Journal of Economics*, *19*(4), 495-515.

Dunne, T., Roberts, M. J., & Samuelson, L. (1989). The growth and failure of us manufacturing plants. *The Quarterly Journal of Economics*, *104*(4), 671–698.

Elston, J. A., & Agarwal, R. (2004). Bank influence, firm performance and survival: Empirical evidence from germany 1970-1986. *Corporate Ownership and Control*, *1*(2), 65–71.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, 848–881.

Ericson, R., & Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, *62*(1), 53–82.

Escalante, C. L., & Luo, T. (2017). Sustaining a healthy farm labor force: issues for policy consideration. *Choices*, *32*(1), 1–9.

Ferman, B. (2021). On the properties of the synthetic control estimator with many periods and many controls. *Journal of the American Statistical Association*, *116*(536), 1764–1772.

Ferman, B., & Pinto, C. (2021). Synthetic controls with imperfect pretreatment fit. *Quantitative Economics*, *12*(4), 1197–1221.

Ferman, B., Pinto, C., & Possebom, V. (2020). Cherry picking with synthetic controls. *Journal of Policy Analysis and Management*, *39*(2), 510–532.

Figueroa, J. F., Wadhera, R. K., Frakt, A. B., Fonarow, G. C., Heidenreich, P. A., Xu, H., … others (2020). Quality of care and outcomes among medicare advantage vs fee-for-service medicare patients hospitalized with heart failure. *JAMA cardiology*, *5*(12), 1349–1357.

Firpo, S., & Possebom, V. (2016). Synthetic control estimator: A generalized inference procedure and confidence sets. *Sao Paulo School of Economics, EESP-FGV*.

Franklin, P. J. (1974). Some observations on exit from the motor insurance industry, 1966-1972. *The Journal of Industrial Economics*, *22*(4), 299–313.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, *33*(1), 1.

Frittelli, J. (2005). Grain transport: modal trends and infrastructure implications..

Fudenberg, D., & Kreps, D. M. (1993). Learning mixed equilibria. *Games and economic behavior*, *5*(3), 320–367.

Fudenberg, D., & Tirole, J. (1986, July). A Theory of Exit in Duopoly. *Econometrica*, *54*(4), 943-960.

Gaynor, M., Ho, K., & Town, R. J. (2015). The industrial organization of health-care markets. *Journal of Economic Literature*, *53*(2), 235–84.

Ghemawat, P., & Nalebuff, B. (1985, Summer). Exit. *RAND Journal of Economics*, *16*(2), 184-194.

Ghemawat, P., & Nalebuff, B. (1990). The devolution of declining industries. *The Quarterly Journal of Economics*, *105*(1), 167–186.

Ghimire, K. M., & Maclean, J. C. (2020). Medical marijuana and workers' compensation claiming. *Health economics*, *29*(4), 419–434.

Gibson, J. K., & Harris, R. I. D. (1996, August). Trade Liberalisation and Plant Exit in New Zealand Manufacturing. *The Review of Economics and Statistics*, *78*(3), 521-529.

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of accounting and economics*, *40*(1-3), 3–73.

Hall, W., & Lynskey, M. (2016). Evaluating the public health impacts of legalizing recreational cannabis use in the united states. *Addiction*, *111*(10), 1764–1773.

Hansen, B., Miller, K., Seo, B., & Weber, C. (2020). Taxing the potency of sin goods: Evidence from recreational cannabis and liquor markets. *National Tax Journal, Forthcoming*.

Hansen, B., Miller, K., & Weber, C. (2017). *The taxation of recreational marijuana: Evidence from washington state* (Tech. Rep.). National Bureau of Economic Research.

Hansen, B., Miller, K., & Weber, C. (2018). Auditing and enforcement in the recreational marijuana industry. In *111th annual conference on taxation.*

Hansen, B., Miller, K., & Weber, C. (2020a). Early evidence on recreational marijuana legalization and traffic fatalities. *Economic inquiry*, *58*(2), 547–568.

Hansen, B., Miller, K., & Weber, C. (2020b). Federalism, partial prohibition, and cross-border sales: Evidence from recreational marijuana. *Journal of Public Economics*, *187*, 104159.

Hansen, B., Miller, K., & Weber, C. (2021a). Up in smoke? the market for cannabis. In K. F. Zimmermann (Ed.), *Handbook of labor, human resources and population economics* (pp. 1–29). Cham: Springer International Publishing. Retrieved from `https://doi.org/10.1007/978-3-319-57365-6_138-1` doi: 10.1007/978-3-319-57365-6_138-1

Hansen, B., Miller, K. S., & Weber, C. (2021b). *Vertical integration and production inefficiency in the presence of a gross receipts tax* (Tech. Rep.). National Bureau of Economic Research.

Hao, Z., & Cowan, B. W. (2020). The cross-border spillover effects of recreational marijuana legalization. *Economic inquiry*, *58*(2), 642–666.

Hart, S., & Mas-Colell, A. (2000). A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, *68*(5), 1127–1150.

Hollenbeck, B., & Uetake, K. (2019). Taxation and market power in the legal marijuana industry. *Available at SSRN 3237729*.

Holmes, S. (2013). *Fresh fruit, broken bodies: Migrant farmworkers in the united states* (Vol. 27). Univ of California Press.

Huang, W. (2003). *Shuttle train adoption strategy* (Unpublished master's thesis). North Dakota State University, Fargo, North Dakota. (North Dakota State University Press)

Hughes, L. A., Schaible, L. M., & Jimmerson, K. (2020). Marijuana dispensaries and neighborhood crime and disorder in denver, colorado. *Justice Quarterly*, *37*(3), 461–485.

Ifrach, B., & Weintraub, G. Y. (2016, 09). A Framework for Dynamic Oligopoly in Concentrated Industries. *The Review of Economic Studies*, *84*(3), 1106-1150. Retrieved from `https://doi.org/10.1093/restud/rdw047` doi: 10.1093/restud/rdw047

Jena, A. B., & Philipson, T. J. (2008). Cost-effectiveness analysis and innovation. *Journal of health economics*, *27*(5), 1224–1236.

Jordan, M. (2020, Apr). Farmworkers, mostly undocumented, become 'essential' during pandemic. *The New York Times*. Retrieved from `https://www.nytimes.com/2020/04/02/us/coronavirus-undocumented-immigrant-farmworkers-agriculture.html`

Jovanovic, B. (1982, May). Selection and the Evolution of Industry. *Econometrica*, *50*(3), 649-670.

Kerr, D. C., Bae, H., Phibbs, S., & Kern, A. C. (2017). Changes in undergraduates' marijuana, heavy alcohol and cigarette use following legalization of recreational marijuana use in oregon. *Addiction*, *112*(11), 1992–2001.

Keys, A. (2020, January 17). Gov. lujan grisham unveils recreational marijuana legalization proposal. *KRQE Albequerque*. Retrieved 2020/07/03, from `https://www.krqe.com/news/politics-government/gov-lujan-grisham-unveils-recreational-marijuana-legalization-proposal/`

Kilmer, B. (2019). How will cannabis legalization affect health, safety, and social equity outcomes? it largely depends on the 14 ps. *The American journal of drug and alcohol abuse*, *45*(6), 664–672.

Kim, H. (2021). Heterogeneous impacts of cost shocks, strategic bidding and pass-through: Evidence from the new england electricity market. *American Economic Journal: Microeconomics*.

Kimura, F., & Kiyota, K. (2006). Exports, fdi, and productivity: Dynamic evidence from japanese firms. *Review of World Economics*, *142*(4), 695–719.

Krissman, F. (2017). *America's largest cannabis labor market.* Humboldt State University, Department of Anthropology.

Larson, E. K., & James, N. F. (2007). Assessing the Effects of Road Pricing on an Industry: The Grain Handling and Transportation System in Saskatchewan. *Journal of the Transportation Research Forum*, *46*(3), 1-19. doi: 10.22004/ag.econ.206895

Larson, E. K., & Nolan, J. F. (2007). Assessing the effects of road pricing on an industry: the grain handling and transportation system in saskatchewan. *Journal of the Transportation Research Forum*, *46*, 69–86.

Lawrence, R., Nolan, J., & Schoney, R. (2016). Simulating contestability in freight transportation: the canadian grain handling and transportation system. *Journal of Transport Economics and Policy*, *50*, 1–25.

Lieberman, M. B. (1990, Winter). Exit from declining industries: 'shakeout' or 'stakeout'? *RAND Journal of Economics*, *21*(4), 538-554.

Liesch, P. W., & Knight, G. A. (1999). Information internalization and hurdle rates in small and medium enterprise internationalization. *Journal of International Business Studies*, *30*(2), 383–394.

Liu, Z.-N., Yu, X.-Y., Jia, L.-F., Wang, Y.-S., Song, Y.-C., & Meng, H.-D. (2021). The influence of distance weight on the inverse distance weighted method for ore-grade estimation. *Scientific Reports*, *11*(1), 1–8.

Luckstead, J., & Devadoss, S. (2019). The importance of h-2a guest workers in agriculture. *Choices*, *34*(1), 1–8.

Maclean, J. C., Ghimire, K. M., & Nicholas, L. H. (2021). Marijuana legalization and disability claiming. *Health Economics*, *30*(2), 453–469.

Mata, J., Portugal, P., & Guimaraes, P. (1995). The survival of new plants: Start-up conditions and post-entry evolution. *International Journal of Industrial Organization*, *13*(4), 459–481.

McGuire, T. G., Newhouse, J. P., & Sinaiko, A. D. (2011). An economic history of medicare part c. *The Milbank Quarterly*, *89*(2), 289–332.

Meyers, D. J., & Trivedi, A. N. (2021). Medicare advantage chart reviews are associated with billions in additional payments for some plans. *Medical Care*, *59*(2), 96–100.

Meyers, D. J., Trivedi, A. N., Wilson, I. B., Mor, V., & Rahman, M. (2021). Higher medicare advantage star ratings are associated with improvements in patient outcomes: Study examines how medicare quality ratings impact patient outcomes for those enrolled in medicare advantage. *Health Affairs*, *40*(2), 243–250.

Michalowsky, B., Hoffmann, W., Kennedy, K., & Xie, F. (2020). Is the whole larger than the sum of its parts? impact of missing data imputation in economic evaluation conducted alongside randomized controlled trials. *The European Journal of Health Economics*, *21*(5), 717–728.

Milgrom, P., & Roberts, J. (1991). Adaptive and sophisticated learning in normal form games. *Games and economic Behavior*, *3*(1), 82–100.

Miller, A. M., Rosenman, R., & Cowan, B. W. (2017). Recreational marijuana legalization and college student use: Early evidence. *SSM-Population Health*, *3*, 649–657.

Miller, K., & Seo, B. (2021). The effect of cannabis legalization on substance demand and tax revenues. *National Tax Journal*, *74*(1), 107–145.

Miller, K. S., Petrin, A., Town, R., & Chernew, M. (2021). *Optimal managed competition subsidies* (Tech. Rep.). National Bureau of Economic Research.

Miller, K. S., & Wilson, W. W. (2018). Governance structure and exit: Evidence from california hospitals. *Review of Industrial Organization*, *53*(1), 31–55.

Motel, S. (2015). Six facts about marijuana. *Pew Research Fact Tank*. Retrieved 10/22/2016, from `http://www.pewresearch.org/fact-tank/2015/04/14/6-facts-about-marijuana/`

Murray, J. S., & Reiter, J. P. (2016). Multiple imputation of missing categorical and continuous values via bayesian mixture models with local dependence. *Journal of the American Statistical Association*, *111*(516), 1466–1479.

Neumark, D., Zhang, J., & Ciccarella, S. (2008). The effects of wal-mart on local labor markets. *Journal of Urban Economics*, *63*(2), 405–430.

Nicholas, L. H., & Maclean, J. C. (2019). The effect of medical marijuana laws on the health and labor supply of older adults: Evidence from the health and retirement study. *Journal of Policy Analysis and Management*, *38*(2), 455–480.

Nolan, J. (2007). Saskatchewan: Geographic perspectives. In *Transportation* (chap. 19). Regina, SK: Canadian Plains Research Centre Press.

Nolan, J. F., & Kerr, W. A. (2012). The estey review's contribution to transportation policy. *Journal of International Law and Trade Policy*, *13*(2), 243–262.

Park, S., White, L., Fishman, P., Larson, E. B., & Coe, N. B. (2020). Health care utilization, care satisfaction, and health status for medicare advantage and traditional medicare beneficiaries with and without alzheimer disease and related dementias. *JAMA network open*, *3*(3), e201809–e201809.

Prater, M. E., Sparger, A., Bahizi, P., & O'Neil, D. J. (2013). Rail Market Share of Grain and Oilseed Transportation. *Journal of the Transportation Research Forum*, *52*(2), 1-25. doi: 10.22004/ag.econ.207350

Reynolds, S. S. (1988). Plant closings and exit behaviour in declining industries. *Economica*, *55*(220), 493–503.

Rutström, E. E., & Wilcox, N. T. (2009). Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test. *Games and Economic Behavior*, *67*(2), 616–632.

Sabia, J. J., & Nguyen, T. T. (2018). The effect of medical marijuana laws on labor market outcomes. *The Journal of Law and Economics*, *61*(3), 361–396.

Sacirbey, O. (2016, Dec). Testing, supply issues 'could be death knell' for oregon mj market. *MJBizDaily*. Retrieved from `https://mjbizdaily.com/testing-supply-issues-could-be-death-knell-for-oregon-mj-market/`

Sarmiento, C., & Wilson, W. W. (2005). Spatial modeling in technology adoption decisions: The case of shuttle train elevators. *American Journal of Agricultural Economics*, *87*(4), 1034–1045.

Schaneman, B. (2019). Marijuana, hemp cross-pollination a budding conflict. *AP News*. Retrieved from https://apnews.com/article/wa-state-wire-ca-state-wire-jay-inslee-cannabis-marijuana-eded257dc4de4ce19a5388a7acbf7c38

Schmiesing, B. H., Blank, S. C., & Gunn, S. P. (1985). The influence of technological change on grain elevator pricing efficiency. *North Central Journal of Agricultural Economics*, *7*(2), 95–107.

Segarra, A., & Callejón, M. (2002). New firms' survival and market turbulence: New evidence from spain. *Review of Industrial Organization*, *20*(1), 1–14.

Selyem, B. K. (2000). The legacy of country grain elevators: A photo essay. *Kansas History*, *23*(1-2).

Shepard, H., Parker, J., Darby, P., & Ainsworth, C. C. (1999). Sex expression in hop (humulus lupulus l. and h. japonicus sieb. et zucc.): floral morphology and sex chromosomes. *Sex determination in plants*, *8*, 139–50.

Sills, E. O., Herrera, D., Kirkpatrick, A. J., Brandão Jr, A., Dickson, R., Hall, S., … others (2015). Estimating the impacts of local policy innovation: the synthetic control method applied to tropical deforestation. *PloS one*, *10*(7), e0132590.

Simmins, G. (2004). Prairie grain elevators: An old purpose in search of a new form. *Challenging Frontiers*, 205-230.

Smart, R., & Pacula, R. L. (2019). Early evidence of the impact of cannabis legalization on cannabis use, cannabis use disorder, and the use of other substances: findings from state policy evaluations. *The American journal of drug and alcohol abuse*, *45*(6), 644–663.

Smith, V. M., Powell, M., Mungeam, D., & Emmons, R. G. (2019). Stakeholder perceptions of the impact of cannabis production on the southern oregon food system. *Journal of Agriculture, Food Systems, and Community Development*, *8*(4), 125–135.

Song, Z., Landrum, M. B., & Chernew, M. E. (2013). Competitive bidding in medicare advantage: Effect of benchmark changes on plan bids. *Journal of health economics*, *32*(6), 1301–1312.

Starfield, B., Shi, L., & Macinko, J. (2005). Contribution of primary care to health systems and health. *The milbank quarterly*, *83*(3), 457–502.

Stoicheff, L. (2018, April 3). The impact of santa cruz county agriculture's labor shortage. *Good Times Santa Cruz*. Retrieved 2020/07/03, from https://goodtimes.sc/cover-stories/impact-santa-cruz-county-agriculture-labor-shortage/

Thomas, D. (2018). License quotas and the inefficient regulation of sin goods: Evidence from the washington recreational marijuana market. *Available at SSRN 3312960*.

Thomas-Lycklama-Nijeholt, G. (2012). *On the road for work: Migratory workers on the east coast of the united states* (Vol. 7). Springer Science & Business Media.

Thraves, B. D. (2007). *Saskatchewan: Geographic perspectives* (No. 52). University of Regina Press.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Ullman, D. F. (2017). The effect of medical marijuana on sickness absence. *Health economics*, *26*(10), 1322–1327.

Vachal, K., & Bitzan, J. (1997). *Implications of a north american grain marketing system for prairie transportation and elevators* (Tech. Rep.). Upper Great Plains Transportation Institute, North Dakota State University.

Valachovic, Y., Quinn-Davidson, L., Stackhouse, J., Butsic, V., et al. (2019). Perceptions of cannabis among humboldt county timberland and ranchland owners. *California Agriculture*, *73*(3), 161–168.

Van Beveren, I. (2007). Footloose multinationals in belgium? *Review of World Economics*, *143*(3), 483–507.

Van Buuren, S., Brand, J. P., Groothuis-Oudshoorn, C. G., & Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. *Journal of statistical computation and simulation*, *76*(12), 1049–1064.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2010). mice: Multivariate imputation by chained equations in r. *Journal of statistical software*, 1–68.

van Ours, J. C., & Williams, J. (2015). Cannabis use and its effects on health, education and labor market success. *Journal of Economic Surveys*, *29*(5), 993–1010.

Wallace, A. (2020, June 25). Why more states could legalize cannabis in 2021. *CNN Business*. Retrieved 2020/07/03, from https://www.cnn.com/2020/06/24/business/cannabis-legalize-2021-recession/index.html

Wallace, D. (1997). *Positioning for the future: Canadian shipper perspective.* Customized Ag Production & Marketing Conference. (Unpublished presentation)

Washburn, K. (2020, February 5). Wine to weed: Wine grape growers lose labor to cannabis farms. *The Sun Gazette*. Retrieved 2020/07/03, from `https://thesungazette.com/article/news/2020/02/05/wine-grape-growers-are-losing-labor-to-commercial-cannabis-farms/`

Whinston, M. D. (1988). Exit with multiplant firms. *The RAND Journal of Economics*, *19*(4), 568–588.

White, T. K., Reiter, J. P., & Petrin, A. (2012). *Plant-level productivity and imputation of missing data in us census manufacturing data* (Tech. Rep.). National Bureau of Economic Research.

White, T. K., Reiter, J. P., & Petrin, A. (2018). Imputation in us manufacturing data and its implications for productivity dispersion. *Review of Economics and Statistics*, *100*(3), 502–509.

Zingales, L. (1998). Survival of the fittest or the fattest? exit and financing in the trucking industry. *The Journal of Finance*, *53*(3), 905–938.