UNIVERSITY OF CALIFORNIA
Los Angeles

Essays on Coordination and Economic Efficiency

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Economics

by

Juan S. Rojas Bohorquez

2022
ABSTRACT OF THE DISSERTATION

Essays on Coordination and Economic Efficiency

by

Juan S. Rojas Bohorquez
Doctor of Philosophy in Economics
University of California, Los Angeles, 2022
Professor Dora L. Costa, Chair

While competition has long been recognized in economic literature as a tenet of economic efficiency, the link between coordination and efficiency is yet to be fully understood. In this dissertation, I study three separate episodes that contribute to our understanding on the ambiguous link between coordination and efficiency. The first episode is set in the early stages of development of the wine industry in California. I show how the disclosure of information fostered coordination between winemakers and vintners, which improved the efficiency in the use of resources and played a crucial role in the industrial growth of one of the major agro-industrial businesses in the state. The second episode presents a negative relationship between coordination and efficiency. We show how a loophole in Medicaid has allowed state and local governments to coordinate with providers to maximize federal matching funds at the expense of efficiency in the provision of care. Lastly, using evidence from pension funds in Colombia we show how ownership acts as a conduit for coordination

ii
between firms having a mixed effect on the efficient use of financial assets.

Chapter 1 studies the effects of information disclosure on coordination and, ultimately, on industrial growth, using evidence from the wine and grape industry in the United States. Between 1960 and 1968, the Raisin Lay Survey gave accurate, weekly information about the size and allocation of the raisin harvest to grape growers, packers, and winemakers. Relying on data from 1950 to 1970 at county-year level from multiple newly digitized sources, I estimate a difference-in-differences model that exploits the fact that the Survey was, for budgetary reasons, only implemented in a reduced area compared with the original plan. The results show that information disclosure led to larger wine production in surveyed areas and increased investment from both grape growers and winemakers. Furthermore, I provide empirical evidence that the main mechanism driving these results is the enhanced coordination between growers and winemakers that resulted from the information disclosure.

Chapter 2 shows theoretically and empirically how loopholes in the institutional design of joint Medicaid funding can lead to coordination between local governments and providers that result in price and volume distortions. Our empirical analysis combines audit, survey, and administrative datasets on skilled nursing facilities (SNFs) from 1999-2017 with two reforms and difference-in-differences models. We first document that states use creative financing schemes to divert federal Medicaid matching funds. Using the case study of Indiana, we then document that these schemes lead to an increase in Medicaid SNF days for dementia patients. The expansion of SNFs of lower match quality leads to an increase in mortality pointing to a misallocation of vulnerable patients to providers.

Finally, Chapter 3 studies the distortions of common ownership on the portfo-
lio allocation of financial entities. We use evidence from Colombian pension funds, which is a highly concentrated market, the largest entities are integrated with large financial conglomerates and manage funds that amount to roughly 25% of GDP and 87% of the market capitalization of the Colombian stock exchange. We use a rich database on the daily portfolio positions of all pension funds, which is collected by the Financial Superintendence of Colombia. We exploit the variation in ownership coming from two mergers where two of the largest pension funds that are part of financial conglomerates acquire other pension funds. We use a difference-in-differences specification to test whether after the merger the pension funds disproportionately increased their holdings of assets issued by other firms in their financial conglomerate. The preliminary results suggest that pension funds disproportionately favor investments in related firms as merge firms more than double their share of commonly owned assets in their portfolios. This reduced-form evidence lays the building blocks to a structural analysis that allows decomposing the observed changes between ownership, market power, and changes in expectations, which would ultimately shed light on the efficiency of the allocation of financial assets.
The dissertation of Juan S. Rojas Bohorquez is approved.

Simon A. Board
Michela Giorcelli
Martin B. Hackmann

Dora L. Costa, Committee Chair

University of California, Los Angeles
2022
Para Paula—el amor de mi vida—
mes papás, mi hermana,
y para resto de mi familia.
Sin ellos, nunca lo hubiera logrado.
Los amo con todas mis fuerzas.
# TABLE OF CONTENTS

1 Does Information Disclosure Foster Industrial Growth? Evidence from the Early California Grape and Wine Industry .............................................. 1

1.1 Introduction ........................................................................................ 2

1.2 The California Grape and Wine Industry (1950-1968) ......................... 7

1.2.1 Institutional Context of the Grape and Wine Industry ................ 7

1.2.2 The Coordination Problem ............................................................ 9

1.2.3 Intervention: The Raisin Lay Survey ........................................... 11

1.3 Data .................................................................................................... 13

1.3.1 Production of Wine ....................................................................... 13

1.3.2 Investment of Upstream Grape Growers ..................................... 14

1.3.3 Investment of Downstream Winemakers .................................... 15

1.4 Empirical Design ................................................................................ 16

1.4.1 Balancing Tests ............................................................................ 20

1.4.2 Parallel Trends Before the Raisin Lay Survey (1950-1959) ........ 21

1.4.3 Placebo Test: Investment in Other Fruit Tress ............................ 22

1.5 Effects of the Raisin Lay Survey on Production and Investment ........... 22

1.5.1 Wine Production .......................................................................... 22

1.5.2 Investment of Upstream Grape Growers .................................... 26

1.5.3 Investment of Downstream Winemakers .................................... 29
2 Creative Financing and Public Moral Hazard: Evidence from Medicaid Supplemental Payments

2.1 Introduction

2.2 Institutional Setting

2.2.1 Provider Reimbursement and Creative Financing Pre-2003

2.2.2 The 2003 Reform of Supplemental Payments

2.2.3 Creative Financing 2.0: Evidence from Indiana

2.3 Conceptual Framework

2.3.1 Baseline Model: No Creative Financing

2.3.2 Joint Funding Under Creative Financing Mechanisms Pre 2003

2.3.3 Creative Financing Mechanisms by Ownership Type

2.4 Data, Empirical Approach, and Results

2.4.1 Creative Financing in U.S. Nursing Homes

2.4.2 Distortionary Effects of Creative Financing

2.4.3 Mechanisms and Patient Outcomes

2.5 Discussion and Conclusion

2.A Appendix: Additional Figures and Tables

2.B Principal Agent Conflict: County Hospitals and Indiana
3 Common Ownership and Distorted Portfolio Allocation: Evidence from Colombian Pension Funds

3.1 Introduction .................................................. 107
3.2 Pension Funds in Colombia ................................. 110
3.3 Data ............................................................. 111
   3.3.1 Daily Portfolio Position ................................ 111
   3.3.2 Balance Sheet Data of Pension Funds ............... 112
   3.3.3 Ownership of Financial Institutions ................. 113
   3.3.4 Ownership of Non-Financial Institutions .......... 113
   3.3.5 Transition Matrices of Pension Fund Savers ........ 114
   3.3.6 News Archives ........................................... 114
3.4 Empirical Design ............................................. 115
3.5 Reduced-Form Evidence on Portfolio Allocation Decisions after Changes in Ownership ........................................ 116
3.6 Conclusion and Discussion .................................. 124
# LIST OF FIGURES

1.1 Vertical supply chain of the grape-wine industry in California .......................... 9

1.2 Map of California Wine Districts ................................................................. 19

1.3 Effects of the Raisin Lay Survey on Wine Production ..................................... 25

1.4 Effects of the Raisin Lay Survey on Grape Growers’ Vineyard Acreage
   by Type of Grape ......................................................................................... 28

1.5 Spillover Effects of the Raisin Lay Survey on Winemakers’ Investment .............. 33

1.A.1 Map of the Fresno and Madera Areas Photographed in the Raisin Lay
   Survey ........................................................................................................ 39

1.A.2 Profit per ton of raisins shipping from Fresno to different locations in
   California ........................................................................................................ 40

1.A.3 Map of California Regions ........................................................................... 41

1.A.4 Histogram of Placebo Coefficients for the Effects of the Raisin Lay
   Survey on Grape Growers’ Vineyard Acreage by Type of Grape ................. 42

1.A.5 Map of the California Wine Districts ............................................................ 43

2.1 Public Markdowns and the Supplemental Payments Reform .............................. 73

2.2 Effect of 2003 Reform on Nursing Homes ..................................................... 75

2.3 Mechanisms: Dementia Patients in the MDS SNF Sample Population ............. 80

2.4 Mechanisms: Dementia Patients in the MedPAR Hospital Sample Popu-
   lation ............................................................................................................. 83

2.A.1 Schemes using UPL Supplemental Payments for Nursing Facilities ......... 87
2.A.2 States Use of UPL-SPs in Nursing Homes in FY 2002

2.A.3 Public Markdowns and the Effective FMAP: Evidence from LTC Focus
2000-2002

2.A.4 Development of Shares of County-Owned Nursing Homes

2.C.1 Public Markdowns and the Effective FMAP—Evidence from Audit
Reports

2.D.1 Public Markdowns and the Effective FMAP: Evidence from the HHC
of Marion County
## LIST OF TABLES

- **1.1** Effects of the Raisin Lay Survey on Wine Production ........................................ 23
- **1.2** Effects of the Raisin Lay Survey on Grape Growers’ Investment .......................... 27
- **1.3** Effects of the Raisin Lay Survey on Winemakers’ Investment ............................... 30
- **1.4** Effects of the Raisin Lay Survey on the Proxy Measures of Coordination between Winemakers and Grape Growers ......................................................... 36
- **1.A.1** Verifying Balance in Vineyard Characteristics in Surveyed and Non-surveyed Counties ............................................................................................................. 44
- **1.A.2** Verifying Balance in Vineyard Characteristics in Madera and Non-surveyed Counties .................................................................................................................. 45
- **1.A.3** Time Trends for Wine Production by Type before the Raisin Lay Survey for Surveyed and Non-Surveyed Counties, 1950-1959 ..................................................... 46
- **1.A.4** Time Trends for Grape Crush by Variety before the Raisin Lay Survey for Surveyed and Non-Surveyed Counties, 1950-1959 ..................................................... 47
- **1.A.5** Time Trends for Bearing Acreage by Variety before the Raisin Lay Survey for Surveyed and Non-Surveyed Counties, 1950-1959 ..................................................... 48
- **1.A.6** Time Trends for Wineries’ Vineyards and Capacity before the Raisin Lay Survey for Surveyed and Non-Surveyed Counties, 1950-1959 ..................................................... 49
- **1.A.7** Placebo Tests for the Acreage of Bearing Fruit Trees Different from Grapes ................. 50
- **1.A.8** Effects of the Raisin Lay Survey on Grape Crush .................................................. 51
1.A.9  Effects of the Raisin Lay Survey on Wine Production using Different Samples for Control and Treatment Groups .......................... 52
1.A.10 Effects of the Raisin Lay Survey on Grape Crush Using Different Samples for Control and Treatment Groups .......................... 53
1.A.11 Effects of the Raisin Lay Survey on Total Grape Vine Acreage Using Different Samples for Control and Treatment Groups .............. 54
1.A.12 Effects of the Raisin Lay Survey on Winemakers’ Investment Using a Balanced Panel ................................................................. 55
1.A.13 Spillover Effects of the Raisin Lay Survey on Winemakers’ Investment 56

2.A.1  LTCFocus Summary Statistics .................................................. 91
2.A.2  Characteristics of Early (<2011) vs. Late SNF Acquisitions (>=2011) 92
2.A.3  Effects of 2003 Reform on Various SNF State-Level Outcomes ....... 94
2.A.4  Difference-in-Differences Models with Gardner 2sDD Correction ...... 95
2.A.5  Mechanisms—Alzheimer Patients in SNFs ................................. 96
2.A.6  Mechanisms—Hospital Discharges of Dementia Patients ............... 97

3.1  Change in Portfolio Shares after the Merger between Povenir and Horizonte (Grupo Aval) .............................................................. 117
3.2  Change in Portfolio Shares after the Merger between Proteccion and ING (Grupo Sura) ............................................................... 118
3.3  Change in Portfolio Shares after the Merger between Povenir and Horizonte (Grupo Aval), by Type of Asset ................................. 119
3.4 Change in Portfolio Shares after the Merger between Proteccion and ING (Grupo Sura), by Type of Asset ........................................ 120

3.5 Change in Portfolio Shares after the Merger between Povenir and Horizonte (Grupo Aval), by Type of Portfolio ................................. 122

3.6 Change in Portfolio Shares after the Merger between Proteccion and ING (Grupo Sura), by Type of Portfolio ........................................ 123
ACKNOWLEDGMENTS

Chapter 2 is based on a paper in preparation for publication with Martin B. Hackmann and Nicolas R. Ziebarth. Chapter 3 is a preliminary version of a paper currently in preparation for publication with Wilmar A. Cabrera.
VITA

JUAN S. ROJAS BOHORQUEZ

EDUCATION

2013-2016. M.A. Economics, Universidad de los Andes, Bogota, Colombia
2008-2013. B.A. Economics, Universidad de los Andes, Bogota, Colombia

WORKING PAPERS

Do Information Disclosures Foster Industrial Growth? Evidence from the Early California Grape and Wine Industry

Creative Financing and Public Moral Hazard: Evidence from Medicaid Supplemental Payments (with Martin Hackmann and Nicolas R. Ziebarth)

Market Power, Common Ownership, and Distorted Portfolio Allocation (with Wilmar Cabrera)

CHAPTERS IN BOOKS


SEMINARS AND CONFERENCES

Do Information Disclosures Foster Industrial Growth?: Rice University (2022), Bryn Mawr College (2022), Universidad Javeriana (2022), Compass Lexecon (2022), Analysis Group (2022), The Brattle Group (2022), Charles Rivers Associates (2022), xvi
NBER Summer Institute (poster session 2021), All-UC Economic History Workshop (2021), Economic History Association (poster session 2021)


**FELLOWSHIPS, GRANTS, AND AWARDS**

2016 - 2022. Ph.D. Fellowship, Department of Economics, UCLA.

2021, 2019, 2018. Best Teaching Assistant Award, UCLA.


2020. Excellence in Online Teaching Award, UCLA.


2020. Lewis L. Clarke Graduate Fellowship, UCLA.

2014. Best Teaching Assistant Award, Universidad de los Andes.

**TEACHING EXPERIENCE**

*Instructor at UCLA:* Microeconomic Theory

*Teaching Assistant at UCLA:* Microeconomic Theory, Game Theory, Industrial Organization, Economics of Technology and E-Commerce, Organization of Firms, Introduction to Econometrics, Development of Economic Institutions in Western Europe.

*Teaching Assistant at Universidad de los Andes:* Game Theory (M.A. and Ph.D. level), Introduction to Macroeconomics.
CHAPTER 1

Does Information Disclosure Foster Industrial Growth? Evidence from the Early California Grape and Wine Industry

This chapter studies the effects of information disclosure on industrial growth, using evidence from the wine and grape industry in the United States. Between 1960 and 1968, the Raisin Lay Survey gave accurate, weekly information about the size and allocation of the raisin harvest to grape growers, packers, and winemakers. Relying on data from 1950 to 1970 at county-year level from multiple newly digitized sources, I estimate a difference-in-differences model that exploits the fact that the Survey was, for budgetary reasons, only implemented in a reduced area compared with the original plan. The results show that information disclosure led to larger wine production in surveyed areas and increased investment from both grape growers and winemakers. Furthermore, I provide empirical evidence that the main mechanism driving these results is the enhanced coordination between growers and winemakers that resulted from the information disclosure.
1.1 Introduction

The supply chain for an industry, even narrowly defined, involves the coordination of multiple firms that must make decisions under uncertainty about the actions of other participants. In some markets, especially agricultural markets, asymmetric information and communication frictions may reduce the effectiveness of prices as coordination devices (Minarelli et al., 2016). This failure in coordination leads to a misallocation of resources in these markets, and lower production and investment. In such cases, ample theoretical literature suggests that information disclosure about firms’ actions could lead to improved coordination between firms, and thus to industrial growth (Vives, 1990; Pavan and Vives, 2015).

However, there is limited empirical evidence from nonexperimental settings regarding the effectiveness of information disclosure in solving coordination problems and fostering industrial growth. The causal effects of information disclosure remain elusive, mainly due to the endogeneity caused by firms mainly disclosing information voluntarily. Further, it is difficult to obtain an accurate notion of the information set of the relevant firms after the information disclosure. Related empirical literature uses experimental evidence to study voluntary disclosure by firms and its impact on firms’ decisions (Ackert et al., 2000; Abramson et al., 2005). Also, an extensive literature studies coordination failures in the development of different agro-industrial supply chains (Libecap and Lueck, 2011; Poray et al., 2002; Frank and Henderson, 1992; Purcell, 1973), but disregard the role of information disclosure as a solution to this problem.

This paper quantifies the effect of information disclosure on the production and
investment of firms in a supply chain, using evidence from California’s wine and grape industry. Firms in this industry face a coordination problem stemming from the multiple uses of raisin grapes and the difficulty winemakers and grape growers face in communicating, and thus anticipating, the production plans of other participants. To solve this coordination problem, between 1960 and 1968, an aerial survey of raisin vineyards measured and disclosed weekly information to all firms about the size and allocation of the harvest. The plan was to cover all raisin-growing areas. However, the survey was limited to two counties due to budget cuts, which provides the quasi-experimental variation needed to assess the effect of information disclosure on firms’ production and investment decisions.


I use a difference-in-differences model to compare production and investment in surveyed and non-surveyed areas to estimate the effect of information disclosure on these outcomes. The identification assumption is that in the absence of information disclosure, the production and investment of firms in surveyed and non-surveyed areas would have followed the same trend. As evidence for this assumption, I show
no evidence of differential pretrends between 1950 and 1959 or differences in the observed characteristics of grape-growing farms in 1950 or 1959. Additionally, a placebo test using fruits other than grapes shows that the surveyed counties did not increase their investment in these products after the survey.

The identification strategy relies on an additional assumption justified by the high transportation costs in the wine market: that markets are local. Given that all winemakers and grape growers received information on the raisin harvest, for the disclosure to primarily affect the decisions of neighboring firms, transactions mostly needed to happen locally. I justify this assumption based the high transportation costs of moving large volumes of grapes together with relatively low profit margins for grape growers. I estimate that the cost per mile to transport fresh grapes was approximately 1% of the profit of grape growers (Black, 1955; Transportation Division, 1964; Burns, 1974). The cost per mile was higher for grapes allocated to the production of table wines, as they were required to be crushed on the same day of harvest, and long-distance transportation needed special containers that protected grapes from getting crushed and fermented (Hale and Stokes, 1960). The relatively high transportation costs generated a strong preference for transactions with neighboring firms, which implies that firms near the surveyed areas would disproportionately benefit from the information disclosure.

I obtain three main results. First, information disclosure increases the production of wine, and mainly of higher quality table wines. The production of table wines more than doubled (108%) in surveyed areas, while the production of fortified wines rose by 18%. Second, in terms of the investment of upstream grape growers, there were increases in the acres of vineyards bearing wine grapes (24%) and raisins (18%). Third, downstream winemakers, in turn, rose their storage capacity by 13%.
Also, winemakers vertically integrated with the production of wine grapes—which act as complements to raisins in producing table wines—by increasing their vineyard acreage by 78%.

Finally, I provide empirical evidence suggesting that the primary mechanism driving increased production and investment is the enhanced coordination between growers and winemakers resulting from the information disclosure. I exploit the perfect complementarity in the production of table wines between raisins and wine grapes to construct two measures of coordination. The first is the correlation in the crush of raisins and wine grapes early in the harvest season, when the production of table wines happens, which is predicted to increase with improved coordination. The second is a measure of waste based on the excess crush of either raisin or wine grapes, which should decline with improved coordination. I estimate a difference-in-differences model that effectively shows that the correlation between crushed wine and raisin grapes increased, and the waste fell disproportionately in surveyed areas during the production of table wines.

This paper contributes to our understanding of the use of information disclosure in vertical supply chains by empirically examining its causal effect on the production and investment of firms in a non-experimental setting. The study of information disclosure is an important issue for anti-trust policy because information disclosure, mainly by trade associations, might increase efficiency in supply chains but could also facilitate collusion (Vives, 1990). This issue has been studied extensively in the theoretical literature (Pavan and Vives, 2015; Angeletos and Pavan, 2007; Rothschild, 1973), but empirical evidence is limited to experimental settings (Abramson et al., 2005; Ackert et al., 2000) and the results still ambiguous. This paper presents an instance in which information disclosure fosters efficiency and industrial growth, and
thus reinforces the current jurisprudence whereby information disclosure is not anti-
competitive per se.

Second, this paper is related to the empirical literature on possible solutions to
the frictions in vertical supply chains by showing the efficacy of information dis-
closures in achieving production and investment growth. There is ample evidence
on the use of vertical integration (Kopp and Sexton, 2021; Hennessy, 1996); con-
tingent contracts (Steiner, 2012; Goodhue et al., 2003; Bellemare and Lim, 2018);
and revenue-sharing agreements (Frank and Henderson, 1992; Poray et al., 2002) to
increase the allocation of resources in vertical supply chains. This paper provides ev-
idence of how information disclosure can alternatively help solve frictions in vertical
supply chains and promote industrial growth, which is especially relevant for policies
that promote the development of the agro-industrial sector of developing economies.

Finally, this paper contributes to our understanding of the supply factors that
underlie the growth of the agro-industrial sector and the wine industry in Califor-
nia. The paper shows the importance of coordination in the development of the
agro-industrial sector, and complements the ample evidence on the role of market
integration (Asher and Novosad, 2020); the diffusion and adoption of technology
(Griliches, 1957; Olmstead and Rhode, 2001); and human capital (Taylor and Mar-
tin, 2001; Huffman, 2001). For the wine industry in California, it sheds light on the
previously unexamined importance of the Raisin Lay Survey to the growth of the
industry during the 1960s (Alston et al., 2018; Pinney, 2009; Lukacs, 2000).
1.2 The California Grape and Wine Industry (1950-1968)

Grape growers and winemakers in the early California wine industry faced a coordination problem that fundamentally stemmed from the multiple uses of raisin grapes and the difficulty winemakers and grape growers encountered in anticipating the production plans of other participants in the vertical supply chain. Since market mechanisms proved ineffective, a survey that disclosed weekly information about the size and allocation of the raisin harvest to all participants addressed the coordination problem. The unexpected budget cuts that allowed only limited survey implementation generated quasi-experimental variation to measure the effect of information disclosure on production and investment.

In this section, I present the industrial context of the post-Prohibition grape and wine supply chain. Then, I explain the coordination problem between grape growers and winemakers and how market and government interventions failed to solve it. Finally, I describe how the Raisin Lay Survey used information disclosure on the size and allocation of raisin grapes to solve the coordination problem.

1.2.1 Institutional Context of the Grape and Wine Industry

In 1950, the wine industry was largely underdeveloped. The capital stock of winemakers was relatively small, and growth in the production of and investment in wine was low and mainly focused on the production of low-quality fortified wine. After a period of buoyancy in the first two decades of the 20th century, first Prohibition and then limitations on crushing raisins during World War II decimated the wine industry (Lukacs, 2000; Mendelson, 2009). This meant that in 1950, the fermenta-
tion and storage capacity of winemakers was only 17% and 23%, respectively, of that of 1919 (Pinney, 2009). Also, the wine production capacity of winemakers remained stagnant, as the production of table and fortified wines only grew by 0.4% and 1.2% yearly on average between 1950 and 1959. Moreover, in 1950 the production of fortified wines, which were lower quality and cheaper than table wines, represented 76% of the total production of wine.

The grape-wine supply chain in California between 1950 and 1968 was composed of upstream grape growers and downstream fruit packers, winemakers, and dried raisin packers (Figure 1.1). Grape growers produced three grape varieties categorized according to their use by downstream firms: table, wine, and raisin grapes. Table grapes were mainly sold to fruit packers and sold fresh. Downstream winemakers crushed wine grapes to produce table and fortified wines. Raisin grapes were more versatile than other varieties, because they could be dried and sold to raisin packers, crushed to produce table and fortified wines, or sold as fresh fruit.

The substitutability between raisins and wine grapes used in the production of wine changed with the type of wine. On the one hand, the production of table wines used raisins and wine grapes as perfect complements. In contrast to the contemporary varieties predominant in California, wine grapes in the 1950s had a strong flavor and a dark juice, which meant that mild-flavored crushed raisin grapes were needed to balance the wine and its alcoholic content. However, only raisins that were picked early in the harvest season\(^1\) had the appropriate sugar content to produce table wine. On the other hand, fortified wines used raisins and wine grapes as perfect

---
\(^1\)Early in the harvest season corresponds to the period between mid-August and the beginning of September; late refers to the period between mid-September and the end of October (Christensen and Peacock, 2000).
substitutes. Raisins that were picked late in the harvest season developed a large amount of sugar, and thus a high alcoholic content, which meant they could no longer be used to balance the intense flavors of the wine grapes. If raisins or wine grapes were unavailable, winemakers crushed more expensive and less productive table grapes to produce fortified wines.

1.2.2 The Coordination Problem

Raisin growers had problems coordinating with downstream firms on the efficient allocation of raisins between its multiple uses, which was particularly important for the production of table wines; during the 1950s, it was common to have surpluses on one use of raisin grapes together with shortages on the alternative uses (Lukacs, 2000). Raisin growers made their harvesting decisions before obtaining complete information on the demand they would obtain downstream for its multiple uses.
and the actual size of the grape harvest. Likewise, downstream firms made their production plans before knowing the availability of raisins and their allocation among multiple uses. The coordination was crucial for the production of table wines, because winemakers required raisins harvested early in the season, with lower sugar content, and had to crush them within hours of harvesting while securing the complementary amount of wine grapes.

The industry consisted of multiple grape-growing farms that were small relative to the production capacity of wineries, which reduced the ability of grape growers to coordinate with winemakers. In 1950, the average raisin-producing county had 209 farms that produced 102 tons of raisins each. Similarly, 153 farms produced wine grapes in these counties, yielding 67 tons of grapes. The average winemaker could produce 361.9 gallons of table wine, which required 32.6 million tons of crushed raisins and 29.0 million tons of wine grapes (Amerine, 1951; Amerine and Joslyn, 1970), which means that the average farm would only supply between 2.3% and 3.1% of the inputs needed by a single winery.

Traditional market mechanisms, such as prices and contingent contracts, proved ineffective in solving the coordination failures between grape growers and winemakers in the production of table wines. First, prices were determined after harvesting decisions were made, which limited flexibility to change quantities as a function of prices. Moreover, prices were volatile and past prices had limited bearing on the allocation decisions of grapes (Renaud, Bertrand, 1966). Also, farmers were subject to idiosyncratic (related to pre-harvest practices) and aggregate shocks (mostly weather or plagues) that made them unable to commit to contracts with other agents in the supply chain (Winkler et al., 1962). Further, high transportation costs limited risk-sharing with other farms.
In addition, the Raisin Advisory Committee (RAC, which is the federally approved trade association of raisin growers) had a Federal Marketing Order in place for dried raisins, which in practice worked as an option that limited the downside of the price. The design of the price floor responded to the market conditions of dried raisins by late September, which is the beginning of the late harvest season. If prices were above some threshold determined by the RAC every year, no intervention occurred, and growers could sell all of their harvest at the higher market price. If prices were below the threshold, the RAC would restrict the sale of part of the harvest to limit local supply and raise prices to the threshold level. The RAC would manage the unsold grapes, pay the growers the threshold price four to five months after harvest, and fund any deficits with federal funds. This price floor gave incentives to raisin growers to delay harvesting and drying to reduce supply and induce a higher price floor. Hence, fewer farmers were willing to harvest early, and thus reduced the availability of raisins to produce table wines (Nef, 1998; Petrucci and Clary, 2002).

1.2.3 Intervention: The Raisin Lay Survey

The RAC and the Wine Advisory Board (WAB, which is the federally approved trade association of California winemakers) designed the Raisin Lay Survey to get an accurate, timely estimate of how much of the raisin crop was dried and how much remained to be crushed into wine. The expectation was that this information would allow all winemakers, grape growers, and dried fruit packers to plan accordingly, thus improving the efficiency in allocating raisin grapes between its multiple uses. The main hope of the WAB was that this would increase the availability of raisin grapes to produce higher-quality table grapes. At the same time, the RAC expected to
reduce the size of federal funds used to subsidize the raisin Marketing Orders.

The initial plan of the Raisin Lay Survey was to aerially photograph the paper trays with grapes laid to dry next to the vineyards of all raisin-growing areas every week. These high-definition pictures were processed using computers to calculate the density of the laid grapes and get an accurate estimate of the dried crop. Since the acreage of vineyards bearing fruit at the end of the previous harvest season was known and distributed by the Crop Reporting Service to all market participants, the Raisin Lay Survey allowed inference of the available grapes to crush. The reported counts were at neighborhood level, which meant that the survey also provided information on the location of available grapes.

The RAC and WAB lobbied both the federal and state government to secure funding to implement the survey starting in 1960. However, the costs to cover a large area were steep, since the program entailed using spy planes with high-definition cameras and securing access to computing facilities to conduct the count for 7 weeks. Consequently, the trade associations could only access funding to implement the program in Fresno and Madera. The choice of this area responded to three main factors: (i) the relatively larger size of Fresno’s agricultural, and particularly viticultural activity; (ii) the proximity to the Fresno airport, which during the 1960s was predominantly used for military operations; and (iii) technical factors related to the optics of the cameras used to conduct the survey. The latter explains why Madera was part of the survey instead of neighboring counties with larger raisin production such as Kings, Tulare, or Merced.

The Survey was conducted for 7 weeks starting in August between 1960 and 1968 and widely distributed to industry participants. The novelty of spy planes
and “electronic brains” to process the images amid the Cold War helped increase the initial diffusion of the survey’s report. This success in distribution guaranteed that the information disclosed was widely received and used by winemakers, grape growers, and packers.

1.3 Data

I collect and digitize data from multiple historical sources on the production and investment of grape growers and winemakers to measure the causal effect of information disclosure on industrial growth. The data contain geographic and temporal information that allows me to measure the differential effects of information disclosure between surveyed and non-surveyed areas. Also, the time series starts 10 years before the intervention, which enables me to test for differential trends before the intervention.

1.3.1 Production of Wine

I collect and digitize data for table and fortified wine production per year at county level between 1950 and 1968. The production of wine is in thousands of gallons and only includes the production between July and December of each year. The primary source is the Marketing California Grapes, Raisins, and Wine Report (henceforth Marketing Report) published by the California Crop and Livestock Reporting Service and the National Agricultural Statistical Service of the United States Department of Agriculture. Historical records from this source have a gap between 1960 and 1963, which I fill by using the Economic Report of the Wine Institute (the trade association of California winemakers), which replicated the information from the
Marketing Report.\textsuperscript{2} I obtain data on 17 wine-producing counties during the 19 years of the sample, for a total of 323 county-year observations.

To complement the data on wine production, I also collect and digitize data on the grape crush of raisin, table, and wine grapes at county level between 1950 and 1965. The grape crush is in tons and indicates the use of inputs to produce wine. The data come from the Marketing Report, and I fill in the missing data using information from the Economic Report of the Wine Institute. The data include information on grape crush for 27 grape-growing counties over 16 years, for a total of 432 county-year observations. In contrast to wine production, the time series stopped in 1965 because, starting in 1966, reports for grape crush were at district level, which aggregates several counties. Also, I have information on grape crush for 10 more counties than for wine production because the data on wine production group counties with smaller production into an Others category.\textsuperscript{3}

The average county produced 2.60 million gallons (s.d. 3.31) of table wine and 5.60 million gallons (s.d. 9.25) of fortified wines. This corresponds to 23.85 million tons (s.d. 67.48) of crushed raisins, 20.88 million tons (s.d. 32.12) of crushed wine grapes, and 12.03 million tons (s.d. 29.61) of crushed table grapes.

\subsection*{1.3.2 Investment of Upstream Grape Growers}

I collect information on the capital stock of grape growers, which is proxied by the acreage of bearing vines of raisin, table and wine grapes and other fruit trees. Data

\textsuperscript{2}I cross-checked production data for 1959, 1964, and 1965—years in which both sources are available—and verified that the figures are identical.

\textsuperscript{3}These counties are Amador, Butte, Merced, Monterey, Placer, Riverside, San Benito, Solano, Ventura, and Yolo.
are available at county-year level between 1953 and 1969. The count of vines and
trees starts once they bear fruit, which means that vines will have a lag of 2 to 3
years between the time of planting and the time they bear fruit (Winkler et al., 1962).
The source for this data is the California Fruit and Nut Acreage Reports issued by
the California Crop and Livestock Reporting Service and the National Agricultural
Statistical Service of the United States Department of Agriculture. The dataset
contains information on 57 counties over 17 years, for a total of 969 county-year
observations.

The average county had a cultivated area of 4.06 thousand acres (s.d. 18.18) of
raisin vines, 1.41 thousand acres (s.d. 5.11) of table grapevines, and 2.20 thousand
acres (s.d. 4.71) of wine grapevines. Most of the counties reported growing grapes
of different varieties; 46 counties had vines with wine grapes and 43 had vines with
raisin and table grapes. However, production was concentrated in the North Coast,
Central Valley, and Southern California; these regions contained all counties with
more than 100 acres of grapevines of any variety.

1.3.3 Investment of Downstream Winemakers

I also collect and digitize information on the capital stock of winemakers measured
by their acreage of wine grapevines and their storage and fermentation capacity. The
data are available at the winery level between 1955 and 1968, with a gap in 1966. As
for the capital stock of grape growers, the vines are those bearing fruit and will be
counted only 2 or 3 years after being planted. Storage and fermenting capacity are
in gallons. The data source is the Wines & Vines Annual Directory, which annually
surveyed all winemakers, distributors, and suppliers of the wine industry in North
America. However, I limit the sample to bonded wineries in California. The wineries are uniquely identified by their permit number with the Bureau of Alcohol, Tobacco, Firearms, and Explosives and their address, which allows me to form a panel between 1955 and 1968.

The dataset contains 573 unique wineries for a total of 3,140 winery-year observations. The average winery has a wine grape vineyard of 391.5 acres (s.d. 3,881.9), a storage capacity of 2.04 million gallons (s.d. 6.90), and fermenting capacity of 401.15 thousand gallons (s.d. 1,489.68).

1.4 Empirical Design

The identification strategy relies on the Raisin Lay Survey’s being conducted exclusively in Fresno and Madera Counties due to budgetary and technical constraints (Appendix Figure 1.A.1). Implementation of the survey was rapid after being agreed upon in late 1959 between trade associations of grape growers and winemakers. The plan was to cover all raisin-growing counties in the Central Valley and Southern California. However, the state and the federal government only offered a limited budget, and thus the aerial Survey was conducted in only a limited zone. The areas surveyed in Fresno County were chosen based on its large agricultural production and proximity to the Fresno Air Terminal, which was primarily for military purposes and thus facilitated the use of spy planes. Meanwhile, Madera was chosen because it was on the route that would best cover Fresno’s vineyards, given the technical characteristics of the planes and the weight of the photographic equipment. This left neighboring counties with larger raisin production than Madera unsurveyed, such as Kings, Tulare, and Merced.
The identification strategy also relies on local markets for grapes. Given that all winemakers and grape growers received information on the raisin harvest, for disclosure to primarily affect the decisions of neighboring firms, transactions mostly needed to happen locally. The high transportation costs of moving large volumes of grapes, together with relatively low profit margins for grape growers, justify this assumption. Using information on transportation costs and profit margins, I estimate that the cost per mile to transport fresh grapes was approximately 1% of the profit of grape growers (Black, 1955; Transportation Division, 1964; Burns, 1974). The cost per mile was almost double for grapes allocated to the production of table wines, since they needed to be crushed on the same day of harvest, and long-distance transportation required special containers that protected grapes from getting crushed and fermented (Hale and Stokes, 1960). The high transportation costs meant that the potential market for grapes harvested in Fresno was small, since they generated a strong preference for transactions with neighboring firms (Appendix Figure 1.A.2). However, this does not exclude the possibility that neighboring counties would partially benefit from the information disclosure, which means that the estimated effects are a lower bound.

I estimate the causal effects of the information disclosure regarding the Raisin Lay Survey on wine production and investment grape growers with the following equation:

$$y_{ct} = \beta \times Surveyed_c \times Post_t + \delta_c + \delta_t + \varepsilon_{ct},$$

where the dependent variable, $y_{ct}$, is grape crush, wine production, and the stock of bearing grapevines in county $c$ at year $t$. County fixed effects, $\delta_c$, control for variation in outcomes across counties that are constant over time, and time fixed effects, $\delta_t$,
account for variation across years that are common across counties. \( \text{Surveyed}_c \) is an indicator variable that equals one if county \( c \) was surveyed, and \( \text{Post}_t \) is an indicator variable that equals 1 starting in 1960. The variable \( \varepsilon_{ct} \) is a random unobserved variable. Standard errors are block-bootstrapped at county level with 1,000 replications to account for possible spatial and serial correlation. \( \beta \) is the parameter of interest that measures the effect of the information disclosure on production and investment outcomes.

To further validate that production and upstream investment increased as a consequence of the information disclosure, I analyze the timing of the investment decisions using a event study design. I estimate the following equation:

\[
y_{ct} = \sum_{h=-5, h\neq-1}^{7} \beta_h \text{Tre} a_t c \times I(t - 1960 = h) + \delta_t + \delta_c + \varepsilon_{ct}, \\
\text{(1.2)}
\]

where subscript \( c \) denotes a county and \( t \) a calendar year. The outcome, \( y_{ct} \), is grape crush, wine production, and the stock of bearing grapevines in county \( c \) at year \( t \). \( \text{Tre} a_t c \) is an indicator variable that is one if county was subject to the Raisin Lay Survey, and \( I(t - 1960 = h) \) is a binary variable that indicates the time period. The variable \( \varepsilon_{ct} \) is a random unobserved variable. Standard errors are block-bootstrapped at county level with 1,000 replications to account for possible spatial and serial correlation. \( \beta_h \) is the parameter of interest and measure the effect of the information disclosure on wine production and on the capital stock \( h \) periods before or after the survey.

For the investment decisions of vintners data is available at the ZIP code, so I
Figure 1.2: Map of California Wine Districts
estimate the following equation:

\[ y_{zt} = \beta \times Surveyed_z \times Post_t + \delta_z + \delta_t + \varepsilon_{zt}, \]  

(1.3)

where the dependent variable, \( y_{zt} \), is the acreage of vineyards owned by wineries and the storage and fermenting capacities in area with ZIP code \( z \) at year \( t \). ZIP code fixed effects, \( \delta_z \), control for variation in outcomes across ZIP codes that are constant over time. \( Surveyed_z \) is an indicator variable that equals one if ZIP code area \( z \) was surveyed. The variable \( \varepsilon_{zt} \) is a random unobserved variable. All other variables are defined as in equation 1.1. Standard errors are block-bootstrapped at county level with 1,000 replications to account for possible spatial and serial correlation. \( \beta \) is the parameter of interest that measures the effect of the information disclosure on investment outcomes. The baseline specification is estimated in the sample that includes all raisin-growing counties (Figure 1.2).

The identifying assumption is that production and investment decisions in surveyed and comparison areas would have been on the same trend in the absence of the Raisin Lay Survey. The rest of the section provides evidence in support of the empirical design and the identifying assumption.

1.4.1 Balancing Tests

I test whether the observable characteristics of vineyards for grape varieties were statistically indistinguishable between surveyed and non-surveyed counties before the Raisin Lay Survey. Using the cross-section for 1950 and 1959 for farm characteristics of the Census of Agriculture, I calculate the conditional average for surveyed and non-surveyed counties for each year and test whether the difference between
the means is statistically significant. Results for of tests show that the number and size of vineyards and the volume of harvests for all grape varieties between treated and comparison counties were statistically indistinguishable at the beginning of the study period in 1950 (Table 1.A.1, column 3) and in the year immediately before the survey in 1959 (Appendix Table 1.A.1, column 6). Despite the differences not being statistically significant, the mean characteristics are systematically larger for surveyed counties, consistent with Fresno being the county with the largest agricultural production in California. To address this issue, I will run robustness tests of the main specifications using only Madera as the treatment unit, since its characteristics suggest that it is similar to the average county in the control group (Appendix Table 1.A.2).

1.4.2 Parallel Trends Before the Raisin Lay Survey (1950-1959)

I use data from 1950 to 1959, before the Raisin Lay Survey, to test whether there are differential time trends in outcomes between surveyed and non-surveyed counties. I estimate a model with a linear time trend interacted with an indicator for counties surveyed for outcomes related to wine production and grape growers’ investment and for ZIP codes in surveyed areas for wineries’ investments. The estimated coefficient on the surveyed counties indicator is not statistically different from zero in all specifications, which confirms the results from the balancing tests presented in Appendix Table 1.A.1. This suggests that there are no statistical differences between surveyed and non-surveyed counties. Estimating a separate time trend for surveyed areas confirms that there are no statistically significant differences (Appendix Tables 1.A.3, 1.A.4, 1.A.5, and 1.A.6).
1.4.3 Placebo Test: Investment in Other Fruit Trees

Finally, a placebo test using investment in fruits other than grapes suggests that no county-wide agricultural policies in surveyed counties act as confounders of the effects of the information disclosure regarding the grape and wine markets. A possible concern is that the surveyed counties, Fresno and Madera, were subject to a policy or intervention that benefited their agriculture disproportionately relative to other counties. To address this issue, I estimate equation 1.1 using the acreage harvested of bearing trees of fruits other than grapes. The results show that the investment in deciduous fruits, citrus, and other fruits in surveyed counties shows no statistically significant differences before and after 1960 (Appendix Table 1.A.7).

1.5 Effects of the Raisin Lay Survey on Production and Investment

Information disclosure increased the production of wine—mainly of higher quality table wines—as well as the investment of both upstream grape growers and downstream winemakers.

1.5.1 Wine Production

Counties surveyed by the Raisin Lay Survey disproportionately increased their production of wine relative to non-surveyed counties, and the effect was larger for table wines. I estimate equation 1.1 using the production of table and fortified wines as the outcome of interest. Results of the estimation show that the information disclo-
Table 1.1: Effects of the Raisin Lay Survey on Wine Production

<table>
<thead>
<tr>
<th></th>
<th>Wine Production (Million Gallons)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Table</td>
<td>Fortified</td>
</tr>
<tr>
<td>Surveyed x Post</td>
<td>5.0708***</td>
<td>2.5754***</td>
<td>2.7606**</td>
</tr>
<tr>
<td></td>
<td>(1.4529)</td>
<td>(0.7899)</td>
<td>(1.2657)</td>
</tr>
<tr>
<td>Observations</td>
<td>253</td>
<td>253</td>
<td>253</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.967</td>
<td>0.821</td>
<td>0.971</td>
</tr>
<tr>
<td>Mean Surveyed: 1950-59</td>
<td>23.07</td>
<td>2.11</td>
<td>20.96</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Wine Industry Economic Report. OLS estimation of equation 1.1 using total production and the production of table and fortified wines as an outcome. Standard errors in parentheses are block-bootstrapped at county level with 1,000 replications.

sure increased the production of wine in surveyed counties by 5.1 million gallons per year, which is 19% more than the 23.1 million gallons that were produced yearly on average before the Survey (Table 1.1). The information disclosure mostly favored the production of table wines, which more than doubled (118%) from the average production before the survey. Meanwhile, the production of fortified wines rose 18% from the average production before the survey (Table 1.1).

Results for the grape crush for wine production in surveyed counties confirm the increased production due to the information disclosure and suggest a more efficient use of grapes as inputs. Estimating equation 1.1 with grape crush as the outcome shows that surveyed counties experienced an increase in the grape crush of 75.5 million tons per year, which is 29% more than the 267.0 million tons per year before the Raisin Lay Survey. This additional grape crush is consistent with the observed increase in wine production (Appendix Table 1.A.8). This change occurs due to
an increase in the crush of both raisin (47%) and wine grapes (13%). Also, the estimation suggests that there is a substitution away from the crush of fresh grapes (-19%), which were a less efficient input in the production of wine and mainly used when other grape varieties were unavailable (Amerine, 1951).

The results are robust to changing the sample of non-surveyed counties included in the control and treatment groups. This exercise addresses the concern that the results might not respond to the information disclosure and are instead due to other shocks that affected the production of wine in distant counties. The estimation with the different samples shows a similar positive effect, which is statistically and economically significant for both wine production (Appendix Table 1.A.9, columns 1 to 5) and grape crush (Appendix Table 1.A.10, columns 1 to 5). Also, I estimate the baseline specification excluding Fresno from the treatment counties to address the concern that Fresno’s agricultural sector was larger than all other counties in California and might not be fully comparable with the control group. The results show that the increase in production and grape crush due to the information disclosure is similar to the rise seen in the main specification, which suggests that the effects are not due to the size of Fresno’s agricultural sector (Appendix Tables 1.A.9 and 1.A.10, column 6).

Finally, the results of the event study are also consistent with the results in Table 1.1, whereby wine production increased due to the information disclosure. I estimate equation 1.2 using total, table, and fortified wine production as outcomes. The estimates for table wine show that production increased 162% 1 year after the Survey and 171% after 5 years (Figure 1.3, Panel B). The results for fortified wines yield estimates that are not statistically different from zero for the first 3 years of the program ($h = 0$, $h = 1$, and $h = 2$). Then, the production of fortified wine increases
Figure 1.3: Effects of the Raisin Lay Survey on Wine Production

Panel A: Total Production

Panel B: Table Wines

Panel C: Fortified Wines

Notes: Data from the Marketing Report and the Economic Report of the Wine Industry. Red dots are the OLS estimates of equation 1.2 using total production of wine and of table and fortified wine as outcomes. Blue lines are the 95% confidence intervals of the OLS estimates. Standard errors are block-bootstrapped at county level with 1,000 replications. Data are at county level.
18% after 3 years and 30% after 5 years (Figure 1.3, Panel C). The timing suggests that the rise in the production of fortified wines mainly responds to increases in the investment described in Subsection 1.5.2, since it takes from 2 to 3 years for vines to start bearing fruit after planting.

1.5.2 Investment of Upstream Grape Growers

Upstream grape growers in surveyed counties disproportionately increased the investment in grapevines used in wine production. Estimates for equation 1.1 using bearing vines as the outcome show that the information disclosure increased the capital stock of grape growers in surveyed counties by 12.7 thousand acres, which is a 15% increase from the average before the Raisin Lay Survey. Decomposition by type of grape shows that the acreage of raisin grapes rose by 18% and of wine grapes by 22%, while the area of table grapes fell by 16% (Table 1.2). The results are consistent with results for the grape crush (Appendix Table 1.A.8), in which the varieties involved in the production of wine (i.e., wine and raisin grapes) substituted less productive table grapes. These results are robust to alternative samples for both the control and treatment groups (Appendix Table 1.A.11).

Event study results are also consistent with the hypothesis that grape growers increased their investment due to the information disclosure. Estimates for all grapes show that bearing vines increased 9% 2 years after the survey and 20% after 7 years (Figure 1.4, Panel A). The first 2 years of the survey ($h = 0$ and $h = 1$) are not statistically different from zero, which is consistent with the time it takes between planting the vines and starting to bear fruit. Also, although there is evidence of a slightly ascending pre-trend, all estimates of the parameters lagging the survey are
Table 1.2: Effects of the Raisin Lay Survey on Grape Growers’ Investment

<table>
<thead>
<tr>
<th>Surveyed x Post</th>
<th>Bearing Trees (Thousand Acres)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) Total</td>
</tr>
<tr>
<td>Observations</td>
<td>969</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.995</td>
</tr>
<tr>
<td>Mean Surveyed: 1950-59</td>
<td>82.60</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the California Fruits and Nut Acreage Report and the Economic Report of the Wine Industry. OLS estimation of equation 1.1 uses total production and of table and fortified wines as an outcome. Standard errors in parentheses are block-bootstrapped at county level with 1,000 replications.

not statistically different from zero. By grape variety, the pattern is similar to the results of the difference-in-differences model: Vines of raisins and wine grapes rose by 11% and 6% after 2 years, respectively, and by 22% and 39% after 7 years (Figure 1.4, Panels B and C). Meanwhile, table grapes show negative point estimates after the survey, but are not statistically different from zero (Figure 1.4, Panel D).

Finally, the placebo tests further validate that the effects of information disclosure on the investment of grape growers are not spurious. To do so, I drop the surveyed counties of Fresno and Madera and randomly assign the treatment to two counties. Then, I estimate equation 1.1 and save the point estimates. I repeat this process for 11,000 iterations. Distributions of the estimates for vines of all grapes and by grape variety have means close to zero. Furthermore, I present values for the point estimates for the main specification and those that exclude Fresno from the treatment
Figure 1.4: Effects of the Raisin Lay Survey on Grape Growers’ Vineyard Acreage by Type of Grape

Panel A: Total

Panel B: Raisins

Panel C: Wine

Panel D: Table

Notes: Data from the California Fruits and Nut Acreage Report and the Economic Report of the Wine Industry. Red dots are the OLS estimates of equation 1.2 using total grape acreage and raisin, wine, and table grape acreage as an outcome. Blue lines are the 95% confidence intervals of the OLS estimates. The gray area represents the expected time for vines to start bearing fruit if they are planted in year 0 (Winkler et al., 1962). Standard errors are block-bootstrapped at county level with 1,000 replications. Data are at county level.
group. Results show that for vines of all grapes and wine and raisins are outside the 5th and 95th percentiles of the distribution of placebo estimates (Appendix Figure 1.A.4).

1.5.3 Investment of Downstream Winemakers

Like upstream grape growers, downstream winemakers in surveyed counties disproportionately increased their investment on wine grapevines and storage capacity after the information disclosure. I estimate equation 1.3 using as outcomes the acreage of wine grapevines owned by winemakers and their storage and fermenting capacity. Results show that winemakers in surveyed areas increased their acreage of wine grapes by 78%, which suggests that the enhanced flow of information on the volume of raisins gave incentives to winemakers to vertically integrate the production of wine grapes (Table 1.3). In addition, winemakers in surveyed areas increased their storage capacity by 13%. They did not alter their fermentation capacity; together with the results on production, this suggests that before the survey, winemakers had excess production capacity (Table 1.3).

These results are robust to restricting the sample of wineries to a balanced panel. Equation 1.3 is estimated using only wineries that existed between 1955 and 1968 and that report having vineyards, storage capacity, or fermenting capacity for each year of the sample. Estimation of the area of wine vines includes 21 winemakers; storage capacity includes 69 winemakers; and fermenting capacity includes 56 winemakers. This estimation shows that for winemakers in surveyed areas, the acreage of wine grapes rose 92% and storage capacity increased by 16%. Fermenting capacity was unaffected (Appendix Table 1.A.12).
Table 1.3: Effects of the Raisin Lay Survey on Winemakers’ Investment

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Vineyard Storage Cap. Ferm. Cap.</td>
<td>(Log Acres)</td>
<td>(Log 1000 Gals)</td>
</tr>
<tr>
<td>Surveyed x Post</td>
<td>0.5754***</td>
<td>0.1182**</td>
<td>-0.007649</td>
</tr>
<tr>
<td></td>
<td>(0.1198)</td>
<td>(0.05258)</td>
<td>(0.02700)</td>
</tr>
<tr>
<td>Observations</td>
<td>601</td>
<td>1,358</td>
<td>1,057</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.985</td>
<td>0.987</td>
<td>0.991</td>
</tr>
<tr>
<td>Mean Surveyed: 1955-59</td>
<td>1426.63</td>
<td>5946.58</td>
<td>1506.89</td>
</tr>
<tr>
<td>Winery FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Wines & Vines Directory and Annual Buyer’s Guide. OLS estimation of equation 1.3 using total wine grape vineyard acreage and storage and fermentation capacity as outcomes. The 77.8% increase in vineyard acreage is calculated by $e^{0.5754} - 1$. Standard errors in parentheses are block-bootstrapped at county level with 1,000 replications.
Finally, I test whether the information disclosure had spillovers to wineries adjacent to the surveyed areas. To do so, I estimate the following equation:

\[ y_{izt} = \beta_1 \text{Distance}_{iz} \times \text{Post}_t + \beta_2 \text{Post}_t + \delta_i + \delta_c + \delta_t + \varepsilon_{izt}, \]  

where subscript \( i \) refers to a winery, \( z \) to a ZIP code, \( t \) to a calendar year, and \( c \) to county. \( y_{izt} \) refers to the acreage of wine grapes owned by winery \( i \) or their storage and fermenting capacity. \( \text{Distance}_{iz} \) is the distance in miles of winery \( i \) to the closest ZIP code subject to the Raisin Lay Survey, and takes the value of zero if winery \( i \) is in a ZIP code subject to the Raisin Lay Survey. \( \text{Post}_t \) is a dummy variable of value 1 starting in 1960. \( \delta_i, \delta_c, \) and \( \delta_t \) are fixed effects for establishment, county, and year, while \( \varepsilon_{izt} \) is a random unobserved variable. Standard errors are block-bootstrapped at county level with 1,000 replications to account for possible spatial and serial correlation. \( \beta \) is the parameter of interest and measures how the effect of the information disclosure on the outcome of interest changes with every mile of distance from the surveyed area.

The results of this exercise show that the effects of the Raisin Lay Survey were stronger in surveyed areas, but neighboring areas still benefited from the information disclosure, albeit at a rate that decreases with distance. The estimates suggest that each mile reduces the additional investment in wine grape vineyards and the storage capacity of winemakers by 0.07% (Appendix Table 1.A.13). If the effects are linear, the positive effects of the Raisin Lay Survey would have only impacted the Central Valley. The effects for both vineyard acreage and storage would disappear entirely north of Sacramento in the north and close to Los Angeles in the south (Figure 1.5). These results are consistent with the rough estimates of transportation costs.
reported in Section 1.4, since transporting grapes for 50 miles to produce table wines would deplete the grape growers’ profits.
Figure 1.5: Spillover Effects of the Raisin Lay Survey on Winemakers’ Investment

Panel A: Wine Grape Vineyard Acreage

Panel B: Storage Capacity

Notes: Data from the Wines & Vines Directory and Annual Buyer’s Guide. Predicted effects on wine grape vineyard acreage and the storage capacity of winemakers in the Raisin Lay Survey using OLS estimates of equation 1.4 and reported in Appendix Table 1.A.13. Data are at winery level.
1.6 Mechanism: Increased Coordination between Firms

I hypothesize that the primary mechanism behind the observed increases in wine production and investment after the information disclosure is the enhanced coordination between firms, especially in the production of table wines. By observing the allocation decisions of raisin growers beforehand, winemakers and wine grape growers could match them. In turn, raisin growers were more likely to harvest early in the season, since they were confident that wine grape growers and winemakers would match their decision. This improvement in coordination means that firms will have greater incentives to invest, because the reduction in the strategic uncertainty faced by all firms raises the expected return of the investment.

To test this hypothesis, I use data on the weekly crush of raisin and wine grapes from the Marketing Report at district level, which is a higher level of aggregation than a country. There are three wine districts in California: North Coast, Central Valley, and Southern California. The Central Valley district includes the surveyed counties of Fresno and Madera and will be the treated unit, while the other two districts will serve as controls (Appendix Figure 1.A.5). Ideally, the data would be at county level; however, it is not available at county level given the higher frequency of the time series.

I build two measures of coordination that rely on the perfect complementarity of raisin and wine grapes in the production of table wines to test for enhanced coordination. The first is the correlation for each district and year between raisin and grape crush for the early harvest season.\(^4\) The rationale behind this measure

\(^4\)Weeks 30 to 36 of the year, which are usually between late July and early September.
is that in the absence of coordination failures, this should be close to one. Thus, if the Raisin Lay Survey increased coordination, we expect that the correlation would disproportionately increase in the Central Valley district.

The second measure is an estimate of waste due to a mismatch between the crush of raisin and wine grapes, as follows:

\[
\text{waste}_{dt} = \max\{\text{raisin}_{dt}, 1.125 \times \text{wine}_{dt}\} - \min\{\text{raisin}_{dt}, 1.125 \times \text{wine}_{dt}\},
\]

where subindex \(d\) refers to a district and \(t\) is a year. \text{raisin}_{dt} and \text{wine}_{dt} are the raisin and wine crush, respectively. Coefficients of the Leontief production function correspond to the proportions recommended in Amerine (1951) and Amerine and Joslyn (1970) for the production of table wines. The expression within the max operator indicates the input for which there is excess availability.

I estimate the effects of the information disclosure from the Raisin Lay Survey on the proxies for coordination using only the weeks that correspond to the early harvest season, when table wines would be produced, with the following equation:

\[
y_{dt} = \beta \times \text{Surveyed}_d \times \text{Post}_t + \Gamma X_{dt} + \delta_d + \delta_t + \varepsilon_{dt},
\]

where the dependent variable, \(y_{dt}\), are the correlation and waste measures for district \(d\) at year \(t\) defined above. District fixed effects, \(\delta_d\), control for variation in outcomes across counties that are constant over time, and time fixed effects, \(\delta_t\), account for variation across years that are common across counties. \(X_{dt}\) is the total yearly grape crush in the district to control for the size of the harvest. \text{Surveyed}_d\) is an indicator variable that equals 1 for the Central Valley district, and \text{Post}_t\) is another indicator
Table 1.4: Effects of the Raisin Lay Survey on the Proxy Measures of Coordination between Winemakers and Grape Growers

<table>
<thead>
<tr>
<th>Measures of Coordination</th>
<th>(1) Correlation</th>
<th>(2) Waste</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveyed x Post</td>
<td>0.1031**</td>
<td>-1,482.3***</td>
</tr>
<tr>
<td></td>
<td>(0.04160)</td>
<td>(43.420)</td>
</tr>
</tbody>
</table>

Observations: 112 112
R-squared: 0.853 0.968
Mean Surveyed: 1950-59: 0.8333 4,169.73
County FE: YES YES
Year FE: YES YES

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Marketing California Grapes, Raisins, and Wine Report. OLS estimation of equation 1.6 using total grape crush as the outcome. The estimation in column (1) uses the correlation between wine and raisin grape crush during late July and the first week of September as the outcome. The estimation in column (2) uses the estimate of waste described in equation 1.5 as outcome. Standard errors in parentheses are clustered at district level.

The results of these estimates suggest that the Raisin Lay Survey indeed improved coordination between winemakers, raisin growers, and wine grape growers. The correlation between grape and raisin crush disproportionately increased by 0.10 points in surveyed counties from a pre-survey average of 0.83 (Table 1.4, Column 1). After the survey, the average correlation is close to 1, as predicted by the theory for perfect complements. The estimate for waste also implies an increase in coordination, because the excess grape crush during the production of table wines fell by
35.5% (Table 1.4, Column 2).

1.7 Conclusion

I estimate the effects of the information disclosure of firms’ allocation decisions on the production and investment decisions of firms in a supply chain. Using evidence from the grape and wine industry in California between 1950 and 1968, I exploit the quasi-experimental variation arising from budget cuts in implementing the Raisin Lay Survey, which limited the surveyed areas to only the counties of Fresno and Madera. The survey reported weekly information about the quantity of dried raisins, which allowed inference of the harvest size and allocation decisions of raisin growers. Although the report was distributed to all participants in the industry, it disproportionately benefited firms in surveyed areas since high transportation costs meant that trading was mainly local. By comparing surveyed and non-surveyed areas, I observe that grape crush and wine production rose disproportionately in surveyed areas, and especially favored the production of higher quality table wines. In addition, winemakers and wine grape and raisin growers in surveyed counties disproportionately increased their investment due to the information disclosure, which accelerated the development of the Californian wine industry. Finally, leveraging the perfect complementarity of wine and raisin grapes in the production of table wines, I provide evidence that the main channel driving the increased production and investment was an improvement in the coordination of firms in the grape-wine supply chain.

The results of this paper might inform agricultural and industrial policies in developing economies. The wine industry in California in 1950 was in its infancy after Prohibition, and restrictions during World War II had decimated firms’ capital and
industrial knowledge, which is similar to the state of several agro-industrial supply chains in developing economies. Also, industries in developing economies face information and communication frictions similar to those of the early Californian grape and wine industry, which limits the coordination and contributes to the underdevelopment of these industries. Furthermore, the technological advances in aerial and satellite imagery render these programs easier for implement than alternative solutions to coordination failures in supply chains. Contingent contracts or vertical integration usually require conditions not present in developing economies, such as a strong judiciary and institutional setting, or larger financial investments by farmers or manufacturers, which renders their use difficult.
1.A  Appendix: Additional Figures and Tables

Figure 1.A.1: Map of the Fresno and Madera Areas Photographed in the Raisin Lay Survey

Notes: Available in the archives of the Wine Institute held at the Special Collections of the Shields Library at the University of California, Davis. U.S. Route 99, the San Joaquin river, and the Kings river are used as anchors to georeference the location of the surveyed areas. Date: April 1960.
Figure 1.A.2: Profit per ton of raisins shipping from Fresno to different locations in California

Panel A: For Fortified Wines

Grape Growers’ Profits (in $ per ton)
-93 to -37
-37 to -16
-16 to -3
-3 to 7
7 to 17
17 to 30
30 to 40

Panel B: For Table Wines

Grape Grower’s Profits (in $ per ton)
-208 to -108
-108 to -67
-67 to -30
-30 to -6
-6 to 6
6 to 17
17 to 24
24 to 30
30 to 40

Notes: Data from Black (1955); Hale and Stokes (1960); Transportation Division (1964); and Burns (1974). The cost in truck of raisins intended to produce fortified wines per mile and ton was between $0.28 and $0.42 in 1956 and between $0.31 and $0.48 in 1963. The cost of transportation for raisins for table wines was 60% higher than for fortified wines. The average profit per ton before shipping was $35 in 1956 and $40 in 1963. The map reflects data for 1963 using the lower bound of the transportation cost.
Figure 1.A.3: Map of California Regions

- **Surveyed Counties (part of San Joaquin Valley)**
- **San Joaquin Valley**
- **Central Valley (excl. San Joaquin Valley)**
- **Southern California**
Figure 1.A.4: Histogram of Placebo Coefficients for the Effects of the Raisin Lay Survey on Grape Growers’ Vineyard Acreage by Type of Grape

Panel A: Total  Panel B: Raisins

Panel C: Wine  Panel D: Table

Notes: Data from the California Fruits and Nut Acreage Report and the Economic Report of the Wine Industry. The solid black line is the point estimate of the effect of the information disclosure on grape growers’ investment, as reported in Table 1.2. The solid red line is the point estimate of the effect of the information disclosure on grape growers’ investment, which only includes Madera in the treatment group. The light gray histograms are the placebo point estimates of 11,000 simulations in which two counties are randomly assigned as treatment and used to estimate equation 1.1 in a sample that excludes Fresno and Madera Counties. The dotted lines are the 95% confidence intervals of the placebo point coefficients. Data are at county level.
Figure 1.A.5: Map of the California Wine Districts
<table>
<thead>
<tr>
<th></th>
<th>1950</th>
<th>1959</th>
<th></th>
<th>1950</th>
<th>1959</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Surv (1)</td>
<td>Non-surv (2)</td>
<td>Diff. (3)</td>
<td>Surv (4)</td>
<td>Non-surv (5)</td>
<td>Diff. (6)</td>
</tr>
<tr>
<td><strong>Raisin Grapes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Farms</td>
<td>2,967</td>
<td>215</td>
<td>2,751</td>
<td>2,529</td>
<td>136</td>
<td>2,393</td>
</tr>
<tr>
<td></td>
<td>(3,536)</td>
<td>(338)</td>
<td>(2,501)</td>
<td>(2,949)</td>
<td>(246)</td>
<td>(2,086)</td>
</tr>
<tr>
<td>Farm area (acres)</td>
<td>73,774</td>
<td>2,573</td>
<td>71,201</td>
<td>80,675</td>
<td>2,430</td>
<td>78,245</td>
</tr>
<tr>
<td></td>
<td>(85,673)</td>
<td>(5,786)</td>
<td>(60,590)</td>
<td>(85,107)</td>
<td>(5,897)</td>
<td>(60,191)</td>
</tr>
<tr>
<td>Production (tons)</td>
<td>443,314</td>
<td>12,934</td>
<td>430,380</td>
<td>557,695</td>
<td>15,342</td>
<td>542,352</td>
</tr>
<tr>
<td></td>
<td>(510,006)</td>
<td>(31,010)</td>
<td>(360,678)</td>
<td>(589,591)</td>
<td>(38,091)</td>
<td>(416,968)</td>
</tr>
<tr>
<td><strong>Wine Grapes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Farms</td>
<td>567</td>
<td>172</td>
<td>395</td>
<td>394</td>
<td>103</td>
<td>290</td>
</tr>
<tr>
<td></td>
<td>(518)</td>
<td>(258)</td>
<td>(369)</td>
<td>(392)</td>
<td>(191)</td>
<td>(279)</td>
</tr>
<tr>
<td>Farm area (acres)</td>
<td>11,080</td>
<td>3,757</td>
<td>7,322</td>
<td>8,640</td>
<td>2,985</td>
<td>5,654</td>
</tr>
<tr>
<td></td>
<td>(9,852)</td>
<td>(8,324)</td>
<td>(7,148)</td>
<td>(6,564)</td>
<td>(6,349)</td>
<td>(4,800)</td>
</tr>
<tr>
<td>Production (tons)</td>
<td>50,760</td>
<td>14,911</td>
<td>35,848</td>
<td>63,200</td>
<td>15,622</td>
<td>47,577</td>
</tr>
<tr>
<td></td>
<td>(43,234)</td>
<td>(32,267)</td>
<td>(31,195)</td>
<td>(46,717)</td>
<td>(31,587)</td>
<td>(33,589)</td>
</tr>
<tr>
<td><strong>Table Grapes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Farms</td>
<td>700</td>
<td>298</td>
<td>401</td>
<td>476</td>
<td>143</td>
<td>332</td>
</tr>
<tr>
<td></td>
<td>(791)</td>
<td>(295)</td>
<td>(562)</td>
<td>(612)</td>
<td>(210)</td>
<td>(434)</td>
</tr>
<tr>
<td>Farm area (acres)</td>
<td>9,826</td>
<td>2,855</td>
<td>6,970</td>
<td>8,314</td>
<td>3,253</td>
<td>5,060</td>
</tr>
<tr>
<td></td>
<td>(12,953)</td>
<td>(7,336)</td>
<td>(9,267)</td>
<td>(10,562)</td>
<td>(8,388)</td>
<td>(7,641)</td>
</tr>
<tr>
<td>Production (tons)</td>
<td>52,349</td>
<td>16,235</td>
<td>36,113</td>
<td>63,200</td>
<td>15,622</td>
<td>47,577</td>
</tr>
<tr>
<td></td>
<td>(68,711)</td>
<td>(44,662)</td>
<td>(49,340)</td>
<td>(46,717)</td>
<td>(31,587)</td>
<td>(33,589)</td>
</tr>
</tbody>
</table>

**Notes:** ***p<0.01, **p<0.05, *p<0.1. Data are from the 1950 and 1959 Census of Agriculture. Balancing tests for 2 counties that were aerially surveyed in the Raisin Lay Survey and 27 counties that were not surveyed. Counties in the Central Valley and Southern California that produce raisins are included. Columns 1-2 and 4-5 present the mean and standard deviation (in parentheses) of characteristics of vineyards in surveyed and non-surveyed counties at the start of the study period (1950) and the year immediately before the Raisin Lay Survey (1959). Columns 3 and 6 present the difference and the standard error (in parentheses) of the means difference t-test between surveyed and non-surveyed counties.
Table 1.A.2: Verifying Balance in Vineyard Characteristics in Madera and Non-surveyed Counties

<table>
<thead>
<tr>
<th></th>
<th>1950</th>
<th></th>
<th>1959</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Madera (1)</td>
<td>Non-surveyed (2)</td>
<td>Diff. (3)</td>
<td>Madera (4)</td>
</tr>
<tr>
<td><strong>Raisin Grapes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Farms</td>
<td>467</td>
<td>215.52</td>
<td>251.48</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td>(338.37)</td>
<td>(338.37)</td>
<td>(245.26)</td>
<td>(245.26)</td>
</tr>
<tr>
<td>Farm area (acres)</td>
<td>7,194</td>
<td>2,572.59</td>
<td>4,618.41</td>
<td>4,049</td>
</tr>
<tr>
<td></td>
<td>(5,786.47)</td>
<td>(5,786.47)</td>
<td>(5,896.54)</td>
<td>(5,896.54)</td>
</tr>
<tr>
<td>Production (tons)</td>
<td>22,686</td>
<td>12,934.26</td>
<td>9,751.74</td>
<td>26,079</td>
</tr>
<tr>
<td></td>
<td>(31,009.99)</td>
<td>(31,009.99)</td>
<td>(38,091.35)</td>
<td>(38,091.35)</td>
</tr>
<tr>
<td><strong>Wine Grapes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Farms</td>
<td>201</td>
<td>172.44</td>
<td>28.56</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>(258.02)</td>
<td>(258.02)</td>
<td>(191.33)</td>
<td>(191.33)</td>
</tr>
<tr>
<td>Farm area (acres)</td>
<td>4,114</td>
<td>3,757.63</td>
<td>356.37</td>
<td>3,998</td>
</tr>
<tr>
<td></td>
<td>(8,324.15)</td>
<td>(8,324.15)</td>
<td>(6,349.58)</td>
<td>(6,349.58)</td>
</tr>
<tr>
<td>Production (tons)</td>
<td>20,189</td>
<td>14,911.81</td>
<td>5,277.19</td>
<td>30,116</td>
</tr>
<tr>
<td></td>
<td>(32,267.05)</td>
<td>(32,267.05)</td>
<td>(31,587.31)</td>
<td>(31,587.31)</td>
</tr>
<tr>
<td><strong>Table Grapes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Farms</td>
<td>141</td>
<td>298.52</td>
<td>-157.52</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>(295.48)</td>
<td>(295.48)</td>
<td>(210.82)</td>
<td>(210.82)</td>
</tr>
<tr>
<td>Farm area (acres)</td>
<td>667</td>
<td>2,855.52</td>
<td>-2,188.52</td>
<td>845</td>
</tr>
<tr>
<td></td>
<td>(7,336.50)</td>
<td>(7,336.50)</td>
<td>(8,388.18)</td>
<td>(8,388.18)</td>
</tr>
<tr>
<td>Production (tons)</td>
<td>3,763</td>
<td>16,235.96</td>
<td>-12,472.96</td>
<td>6,830</td>
</tr>
<tr>
<td></td>
<td>(44,662.77)</td>
<td>(44,662.77)</td>
<td>(31,587.31)</td>
<td>(31,587.31)</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data are from the 1950 and 1959 Census of Agriculture. Balancing tests for 2 counties that were aerially surveyed in the Raisin Lay Survey and 27 counties that were not surveyed. Counties in the Central Valley and Southern California that produce raisins are included. Columns 1 and 4 present the values for Madera County. Columns 2 and 5 present the mean and standard deviation (in parentheses) of characteristics of vineyards in non-surveyed counties at the start of the study period (1950) and the year immediately before the Raisin Lay Survey (1959). Columns 3 and 6 present the difference and the standard error (in parentheses) of the means difference t-test between surveyed and non-surveyed counties.
Table 1.A.3: Time Trends for Wine Production by Type before the Raisin Lay Survey for Surveyed and Non-Surveyed Counties, 1950-1959

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Table</td>
<td>Fortified</td>
</tr>
<tr>
<td>Trend</td>
<td>0.056</td>
<td>0.038**</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.017)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>Surveyed</td>
<td>-293.88</td>
<td>-101.08</td>
<td>-242.95</td>
</tr>
<tr>
<td></td>
<td>(774.21)</td>
<td>(207.73)</td>
<td>(629.49)</td>
</tr>
<tr>
<td>Trend x Surveyed</td>
<td>0.1532</td>
<td>0.05151</td>
<td>0.1274</td>
</tr>
<tr>
<td></td>
<td>(0.3961)</td>
<td>(0.1063)</td>
<td>(0.3221)</td>
</tr>
<tr>
<td>Observations</td>
<td>160</td>
<td>160</td>
<td>160</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.958</td>
<td>0.922</td>
<td>0.965</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Economic Report of the Wine Industry. OLS regression predicting the production of table and fortified wines on a time trend interacted with an indicator variable for surveyed and non-surveyed counties. Data are at county level. Standard errors in parentheses are block-bootstrapped at the county level with 1,000 replications.
Table 1.A.4: Time Trends for Grape Crush by Variety before the Raisin Lay Survey for Surveyed and Non-Surveyed Counties, 1950-1959

<table>
<thead>
<tr>
<th></th>
<th>(1) Total</th>
<th>(2) Raisins</th>
<th>(3) Wine</th>
<th>(4) Table</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trend</strong></td>
<td>-0.04493</td>
<td>0.1310</td>
<td>-0.01968</td>
<td>-0.1563</td>
</tr>
<tr>
<td></td>
<td>(0.2827)</td>
<td>(0.1132)</td>
<td>(0.1130)</td>
<td>(0.1503)</td>
</tr>
<tr>
<td><strong>Surveyed</strong></td>
<td>-6,360.9</td>
<td>-10,710</td>
<td>-6.5078</td>
<td>4,355.7*</td>
</tr>
<tr>
<td></td>
<td>(9,647.5)</td>
<td>(9,818.3)</td>
<td>(1,203.9)</td>
<td>(2,562.8)</td>
</tr>
<tr>
<td><strong>Trend x Surveyed</strong></td>
<td>3.2854</td>
<td>5.5025</td>
<td>0.008306</td>
<td>-2.2254</td>
</tr>
<tr>
<td></td>
<td>(4.9357)</td>
<td>(5.0231)</td>
<td>(0.6159)</td>
<td>(1.9112)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>268</td>
<td>268</td>
<td>268</td>
<td>268</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.962</td>
<td>0.933</td>
<td>0.978</td>
<td>0.914</td>
</tr>
<tr>
<td><strong>County FE</strong></td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Marketing California Grapes, Raisins, and Wine Report and the Economic Report of the Wine Industry. OLS regression predicting grape crush by variety on a linear time trend interacted with an indicator variable for surveyed and non-surveyed counties. Data are at the county level. Standard errors in parentheses are block-bootstrapped at county level with 1,000 replications.
Table 1.A.5: Time Trends for Bearing Acreage by Variety before the Raisin Lay Survey for Surveyed and Non-Surveyed Counties, 1950-1959

<table>
<thead>
<tr>
<th>Bearing Vines (Thousand Acres)</th>
<th>(1) Total</th>
<th>(2) Raisins</th>
<th>(3) Wine</th>
<th>(4) Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trend</td>
<td>-0.1155***</td>
<td>-0.01901**</td>
<td>-0.08596***</td>
<td>-0.01050</td>
</tr>
<tr>
<td></td>
<td>(0.02599)</td>
<td>(0.007384)</td>
<td>(0.01902)</td>
<td>(0.006755)</td>
</tr>
<tr>
<td>Surveyed</td>
<td>1,915.2</td>
<td>1,206.1</td>
<td>255.71</td>
<td>453.38</td>
</tr>
<tr>
<td></td>
<td>(2,197.2)</td>
<td>(1,781.0)</td>
<td>(240.76)</td>
<td>(281.49)</td>
</tr>
<tr>
<td>Trend x Surveyed</td>
<td>-0.9710</td>
<td>-0.6090</td>
<td>-0.1304</td>
<td>-0.2316</td>
</tr>
<tr>
<td></td>
<td>(1.1234)</td>
<td>(0.9105)</td>
<td>(0.1231)</td>
<td>(0.1439)</td>
</tr>
<tr>
<td>Observations</td>
<td>399</td>
<td>399</td>
<td>399</td>
<td>399</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.996</td>
<td>0.997</td>
<td>0.989</td>
<td>0.998</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the California Fruits and Nut Acreage Report and the Economic Report of the Wine Industry. OLS regression predicting the acreage of bearing vines by variety on a linear time trend interacted with an indicator variable for surveyed and non-surveyed counties. Data are at county level. Standard errors in parentheses are block-bootstrapped at the county level with 1,000 replications.
Table 1.A.6: Time Trends for Wineries’ Vineyards and Capacity before the Raisin Lay Survey for Surveyed and Non-Surveyed Counties, 1950-1959

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Vineyard Storage Cap.</td>
<td>Fermentation Cap.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Log Acres)</td>
<td>(Log Thousand Gals)</td>
<td>(Log Thousand Gals)</td>
</tr>
<tr>
<td>Time Trend</td>
<td>0.01161*</td>
<td>0.007076*</td>
<td>0.001683</td>
</tr>
<tr>
<td></td>
<td>(0.006624)</td>
<td>(0.003662)</td>
<td>(0.003415)</td>
</tr>
<tr>
<td>Trend x Surveyed</td>
<td>0.03414</td>
<td>0.09761</td>
<td>0.01846</td>
</tr>
<tr>
<td></td>
<td>(0.02492)</td>
<td>(0.06197)</td>
<td>(0.01384)</td>
</tr>
<tr>
<td>Observations</td>
<td>240</td>
<td>580</td>
<td>449</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.998</td>
<td>0.990</td>
<td>0.999</td>
</tr>
<tr>
<td>Winery FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Wines & Vines Annual Directory. OLS regression predicting the acreage of bearing wines vines owned by winemakers and their storage and fermenting capacity on a linear time trend interacted with an indicator variable for surveyed and non-surveyed counties. Data are at winery level. Standard errors in parentheses are block-bootstrapped at county level with 1,000 replications.
Table 1.A.7: Placebo Tests for the Acreage of Bearing Fruit Trees Different from Grapes

<table>
<thead>
<tr>
<th>Bearing Trees ( Thousand Acres)</th>
<th>Deciduous</th>
<th>Citrus</th>
<th>Other Fruits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveyed x Post</td>
<td>551.461</td>
<td>5,046.889</td>
<td>19.198</td>
</tr>
<tr>
<td></td>
<td>(1,121.806)</td>
<td>(3,270.209)</td>
<td>(274.805)</td>
</tr>
<tr>
<td>Observations</td>
<td>435</td>
<td>435</td>
<td>435</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.962</td>
<td>0.907</td>
<td>0.982</td>
</tr>
<tr>
<td>Mean Surveyed: 1950-59</td>
<td>13,320.10</td>
<td>2,366.20</td>
<td>9,024.30</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

**Deciduous:** Apples, apricots, cherries, nectarines, peaches, pears, plums, prunes.

**Citrus:** Grapefruits, lemons, limes, oranges.

**Other:** Avocados, dates, figs, olives, persimmons, pomegranates, quinces.

*Notes:* ***p<0.01, **p<0.05, *p<0.1. Data from California Fruits and Nut Acreage Report and the Economic Report of the Wine Industry. OLS estimation of equation 1.1 using the acreage of trees bearing fruit as the outcome. Data is at county level. Standard errors in parentheses are block-bootstrapped at county level with 1,000 replications.
Table 1.A.8: Effects of the Raisin Lay Survey on Grape Crush

<table>
<thead>
<tr>
<th></th>
<th>Crushed (Thousand Tons)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Raisins</td>
<td>Wine</td>
<td>Table</td>
</tr>
<tr>
<td>Surveyed x Post</td>
<td>75.538***</td>
<td>77.254***</td>
<td>7.2266***</td>
<td>-8.9426***</td>
</tr>
<tr>
<td></td>
<td>(27.769)</td>
<td>(24.527)</td>
<td>(2.2431)</td>
<td>(3.8111)</td>
</tr>
<tr>
<td>Observations</td>
<td>368</td>
<td>368</td>
<td>368</td>
<td>368</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.964</td>
<td>0.939</td>
<td>0.972</td>
<td>0.935</td>
</tr>
<tr>
<td>Mean Surveyed:</td>
<td>1950-59</td>
<td>267.04</td>
<td>164.36</td>
<td>55.72</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Marketing California Grapes, Raisins, and Wine Report, and the Economic Report of the Wine Industry. OLS estimation of equation 1.1 using total grape crush, and raisin, wine, and table grape crush as the outcome. Standard errors in parentheses are block-bootstrapped at county level with 1,000 replications.
Table 1.A.9: Effects of the Raisin Lay Survey on Wine Production using Different Samples for Control and Treatment Groups

<table>
<thead>
<tr>
<th>Sample</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surv x Post</td>
<td>5.07***</td>
<td>2.36**</td>
<td>2.74**</td>
<td>4.16***</td>
<td>6.20***</td>
<td>3.56***</td>
</tr>
<tr>
<td>R-2</td>
<td>0.967</td>
<td>0.962</td>
<td>0.966</td>
<td>0.967</td>
<td>0.974</td>
<td>0.954</td>
</tr>
<tr>
<td>Mean Sur</td>
<td>23.07</td>
<td>23.07</td>
<td>23.07</td>
<td>23.07</td>
<td>23.07</td>
<td>6.90</td>
</tr>
<tr>
<td>Obs</td>
<td>253</td>
<td>96</td>
<td>109</td>
<td>157</td>
<td>189</td>
<td>237</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Economic Report of the Wine Industry. OLS estimation of equation 1.1 using total production as the outcome. The estimation in column (1) uses the main specification, which includes all raisin-growing counties. The estimations from columns (2) to (5) vary the counties included in the control group, such that (2) includes only counties in the San Joaquin Valley (SJV); (3) only in the Central Valley (CV); (4) the Central Valley and Southern California (CV+SC); and (5) excludes counties in the SJV that are not surveyed. Column 6 excludes Fresno from the treatment group (i.e., uses only Madera County as the treatment unit). Standard errors in parentheses are block-bootstrapped at county level with 1,000 replications.
Table A.10: Effects of the Raisin Lay Survey on Grape Crush Using Different Samples for Control and Treatment Groups

<table>
<thead>
<tr>
<th>Total Grape Crush (Thousand Tons)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surv x Post</td>
<td>75.53***</td>
<td>53.13**</td>
<td>54.95**</td>
<td>70.14***</td>
<td>82.73***</td>
<td>43.44***</td>
</tr>
<tr>
<td></td>
<td>(27.67)</td>
<td>(25.70)</td>
<td>(22.91)</td>
<td>(26.67)</td>
<td>(29.03)</td>
<td>(15.05)</td>
</tr>
<tr>
<td>Sample</td>
<td>All</td>
<td>SJV</td>
<td>CV</td>
<td>CV+SC</td>
<td>All ex. SJV</td>
<td>ex. Fresno</td>
</tr>
<tr>
<td>Obs2</td>
<td>368</td>
<td>101</td>
<td>139</td>
<td>221</td>
<td>285</td>
<td>353</td>
</tr>
<tr>
<td>R-2</td>
<td>0.964</td>
<td>0.960</td>
<td>0.966</td>
<td>0.964</td>
<td>0.965</td>
<td>0.960</td>
</tr>
<tr>
<td>Mean Surv</td>
<td>267.04</td>
<td>267.04</td>
<td>267.04</td>
<td>267.04</td>
<td>267.04</td>
<td>69.32</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Notes:</td>
<td>***p&lt;0.01, **p&lt;0.05, *p&lt;0.1. Data from Marketing California Grapes, Raisins, and Wine Report and the Economic Report of the Wine Industry. OLS estimation of equation 1.1 using total grape crush as the outcome. The estimation in column (1) uses the main specification, which includes all raisin-growing counties. The estimations from columns (2) to (5) vary the counties included in the control group, such that (2) includes only counties in the San Joaquin Valley (SJV); (3) only in the Central Valley (CV); (4) the Central Valley and Southern California (CV+SC); and (5) excludes counties in the SJV that are not surveyed. Column 6 excludes Fresno from the treatment group (i.e., uses only Madera County as the treatment unit). Standard errors in parentheses are block-bootstrapped at county level with 1,000 replications.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1.A.11: Effects of the Raisin Lay Survey on Total Grape Vine Acreage Using Different Samples for Control and Treatment Groups

<table>
<thead>
<tr>
<th></th>
<th>Total Grape Vine Acreage (Thousand Acres)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Surv x Post</td>
<td>12.68*** 2.18</td>
</tr>
<tr>
<td>Sample</td>
<td>All</td>
</tr>
<tr>
<td>Obs</td>
<td>969</td>
</tr>
<tr>
<td>R-2</td>
<td>0.995</td>
</tr>
<tr>
<td>Mean Surv</td>
<td>82.60</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from Marketing California Grapes, Raisins, and Wine Report and the Economic Report of the Wine Industry. OLS estimation of equation 1.1 using total grape vine acreage as the outcome. The estimation in column (1) uses the main specification, which includes all raisin-growing counties. The estimations from columns (2) to (5) vary the counties included in the control group, such that (2) includes only counties in the San Joaquin Valley (SJV); (3) only in the Central Valley (CV); (4) the Central Valley and Southern California (CV+SC); and (5) excludes counties in the SJV that are not surveyed. Column 6 excludes Fresno from the treatment group (i.e., uses only Madera County as the treatment unit). Standard errors in parentheses are block-bootstrapped at county level with 1,000 replications.
Table 1.A.12: Effects of the Raisin Lay Survey on Winemakers’ Investment Using a Balanced Panel

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveyed x Post</td>
<td>0.6522***</td>
<td>0.1490***</td>
<td>-0.01039</td>
</tr>
<tr>
<td></td>
<td>(0.09319)</td>
<td>(0.0628)</td>
<td>(0.03252)</td>
</tr>
<tr>
<td>Observations</td>
<td>273</td>
<td>896</td>
<td>732</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.991</td>
<td>0.983</td>
<td>0.988</td>
</tr>
<tr>
<td>Mean Surveyed: 1955-59</td>
<td>1426.63</td>
<td>5946.58</td>
<td>1506.89</td>
</tr>
<tr>
<td>Winery FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Wines & Vines Annual Directory and Buyer’s Guide. OLS estimation of equation 1.4 using wine grape vineyard acreage and the storage and fermenting capacity of winemakers as the outcomes. Standard errors are block-bootstrapped at county level with 1,000 replications. Data are at winery level.
Table 1.A.13: Spillover Effects of the Raisin Lay Survey on Winemakers’ Investment

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Vineyard (Log Acres)</td>
<td>Storage Cap. (Log 1000 Gals)</td>
<td>Ferm. Cap. (Log 1000 Gals)</td>
</tr>
<tr>
<td>Post</td>
<td>0.1290**</td>
<td>0.1584***</td>
<td>0.01306</td>
</tr>
<tr>
<td></td>
<td>(0.05111)</td>
<td>(0.02659)</td>
<td>(0.02480)</td>
</tr>
<tr>
<td>Post x Distance (miles)</td>
<td>-0.00067*</td>
<td>-0.00067***</td>
<td>0.00015</td>
</tr>
<tr>
<td></td>
<td>(0.0003)</td>
<td>(0.0002)</td>
<td>(0.00019)</td>
</tr>
<tr>
<td>Observations</td>
<td>601</td>
<td>1,358</td>
<td>1,057</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.983</td>
<td>0.987</td>
<td>0.990</td>
</tr>
<tr>
<td>Mean Surveyed:</td>
<td>1955-59</td>
<td>1426.63</td>
<td>5946.58</td>
</tr>
<tr>
<td>Sample</td>
<td>CV+SC</td>
<td>CV+SC</td>
<td>CV+SC</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Wines & Vines Annual Directory and Buyer’s Guide. OLS estimation of equation 1.4 using wine grape vineyard acreage and the storage and fermenting capacity of winemakers as the outcomes. Standard errors are block-bootstrapped at county level with 1,000 replications. Data are at winery level.
CHAPTER 2

Creative Financing and Public Moral Hazard: Evidence from Medicaid Supplemental Payments

This chapter shows theoretically and empirically how joint Medicaid funding can result in price and volume distortions. Our empirical analysis combines audit, survey, and administrative datasets on skilled nursing facilities (SNFs) from 1999-2017 with two reforms and difference-in-differences models. We first document that states use creative financing schemes to divert federal Medicaid matching funds. Using the case study of Indiana, we then document that these schemes lead to an increase in Medicaid SNF days for dementia patients. The expansion of SNFs of lower match quality leads to an increase in mortality pointing to a misallocation of vulnerable patients to providers.

2.1 Introduction

Joint funding of public programs has been a cornerstone of U.S. federalism. In 2019, the federal government provided states and local governments about $750bn in federal grants to fund a wide range of public policies, including health care, education, transportation, and environmental protection (Dilger and Cecire, 2019). In many instances, the joint funding involves federal matching grants to subsidize the spending
decisions made by lower levels of government. Inter-governmental matching grants ensure local control to balance heterogeneous preferences over public goods while internalizing externalities across local jurisdictions (Oates, 1999). Whether and how such joint funding affects the allocation of program resources and the quality and quantity of services is of central importance for the design of efficient policy in federal economies.

In this paper, we study these questions in the context of the Medicaid program. Medicaid, public insurance for low-income populations, absorbs the lion's share (60%) of overall federal grants to states and local government. The federal government matches state spending through the Federal Medicaid Assistance Percentage (FMAP), which denotes the federal cost share of total spending. However, states use “creative financing schemes (CFS)” to divert these federal funds, as shown by Baicker and Staiger (2005), resulting in an unintended transfer of resources between programs. This paper is the first to show, theoretically and empirically, that these schemes provide incentives to increase the volume of health care services. This leads to increased total (state and federal) Medicaid spending and distorts investments and the allocation of patients towards health care providers that are amenable to CFS.

We start with a basic theoretical point on how CFS can distort health care prices and volume in opposing directions. When states can transfer matching funds across programs at no cost, on the margin, they internalize changes in Medicaid provider reimbursement rates in full. The effective (marginal) FMAP for prices decreases to 0% (Baicker and Staiger, 2005). States have now incentives to lower the effectively paid reimbursement rate, which creates a “public markdown” between nominal and effective rates. By contrast, and novel to our analysis, the ability to transfer funds
increases the effective (marginal) FMAP on quantities beyond the nominal FMAP. States have now incentives to increase the volume of care, covered by Medicaid, to increase federal matching funds, which are (quantity) open-ended. We refer to this mechanism as “public moral hazard.”

While, in principle, this conceptual point applies broadly, its practical importance remains an empirical question. We quantify the diversion of federal matching funds, the associated public markdown, and the public moral hazard mechanism in an important, yet understudied, setting: Medicaid supplemental payments (SPs) for skilled nursing facilities (SNFs). SPs top up Medicaid base payments and are typically lump-sum payments, making them particularly susceptible for creative financing practices. In 2019, SPs accounted for seven percent of total Medicaid FFS nursing facility payments (MACPAC, 2019b).

Our empirical analysis proceeds in two steps. First, we present several novel facts on the prevalence of CFS and the diversion of Medicaid funds in the nursing home industry. We combine survey and audit datasets from 2000-2017 and evaluate a national reform to end these practices in 2003. Prior to 2003, 15 states used schemes to divert SPs from private and public nursing homes to fund other programs. The reform disallowed the diversion of SPs from private SNFs, limiting these practices to the small fraction of public SNFs. Using a difference-in-differences (DD) design, we find an average public markdown of 20%, which implies that only 80% of total nominal spending was effectively spent on nursing homes in states with a SP scheme between 2000 and 2002.

Second, to explore volume distortions, we focus on a specific CFS in Indiana. It allows us to exploit detailed institutional information and quasi-experimental varia-
tion. In an effort to side-step the 2003 regulation, the state of Indiana encouraged county hospitals to acquire private SNFs to divert federal matching funds. Between 2000 and 2017, the share of county-owned nursing homes increased from 5% to about 90%. Using administrative micro data, we study the impact of SNF acquisitions on the allocation of care in an event study framework. Consistent with our theoretical predictions, we find volume increases. The number of Medicaid days increases by 7%, driven by newly admitted dementia patients. These vulnerable patients are often admitted to SNFs of low “match quality” for dementia care. While these lower quality SNFs invest in Alzheimer units and staffing following the municipalization, we still find a meaningful increase in dementia patient mortality. This points to a misallocation of vulnerable patients across long-term care providers.

Our findings contribute to several literatures. First, we provide novel insights on Medicaid CFS, which mostly received attention in the health policy literature (Coughlin and Zuckerman, 2003a; Mitchell, 2018), law journals (Hatcher, 2017), and the popular press (Evans et al., 2020a), but thus far little attention in the economics literature. One exception is the seminal paper by Baicker and Staiger (2005), who document the diversion of matching funds in the context of Disproportionate Share Hospital (DSH) payments. They show that these schemes increase the cost of the program to the federal government and decrease the effect of the matching grant on total program resources, as states can avoid increasing their own contributions.

We build on and contribute to Baicker and Staiger (2005) in important ways. We document the extent of CFS in a different industry that critically relies on Medicaid funds. More importantly, we derive theoretically and document empirically how CFS can increase Medicaid volume and hence total (state and federal) spending. This biases investments and the allocation of Medicaid patients toward SNF care, that is
particularly amenable to CFS. We leverage administrative micro data to document how these schemes affect the care sequence and subsequent health outcomes among dementia patients, who constitute 46% of SNF residents and 21% of seniors living in the community (Garfield et al., 2015).

Second, our paper contributes to how ownership structures interact with health care regulation (Duggan, 2000; Grabowski and Hirth, 2003; Grabowski et al., 2013; Gupta et al., 2021; Gandhi et al., 2021). A distinct feature of Indiana’s scheme is that county hospitals were able to appropriate the incremental SPs and to leave the SNF management unchanged. This suggests that county hospitals (and the acquired SNFs) internalize the financial incentives provided by the CFS. This also suggests that seminal theories on organizational differences between public and private organizations—e.g. soft budgets and altruistic motives—are less applicable to our setting, reconciling why our findings are qualitatively consistent with how private hospitals responded to increases in Medicaid reimbursement rates through the DSH program. Specifically, Duggan (2000) shows that private hospitals increased Medicaid volume without improving the quality of care.

Third, we contribute to a large literature on how cost-sharing between patients, providers, and insurers (and the resulting party’s financial incentives) affect health care utilization and outcomes (Finkelstein et al., 2012a; Baicker et al., 2015; Brot-Goldberg et al., 2017). Our analysis shows how program cost-sharing between governments may affect policies and consequently the delivery of care. We show how CFS may distort the quality-quantity tradeoff in health care utilization, result in an expansion of low-value care, and affect patient mortality.

Finally, our analysis contributes to a literature on the role of intergovernmental
tal interactions in federal and state spending for public policies (Baicker et al., 2012; Gordon and Cullen, 2012). This includes important work on the link between federal matching grants, state spending and revenues (Baicker, 2001, 2005). Our analysis shows how CFS distort price setting in health care markets and result in volume increases; that is, state-induced overservicing. Thus, our analysis enriches an active debate on Medicaid financing reform, including increased monitoring (Federal Register, 2019), changes in cost sharing (Rudowitz, 2017; Clemens and Ippolito, 2018), and their broader fiscal impact (Gruber and Sommers, 2020).

2.2 Institutional Setting

2.2.1 Provider Reimbursement and Creative Financing Pre-2003

State Medicaid programs make base payments to SNFs on a per patient and day (per diem) basis (MACPAC, 2019a). To top up the standard Medicaid per diem base rate, states can authorize SPs to nursing homes (and hospitals). The combined payments, per diem plus SP, are then matched with federal funds according to the FMAP, as long as the combined rate per service remains below the “Upper Payment Limit” (UPL) (Mitchell, 2018). The UPL corresponds to the rate that Medicare would pay for the service. Typically, it is at least twice the Medicaid per diem rate.¹

Prior to 2003, states could pay the total UPL funds accrued by public and private nursing homes as a lump-sum to county-owned nursing homes. Specifically, the regulations pooled county and private SNFs in one category and left states full discretion

---

¹In 2018, the average Medicare rate per patient day equaled $515 ($427 for managed Medicare) compared to only $209 for Medicaid (Spanko, 2019).
in how to allocate SPs across providers in that coarse category. As a result, states could pay hundreds of millions in Medicaid funds to public nursing homes, triggering further hundreds of millions in federal matching funds to the same public nursing homes. These payments were not tied to specific services. This allowed states to funnel large portions back to the state via Inter-Governmental Transfers (IGTs).

Panel A of Appendix Figure 2.A.1 shows an example of such a scheme (Mangano, 2001; Coughlin and Zuckerman, 2003a). In this case, Pennsylvania paid $300m (figures rounded) in UPL-SPs to 23 county-owned nursing homes. Given the FMAP of 56%, this action triggered $400m in federal matching funds for a total of $700m in UPL-SPs to the county nursing facilities. However, only $1.5m remained with the providers, while the rest was transferred to the Pennsylvania government via an IGT. That means that the state turned a net surplus of almost $400m and diverted $700m in Medicaid funds.

2.2.2 The 2003 Reform of Supplemental Payments

Twenty-nine states operated IGT programs in 2001, many for nursing homes (U.S. Department of Health and Human Services, 2001). In 2000, in response to these creative financing schemes, Congress issued a directive to the Secretary of Health and Human Services (HHS) to limit federal Medicaid spending on SPs (Ku and Park, 2001; Federal Register, 2001). Effective 2003, states could no longer pay SPs accrued by private SNFs to county-owned SNFs. While states could still divert SPs accrued by public SNFs, the small market share of public SNFs significantly limited
the potential for redirecting money via IGTs.²

The reform affected states differently, depending on their pre-reform creative financing activities. Survey responses collected by Coughlin et al. (2004a) indicate that 23 out of 34 responding states used UPL-SPs in FY 2002. 15 states used UPL-SPs schemes in the nursing home industry, see Figure 2.A.2. We exploit these differences in a differences-in-differences research design to quantify the amounts of funds that were diverted prior to the reform.

2.2.3 Creative Financing 2.0: Evidence from Indiana

Following the 2003 reform, several states explored alternative CFS to divert federal matching funds. One prominent example are provider taxes that operate analogously to the pre-reform schemes, see e.g. Fosdick (2007); Miller and Wang (2009a) for detailed discussion.³

We study a CFS in Indiana. In efforts to sidestep the reform, Indiana encouraged county-owned hospitals to acquire private SNFs (turning them into county-owned SNFs). The goal of these acquisitions was to increase federal Medicaid funds that can be transferred between programs via IGTs. Specifically, the state would make per diem and SPs to the acquired SNF, that would be paired with federal matching funds. The county hospital could then divert any amount, potentially even exceeding the SPs, via IGTs in order to finance its hospital operations (Hatcher,

²The main provision dividing the existing category into public (county-owned) and private providers, but leaving states full discretion in the allocation of SPs within the refined categories.

Panel B of Appendix Figure 2.A.1 illustrates this mechanism.

This model was pioneered by the Health and Hospital Corporation (HHC) of Marion County—a government entity—which acquired 17 nursing homes in 2003. By the end of 2009, HHC owned 40 SNFs throughout the state. In most cases, the county hospital would outsource the operations to a private supplier who kept between 5% and 20% of the UPL payments (Hatcher, 2017). In 2010, Indiana allowed county hospitals to acquire SNFs using a “lease structure” instead of a straight acquisition (Pahud and Myers, 2016), facilitating further acquisitions of SNFs as we show below. Under a straight acquisition, the hospital purchases the provider’s entire assets required for operating the facility. Under a lease structure, by contrast, the hospital just acquires the license for the SNF and leases the assets to trigger a change of ownership, leaving the SNF management unchanged. Several states have since tried to copy Indiana’s model (Evans et al., 2021).

A distinct feature of the Indiana model is that the county hospitals were allowed to retain the incremental funds, suggesting that they internalize the financial incentives provided by the CFS. In our conceptual framework, detailed below, we assume that hospital and state objectives are aligned and hence abstract away from the county hospitals as separate stakeholders. We return to a brief discussion of potential conflicts of interest between county hospitals and the state at the end of the following section.

### 2.3 Conceptual Framework

CFS can distort health care prices and volume in opposing directions. We formalize these points in a stylized framework that illustrates the economic incentives faced by
states in their Medicaid spending decisions.

2.3.1 Baseline Model: No Creative Financing

We model the benefit of SNF health care provision as $B(\theta, Q)$, in dollars. The health benefits depend on the effective reimbursement rate, $\theta \geq 0$, which may be positively related to inputs (e.g. the nurse to patient ratio). Health benefits also depend on the volume of services provided $Q \geq 0$. We assume diminishing marginal benefits in $\theta$ and $Q$ and $B_\theta(0, Q) > 0$ and $B_Q(\theta, 0) > 0$. $B(\theta, Q)$ may include patient utility over nursing home care relative to the next best alternative as well as potential cost externalities (cost savings) from substitutes to nursing home care, e.g., hospital, formal and informal home care. $\theta \cdot Q$ denotes total funding in dollars.

States fund a share $(1 - FMAP)$ of total expenses as long as the reimbursement rate falls short of the UPL. (Below, when we introduce CFS, we consider potentially binding reimbursement rate ceilings.) To simplify the analysis, we assume that states not only choose $\theta$ but also $Q$ directly. Real world policies that directly affect demand and supply of SNF care include state home-and-community based services (HCBS) waiver programs for SNF care, Medicaid eligibility criteria, or admission decisions to public SNFs. Our model predictions therefore speak more directly to these policy tradeoffs. States solve:

$$\max_{\theta, Q} B(\theta, Q) - (1 - FMAP) \cdot \theta \cdot Q ,$$

which implies
\[ B_\theta(\theta, Q)/Q = 1 - FMAP \]  
\[ B_Q(\theta, Q)/\theta = 1 - FMAP \]

Absent CFS, states spend \( 1 - FMAP \) of every marginal dollar in Medicaid treatment quantities. In the optimum, this equals the marginal benefit of Medicaid rates per unit of quantity, as in equation (2.2). In the optimum, \( 1 - FMAP \) also equals the marginal benefit of quantity per dollar of Medicaid rates, as in equation (2.3).

### 2.3.2 Joint Funding Under Creative Financing Mechanisms Pre 2003

We now allow for CFS by introducing a separate nominal billing rate, \( P \), which is billed to the federal government. In contrast, \( \theta \) is the effective rate actually paid to the provider. We allow for \( \theta \neq P \), and refer to \( \mu = \frac{P - \theta}{P} \) as the “public markdown.” \((P - \theta) \cdot Q \) is the diverted transfer amount. Now states optimize:

\[
\max_{\theta, Q, P} B(Q, \theta) - Q \cdot \theta + FMAP \cdot P \cdot Q,
\]

subject to \( P \leq \bar{P} \), where \( \bar{P} \) denotes the UPL. It is clear from this representation that \( P = \bar{P} \). The optimality conditions are then:

\[ B_\theta(\theta, Q)/Q = 1 \]
\[ B_Q(\theta, Q)/\theta = 1 - FMAP \cdot \frac{\bar{P}}{\theta} = 1 - \frac{FMAP}{1 - \mu} \].
The analysis yields two main predictions. First, states have an incentive to maximize the nominal rate if they can transfer unlimited amounts at no costs. Importantly, the ability to transfer funds renders the FMAP’s subsidy effect on rate setting ineffective. As indicated by equation (2.5), states internalize the full cost of rate increases on the margin. This provides states an incentive to lower the effective rate paid to providers. Baicker and Staiger (2005) first developed and tested this prediction.

**Hypothesis 1 (H1)** *Public markdown*: The ability to transfer funds yields a positive wedge between the nominal rate \( P \) and the effective rate \( \theta \), captured by the public markdown \( \mu = \frac{P - \theta}{P} \). This provides incentives for states to decrease the effective rate.

Second, and in contrast to the rate setting tradeoff, the ability to transfer funds amplifies the federal subsidy effect on the volume of care. As indicated by equation (2.6), this increases the (effective) quantity FMAP by \( \frac{P}{\theta} \) and incentivizes a higher volume of care. In principle, the effective quantity FMAP may exceed 100%. Then, on the margin, the state turns a net surplus from expanding Medicaid services.

**Hypothesis 2 (H2)** *Public moral hazard*: The ability of states to transfer funds increases the (effective) quantity FMAP to \( \frac{F_{MAP}}{1 - \mu} = F_{MAP} \cdot \frac{P}{\theta} \). This provides incentives for states to increase the volume of care.

### 2.3.3 Creative Financing Mechanisms by Ownership Type

To tie the predictions closer to Indiana’s scheme, we extend the framework by introducing different ownerships, \( \tau = \{pr, pub\} \) for private and public SNFs. States
optimally choose the share of public SNFs, $\rho$. They can choose different rates and quantities by ownership. States then maximize:

$$
\max_{\theta^r, Q^r, P^p, \rho} \quad (1 - \rho) \cdot \left[ B(Q^{pr}, \theta^{pr}) - Q^{pr} \cdot \theta^{pr} + FMAP \cdot P^{pr} \cdot Q^{pr} \right] - \rho \cdot \left[ B(Q^{pub}, \theta^{pub}) - Q^{pub} \cdot \theta^{pub} + FMAP \cdot P^{pub} \cdot Q^{pub} \right] - \kappa(\rho) \tag{2.7}
$$

where the first line corresponds to private nursing homes and the second line to public nursing homes. The correction term $\kappa(\rho)$ denotes benefits from care access, net of cost inefficiencies from operating public nursing homes, with $\kappa'(1) > 0 > \kappa'(0)$ and $\kappa''(\rho) > 0$. Prior to the 2003 reform, prices and quantities were symmetric between public and private SNFs. $\kappa'(\rho) = 0$ determined the optimal share of public SNFs.

The 2003 SP reform then prohibited the transfer of funds accrued by private providers. To this end, we impose:

$$
P^{pr} = \theta^{pr}. \tag{2.9}
$$

For private nursing homes, the optimality conditions in Section 2.3.1 determine the rate and quantities. For public nursing homes, the optimality conditions in Section 2.3.2 determine the rate and quantities. As a result, the predictions from H2 apply to public nursing homes. They imply volume differences between public and private nursing homes.

The constraint also affects the incentives to municipalize private nursing homes.
The optimal share of public nursing homes is now:

\[ 0 = \Delta \Pi (\text{pub}, \text{pr}) - \kappa'(\rho) \geq FMAP \cdot (\bar{P} - \theta^{pr}) \cdot Q^{pr} - \kappa'(\rho) \]

where \( \Delta \Pi (\text{pub}, \text{pr}) \) is the difference in net benefits between public and private nursing homes, ignoring the correction term \( \kappa(\rho) \).\(^4\) The inequality leverages the observation that the net benefit from municipalization must be at least as large as the incremental federal matching funds from increasing \( P \) to \( \bar{P} \) only, that is without optimally adjusting rates and quantities for public SNFs. As \( FMAP \cdot (\bar{P} - \theta^{pr}) \cdot Q^{pr} > 0 \) and \( \kappa''(\rho) > 0 \), it follows that the optimal share of public nursing homes will increase.

**Hypothesis 3 (H3) Public moral hazard:** The ability of states to transfer funds from public (but not private) nursing homes provides incentives to municipalize private nursing homes. The incentive to municipalize is bounded from below by the federal share of nominal Medicaid spending multiplied by the public markdown: \( FMAP \cdot \bar{P} \cdot Q^{pr} \cdot (1 - \frac{\theta^{pr}}{P}) = FMAP \cdot \bar{P} \cdot Q^{pr} \cdot \mu. \)

In an extension of the model in Appendix 2.B, we add a principal agent conflict between the state and the county hospital that acquires the SNFs in the context of Indiana. The state chooses \( P \), which is paid to the hospital, and the hospital chooses \( \theta \) and \( Q \) and transfers the surplus from the SNF operations, \( (P - \theta) \times Q \), towards its

\(^4\Delta \Pi (\text{pub}, \text{pr}) = \left[ B(Q^{\text{pub}}, \theta^{\text{pub}}) - Q^{\text{pub}} \cdot \theta^{\text{pub}} + FMAP \cdot P^{\text{pub}} \cdot Q^{\text{pub}} \right] - \left[ B(Q^{\text{pr}}, \theta^{\text{pr}}) - Q^{\text{pr}} \cdot \theta^{\text{pr}} + FMAP \cdot \theta^{\text{pr}} \cdot Q^{\text{pr}} \right].\)
hospital operations. The key difference to SNF profits earned by a private investor is that the state internalizes the SNF surplus of the public hospital in full. Instead, the conflict of interest between the hospital and the state is (only) that the hospital does not internalize the state’s cost of financing \((1 - FMAP) \times P \times Q\). We find that the state still chooses \(P = \bar{P}\) as long as the hospital’s quantity decision does not respond too elastically to \(P\). The quality tradeoff of the hospital coincides with equation (2.5) but the incentives to increase quantity are even larger, reinforcing the public moral hazard channel.

2.4 Data, Empirical Approach, and Results

Our empirical strategy proceeds in three steps at different levels of aggregation. We organize the discussion of the data, the empirical strategy, and results accordingly. First, using the 2003 reform along with state-year level SNF data in a (reverse) difference-in-differences framework, we quantify the diversion of Medicaid funds and the “public markdown.” Second, using variation in the timing of SNF acquisitions in Indiana in an event study framework, we quantify the distortionary effects of creative financing on nursing home operations, and specifically Medicaid volume. Third, using administrative micro data from Indiana, we turn to the mechanisms and study changes in health care utilization and patient outcomes.

2.4.1 Creative Financing in U.S. Nursing Homes

To estimate the effect of the 2003 reform on the diversion of Medicaid funds, we use LTC Focus data at the state-year level from 2000-2017. The data cover all Medicaid and Medicare certified SNFs (98% of all SNFs) nationwide. The data provide facility-
level information on staffing levels, patient demographics, facility characteristics, and the state-year level average Medicaid base per diem rate.

To estimate how much Medicaid effectively spends on SNF care, \( \theta \cdot Q \), we multiply the Medicaid per diem rate with total Medicaid days. We compare this estimate to total nominal spending, \( P \cdot Q \), which comprises spending on SNFs and intermediate care facilities (ICFs) in Centers for Medicare & Medicaid Services (2020a). We interpret the difference as the size of the transfers returned to the states. We then construct the markdown as:

\[
\mu_{st} = \frac{P \cdot Q - \theta \cdot Q}{P \cdot Q} = \frac{P - \theta}{P}.
\]

To test hypothesis \( H_1 \), we run a DD model that compares the markdown among the 15 states with pre-2003 creative financing schemes (treatment group) to those 16 states without (control group), see Figure 2.A.2.

**Results:** Figure 2.1 displays the average (negative) public markdown, \(-\frac{P - \theta}{P}\). Consistent with hypothesis \( H_1 \), we find an average public markdown of 20% among treatment states with a SP scheme in pre-reform years. Concurrent with the timing of the reform, the negative markdown increases swiftly to 0. More formally, our DD regression estimate implies a 19 percentage point (ppt) increase.\(^5\)

Considering the nominal Medicaid spending of $58bn on nursing home care between 2000 and 2002 among states with a SP scheme, this average pre-reform markdown of 19% suggest that states diverted 19%*$57bn=$11bn. Further, Table 2.A.3 (Appendix) shows no staffing changes, despite the lower total (state and federal)

\(^5\)Differences between states are shown in Appendix Figure 2.A.3.
Figure 2.1: Public Markdowns and the Supplemental Payments Reform

Source: Long-Term Care: Facts on Care in the U.S. (2020); Centers for Medicare & Medicaid Services (2020a). This figure presents the negative markdown, $-\frac{\theta - \theta_0}{\theta}$ (equation (2.10)), which indicates the (negative) share of nominal SNF Medicaid funds diverted towards other uses, for states with and without a supplemental payment scheme prior to the national reform, see Figure 2.A.2. The vertical line denotes the time of the supplemental payments reform.
nominal Medicaid spending following the reform (-$261m per state/year). These findings reinforce that those marginal nominal funds were not effectively spent on nursing homes but instead (mis)used for other public programs.

To lend further support to these estimates, we digitized and analyzed 24 audit reports from 18 states in the pre-reform period 1997-2003. They provide independent estimates on the diversion of funds for a subset of states. The reports show an average markdown of 17.9% in states with creative financing schemes (Appendix Figure 2.C.1). Appendix Section 2.C provides details.

2.4.2 Distortionary Effects of Creative Financing

To circumvent the 2003 reform, and consistent with hypothesis H3, Indiana encouraged county hospitals to acquire private SNFs. Figures 2.2a and b show that, right after the enactment of the reform, the number of county-owned nursing homes more than doubled from 14 to 29 within a year. It almost doubled again by 2010 and by 2017, 90% of all nursing homes were county-owned. Appendix Figure 2.A.4 shows a map of all 92 counties in Indiana along with the share of county-owned nursing homes from 2000 to 2017.

To bolster the claim that the objective of these acquisitions was the diversion of Medicaid funds, using data from annual HHC financial reports, we identify direct money transfers. The numbers yield an average public markdown of 32% between 2008 and 2019, suggesting that 32% of all nominal Medicaid funds and about 78% of SPs to the SNF were diverted. We estimate an effective quantity $F_{MAP}$, $F_{MAP} \cdot \frac{P}{\theta}$, of 102%, exceeding the statutory maximum of 83%, and providing incentives to increase the quantity of care, see Appendix Figure 2.D.1 and Appendix Section 2.D.
Figure 2.2: Effect of 2003 Reform on Nursing Homes

Source: Centers for Medicare & Medicaid Services (2020b) and Indiana Department of Health (2020). The bottom four subgraphs show event studies as in equation (2.11) where the year before the SNF municipalization is the reference period (solid gray line). The dashed vertical lines indicate the beginning and the end of the implementation period. The pooled effect corresponds to the post-reform effect, pooled over 1-5 years after the conversion.
2.4.2.1 Effect on SNF Operations

To test hypothesis H2 and to estimate the impact of Indiana’s scheme on Medicaid volume and SNF operations, we next exploit the timing of the SNF acquisitions in an event study framework. We extract the timing information on municipalizations from Nursing Home Compare (Centers for Medicare & Medicaid Services, 2020b), the Indiana Department of Health (2020) data as well as reports by the IndyStar newspaper for validation. We then merge the timing information to the SNF-year level data from LTC Focus for the state of Indiana. Specifically, we estimate:

\[ y_{s\tau} = \sum_{\tau=-5}^{-2} \phi_{\tau} + \sum_{\tau=0}^{+T} \phi_{\tau} + \delta_s + \delta_t + \Gamma X_{s,t} + \varepsilon_{j\tau} \]  

(2.11)

where \( s \) indexes a SNF and \( \tau \) the years relative to the public acquisition (with \( \tau = 0 \) denoting the acquisition year). \( \delta_s \) and \( \delta_t \) are SNF and calendar year fixed effects and \( X_{s,t} \) are time varying SNF-level controls. The key coefficients of interest are indicators for years relative to the Medicaid transition, \( \phi_{\tau} \), where \( \phi_{-1} \) is the reference category.

To avoid that compositional changes in SNFs over the event period contaminate our estimates, for example see Callaway and Sant’Anna (2021), we balance our sample of SNFs in event time. That is, we focus on SNFs that we observe for full five years prior to the conversion and for full five years after the conversion (as well as never treated SNFs).\(^6\) Our final sample contains 2,203 SNF-year observations and

---

\(^6\)Thus, we exclude SNFs that were converted very early (pre 2005) or very late (2013-2017).
covers 2000-2017. Table 2.A.2 (Appendix) shows variable means, t-statistics and the normalized difference for SNFs that were converted early vs. late but are part of our sample. Except for three indicators, all normalized differences are below the threshold of 0.25 indicating imbalances according to Imbens and Wooldridge (2009). Early converted SNFs have lower occupancy rates, lower RN to nurse rations but higher shares of restrained patients.

We also present DD estimates that pool the post-acquisition years and treat the acquisition year \( \tau = 0 \) as an implementation period. We cluster all standard errors at the SNF level. In addition, we implement Two-Stage DD to correct for possible biases in staggered DD designs due to effect heterogeneity (Gardner, 2021; Butts and Gardner, 2022).

**Results:** Figure 2.2c-f presents the effects of ownership conversion on (c) for-profit status, (d) Medicaid volume, (e) Alzheimer units, and (f) the presence of a nurse practitioner or physician assistant (NP/PA). As expected, we see a large drop in not-for-profit ownership status in the year of acquisition (Figure 2.2c). The post-acquisition estimates range around -0.6 implying that 60% of the acquired nursing homes were formerly for-profit. The patterns also suggests that conversions are only fully realized one year after the official acquisition date.

Importantly, confirming hypothesis **H2**, Figure 2.2d shows a significant increase in Medicaid days following the municipalization of nursing homes. On average, Medicaid days increase by 7% or 1373 days per SNF and year over the post-conversion period (Panel B, column 2, Table 2.A.4). The increase contributes to a roughly 5ppt higher occupancy rate (Panel A, columns 5-8, Table 2.A.4).
Regarding mechanisms, Figure 2.2e shows an increase in Alzheimer units. Apparently, to increase Medicaid volume, SNFs targeted dementia patients who account for 38% of all Medicaid days in our sample. Consistent with this, in Figure 2.2f and Panel A (columns 1-4) of Table 2.A.4, we find a significant increase in NP/PAs who play a critical role in the delivery of care for dementia patients (Joyce et al., 2018). Two details concerning Indiana’s Medicaid reimbursement design may contribute to these patterns. First, like most states, Indiana risk-adjusts their Medicaid per diem rates, considering the acuity of the patient mix. That means that federal matching funds scale in patient severity, providing incentives to admit complex Medicaid patients after the municipalization. Second, Indiana pays higher Medicaid rates for patients in dementia special care units, see MACPAC (2019a).

2.4.3 Mechanisms and Patient Outcomes

2.4.3.1 Evidence from MDS SNF Stays

To further explore the mechanisms of increases in Medicaid volume, we now use high-quality patient-level assessment data from the Long-Term Care Minimum Data Set (MDS). This is administrative data on all Medicaid and Medicare certified nursing homes in Indiana, from 1999-2015. Given the evidence above, we identify dementia patients from the initial SNF admission assessment.\(^7\) Using the admission and discharge dates along with information on payers, we then construct “dementia admissions” and “dementia Medicaid days” per SNF and year. We drop SNFs not included in the balanced SNF sample in Section 2.4.1 to ease the comparison between

\(^7\)We combine the MDS 2.0 with MDS 3.0 and keep patients who have Section I diagnosis indicator I1q or I1u turned on (MDS 2.0 terminology).
analyses. We consider five years prior to the conversion but only three years after
the conversion as the MDS sample ends in 2015, reducing the number of SNF-year
observations to 1,882.

**Results:** Figure 2.3a documents an increase in admissions of dementia patients
from hospitals. Pooled over all post-conversion years, we estimate an increase of
3.9 admissions from acute care hospitals per year and SNF (Panel A, column 9, Ta-
ble 2.A.5). This corresponds to 71% of the overall increase (column 1) in dementia
patient admissions. In contrast, we find no evidence that the general increase in
admissions is driven by more admissions from the community (Figure 2.3b). Im-
portantly, and strongly reinforcing our findings above, Figure 2.3c documents an
increase in Medicaid dementia patient days with the effect size increasing over time.
Two and three years after reform, we estimate that Medicaid days increase by 1,430
(se=400) and 1,542 days (se=542), respectively. This implies a 15% increase relative
to the mean (or a 6% increase relative to all Medicaid days). Note the similarity of
this increase compared to the overall increase in Medicaid days in Figure 2.2d. We
infer that the large majority of the increase in Medicaid patient volume stems from
dementia patients.

Finally, we test for changes in the quality of care following the conversion. To this
end, we analyze the average one-year mortality (relative to the SNF admission date)
among dementia patients admitted to SNF s in year t in Figure 2.3d. The imprecisely
estimated coefficients yield no evidence for significant changes in mortality.
Figure 2.3: Mechanisms: Dementia Patients in the MDS SNF Sample Population

Source: The figures show event study coefficients as in equation (2.11) where the year before the SNF municipalization is the reference period (solid gray line). The dashed vertical lines indicate the beginning and the end of the implementation period. All event studies show 90% confidence intervals. The pooled effect corresponds to the post-reform effect, pooled over 1-3 years after the conversion. Equivalent DD regression results with two-stage Gardner (2021) correction are shown in Tables 2.A.5 and 2.A.6. Panel A shows results for the MDS (2015) SNF patient population. Panel B shows results for the MEDPAR (2015) hospital patient population, where patients are assigned to their closest SNF, see the text for details.
2.4.3.2 Evidence from MedPAR Hospital Stays

While we find no evidence for an improvement in SNF quality following conversions, the increase in Medicaid volume may reflect better access to inpatient care options and improve the allocation of dementia patients across inpatient and outpatient providers. If so, patient welfare may increase. If, on the other hand, converted SNFs are not able to provide high-quality dementia care, then incremental patients may be worse off in SNF care and welfare may fall.

To shed some light on these questions, we now switch perspective, and implement an approach similar to Einav et al. (2019). Namely, we cut an estimation sample at the hospital stay level. The idea is to study the counterfactual care choices and health outcomes of hospitalized dementia patients, who are at an increased risk of being discharged to a SNF once it is municipalized. To this end, we use inpatient fee-for-service Medicare claims from the CMS Medicare Provider Analysis and Review MedPAR file, spanning the years 1999-2015. We keep acute care hospital claims from Indiana and focus on dementia patients, who we identify based on the diagnosis codes.\footnote{We identify (and keep) dementia patients as patients who belong to Clinical Classification Software (CCS) code 653 “Delirium, dementia, and amnestic and other cognitive disorders”, see https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp for details.} We then assign each hospital stay of each dementia patient to a unique focal SNF. Focal SNFs are within 5 miles of the patient’s former address and have the earliest conversion date of SNFs in that radius. We pick the closest SNF if multiple SNFs were converted in the same year. We also assign the closest SNF if none of the SNFs within 5 miles was converted. These hospital stays define the control group. We drop hospital stays assigned to SNFs that are not included in the balanced SNF sample. The assignment of SNFs to patients, regardless of their...
actual discharge destination, allows us to study how the municipalization of a local SNF affects hospital discharge outcomes and also patient mortality.

Next, we merge the inpatient sample with the MDS at the patient level. We construct an indicator for “discharges to the focal SNF.” It turns on if we find a subsequent MDS stay in the focal SNF within two weeks of the hospital discharge. To explore potential substitution patterns, we construct analogues indicators for discharges to “any SNF” and discharges to “any non-assigned SNF.” We also construct an indicator for “discharges to lower care level providers.” It turns on if the MedPAR claim indicates a home discharge, or a discharge to a rehabilitative or an ICF. Finally, we construct the one-year mortality dated back to the admission date of the index hospital stay.

Results: Consistent with the evidence on increased SNF admissions from hospitals, Figure 2.4e shows an increased risk of being discharged to the focal SNF following its municipalization. The pooled post-conversion increase equals 2.2ppt, a 26% increase (Panel A, column 1 in Table 2.A.6). However, we find no significant changes in discharges to “any SNF”, and no changes in the one-year mortality rate.

However, the overall effects may mask important differences based on SNF’s quality of dementia care. To explore this, we flag SNFs as “lower quality” if they had no Alzheimer Unit and no NP/PA on their staff in the year before conversion. We then revisit the analysis on the subsample of patients who are assigned (not necessarily discharged) to these lower quality SNFs (Figures 2.4f, h and Panel B of Table 2.A.6). We again find significant increases in discharges to “focal SNFs”

---

9We exclude hospital stays initiated by the SNF, by dropping index stays that start within two weeks of a preceding SNF stay.
Figure 2.4: Mechanisms: Dementia Patients in the MedPAR Hospital Sample Population

Source: The figures show event study coefficients as in equation (2.11) where the year before the SNF municipalization is the reference period (solid gray line). The dashed vertical lines indicate the beginning and the end of the implementation period. All event studies show 90% confidence intervals. The pooled effect corresponds to the post-reform effect, pooled over 1-3 years after the conversion. Equivalent DD regression results with two-stage Gardner (2021) correction are shown in Tables 2.A.5 and 2.A.6. Panel A shows results for the MDS (2015) SNF patient population. Panel B shows results for the MEDPAR (2015) hospital patient population, where patients are assigned to their closest SNF, see the text for details.
(3.3ppt or 46%), see Figure 2.4f and Panel B, column 1, Table 2.A.6. We also find increases in discharges to “any SNF”, and decreases in discharges to lower level of care providers (Panel B, columns 5-8, Table 2.A.6). Importantly, we find an increase in the one-year mortality rate in this population, see Figure 2.4h. Mortality increases significantly by 3.6ppt or 11% following the conversion of lower quality SNFs (Panel B, columns 9-10, Table 2.A.6). This finding points to a misallocation of vulnerable dementia patients to lower quality SNFs.

2.5 Discussion and Conclusion

This paper argues that joint Medicaid funding can distort health care prices and inpatient volume. Using a range of methods and datasets, we estimate that states have used CFS to divert 20-30% of nominal Medicaid SNF spending. We show theoretically and empirically that these practices do not only lead to an unintended transfer of resources between programs, as shown already by Baicker and Staiger (2005), but they also distort the allocation of patients and resources across health care providers.

Using the case study of Indiana, we illustrate how the prospect of diverting federal matching funds provides perverse incentives to convert private into county-owned SNFs. The share of public SNFs increased from 5% to more than 90% between 2000 and 2017. Following the municipalization of SNFs, in line with our theoretical predictions, we document an increase in the volume of Medicaid dementia patients. We find particularly large increases in admissions among SNFs with limited experience and no Alzheimer units pre-conversion. Absent creative financing and municipalization, the incremental dementia patients would have received care at home or in an
intermediate or rehabilitative care facility. These less experienced SNFs did invest in Alzheimer units and NP/PAs post-conversion, contributing to the large increase in SNFs with Alzheimer units in Indiana over our sample period (the share of SNFs with Alzheimer units increased from 23% to 36% between 2000 and 2017, making it the highest nationwide). Nevertheless, we still find an increase in dementia patient mortality over three post-conversion years in these less experienced SNFs, pointing to a misallocation of vulnerable patients to providers.

Going beyond our empirical setting, the ability to transfer matching funds provides an explanation for persistently low Medicaid reimbursement rates that have contributed to the chronic quality shortfalls in the nursing home industry (Grabowski, 2001a; Hackmann, 2019). The public moral hazard mechanism may bias states towards Medicaid policies that encourage inpatient long-term care use. According to Evans et al. (2020a): “The state’s [Indiana’s] elder care system is now so skewed to nursing homes [...] that the expansion of alternative options such as in-home care has been stifled.” These policies include (i) Medicaid fee-for-service reimbursement (Hackmann et al., 2021), (ii) generous Medicaid eligibility standards (Grabowski and Gruber, 2007a; Mommaerts, 2018), and (iii) a lack of home and community-based care alternatives to nursing home care (Muramatsu et al., 2007; Guo et al., 2015a; Wang et al., 2020).

Finally, the presence of these schemes may also provide an explanation for the relative small market share of Medicaid managed care plans in long-term care that may help rebalance care delivery away from inpatient long-term care (Kaiser Commission on Medicaid and the Uninsured, 2015; Reaves and Musumeci, 2015b) and inpatient care more broadly (Duggan et al., 2018; Abaluck et al., 2021; Geruso et al., 2020). Specifically, actuarial soundness rules prohibit states from making supple-
mental payments for services covered under a managed care contract. According to MACPAC (2020) this led some states to exclude certain services or populations from managed care.
2.A Appendix: Additional Figures and Tables

Figure 2.A.1: Schemes using UPL Supplemental Payments for Nursing Facilities

(a) Panel A: Pennsylvania in FY2002

(b) Panel B: Indiana after 2003

Notes: Figures in round brackets reflect the final net position of the UPL payments. Figures in bold and square brackets reflect what was reported as Medicaid expenditures in the CMS. Figures for Panel A are based on the data from Mangano (2001) and depicted in Coughlin and Zuckerman (2003a). Figures for Panel B are based on Hatcher (2017)
Figure 2.A.2: States Use of UPL-SPs in Nursing Homes in FY 2002

Notes: This map illustrates which states used an upper payment limit supplemental payment (UPL-SP) scheme in FY 2002 based on the survey from Coughlin et al. (2004a). Pennsylvania is marked as using a UPL-SP scheme based on Coughlin and Zuckerman (2003a).
Figure 2.A.3: Public Mark downs and the Effective FMAP: Evidence from LTC Focus 2000-2002

Notes: The top figure presents the negative markdown, $-\frac{P - \theta}{P}$, which indicates the (negative) share of nominal SNF Medicaid funds diverted towards other uses. The bottom figure presents the nominal FMAP and the effective quantity FMAP: $FMAP - \frac{\theta}{P}$. The presented findings build on the following data sources from 2000-2002: Long-Term Care: Facts on Care in the U.S. (2020); Centers for Medicare & Medicaid Services (2020a).
Figure 2.A.4: Development of Shares of County-Owned Nursing Homes

Source: Centers for Medicare & Medicaid Services (2020b) and Indiana Department of Health (2020).
Table 2.A.1: LTCFocus Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>County-owned SNF</td>
<td>0.389</td>
<td>0.488</td>
<td>2203</td>
</tr>
<tr>
<td><strong>Panel A: Main outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>For-profit</td>
<td>0.586</td>
<td>0.493</td>
<td>2203</td>
</tr>
<tr>
<td>Occupancy rate</td>
<td>76.363</td>
<td>14.268</td>
<td>2203</td>
</tr>
<tr>
<td>Number Medicaid days</td>
<td>18590.507</td>
<td>10416.805</td>
<td>2203</td>
</tr>
<tr>
<td>Alzheimer’s Special Care Unit (SCU)</td>
<td>0.383</td>
<td>0.486</td>
<td>2203</td>
</tr>
<tr>
<td>Has Nurse practitioner or physician’s assistant</td>
<td>0.385</td>
<td>0.487</td>
<td>2203</td>
</tr>
<tr>
<td><strong>Panel B: Controls</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total county population</td>
<td>218,869</td>
<td>281,418</td>
<td>2203</td>
</tr>
<tr>
<td>County population above 65</td>
<td>26,962</td>
<td>31,511</td>
<td>2203</td>
</tr>
<tr>
<td>Share restrained</td>
<td>2.808</td>
<td>5.051</td>
<td>2203</td>
</tr>
<tr>
<td>Share w/ hypertension</td>
<td>62.666</td>
<td>17.047</td>
<td>2055</td>
</tr>
<tr>
<td>RN/Nurses Ratio</td>
<td>11.48</td>
<td>1.427</td>
<td>2203</td>
</tr>
<tr>
<td>Average Acuity Index</td>
<td>0.274</td>
<td>0.156</td>
<td>2203</td>
</tr>
<tr>
<td>Direct-care staff hours per resident day</td>
<td>3.518</td>
<td>1.471</td>
<td>2203</td>
</tr>
<tr>
<td>Registered Nurses hours per resident day</td>
<td>0.447</td>
<td>0.704</td>
<td>2203</td>
</tr>
<tr>
<td>Licensed Practical Nurse hours per resident day</td>
<td>0.976</td>
<td>0.426</td>
<td>2203</td>
</tr>
<tr>
<td>Certified Nursing Assistant hours per resident day</td>
<td>2.095</td>
<td>0.798</td>
<td>2203</td>
</tr>
<tr>
<td>SNF is hospital-based</td>
<td>0.037</td>
<td>0.189</td>
<td>2203</td>
</tr>
</tbody>
</table>

Source: Long-Term Care: Facts on Care in the U.S. (2020), authors’ own calculation.
<table>
<thead>
<tr>
<th>Variable Names</th>
<th>Early (mean)</th>
<th>Early (sd)</th>
<th>Late (mean)</th>
<th>Late (sd)</th>
<th>tstat</th>
<th>Norm Diff</th>
<th>tstat</th>
<th>Norm Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Main outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>For-profit</td>
<td>0.889</td>
<td>0.316</td>
<td>0.854</td>
<td>0.353</td>
<td>0.953</td>
<td>0.073</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occupancy rate</td>
<td>69,572</td>
<td>12,587</td>
<td>20,066</td>
<td>8,912</td>
<td>-5,655</td>
<td>0.471</td>
<td>-5.655</td>
<td>0.471</td>
</tr>
<tr>
<td>Medicaid days</td>
<td>9,685</td>
<td>20,066</td>
<td>20,066</td>
<td>8,912</td>
<td>0.438</td>
<td>0.114</td>
<td>1.384</td>
<td>0.114</td>
</tr>
<tr>
<td>Alzheimer’s disease Special Care Unit (SCU)</td>
<td>0.422</td>
<td>0.497</td>
<td>0.394</td>
<td>0.489</td>
<td>0.512</td>
<td>0.041</td>
<td>0.512</td>
<td>0.041</td>
</tr>
<tr>
<td>Nurse practitioner or physician’s assistant</td>
<td>0.244</td>
<td>0.432</td>
<td>0.312</td>
<td>0.464</td>
<td>-1.373</td>
<td>0.107</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total county population</td>
<td>247,939</td>
<td>31,413</td>
<td>22,700</td>
<td>11,112</td>
<td>11.115</td>
<td>0.147</td>
<td></td>
<td></td>
</tr>
<tr>
<td>County population above 65</td>
<td>20,066</td>
<td>8,912</td>
<td>20,066</td>
<td>8,912</td>
<td>0.409</td>
<td>0.118</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Acuity Index</td>
<td>11.115</td>
<td>0.207</td>
<td>3.171</td>
<td>0.682</td>
<td>1.344</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RN/Nurses Ratio</td>
<td>0.175</td>
<td>1.049</td>
<td>0.250</td>
<td>0.146</td>
<td>-3.721</td>
<td>0.248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct-care staff hours resident day</td>
<td>3.063</td>
<td>0.210</td>
<td>0.250</td>
<td>0.146</td>
<td>-3.721</td>
<td>0.248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Registered Nurses hours resident day</td>
<td>0.972</td>
<td>0.431</td>
<td>0.250</td>
<td>0.146</td>
<td>-3.721</td>
<td>0.248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Licensed Practical Nurse hours resident day</td>
<td>0.972</td>
<td>0.431</td>
<td>0.250</td>
<td>0.146</td>
<td>-3.721</td>
<td>0.248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Certified Nursing Assistant hours resident day</td>
<td>0.972</td>
<td>0.431</td>
<td>0.250</td>
<td>0.146</td>
<td>-3.721</td>
<td>0.248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share restrained</td>
<td>6.638</td>
<td>7.29</td>
<td>5.53</td>
<td>4.576</td>
<td>3.456</td>
<td>0.373</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page...
<table>
<thead>
<tr>
<th>Variable Names</th>
<th>Early (mean)</th>
<th>Early (sd)</th>
<th>Late (mean)</th>
<th>Late (sd)</th>
<th>tstat</th>
<th>Norm Diff</th>
</tr>
</thead>
</table>

Source: Long-Term Care: Facts on Care in the U.S. (2020), authors’ own calculation. “Early” refers to SNFs acquired before 2011 and “late” to those acquired starting in 2011. The sample is our main sample and a panel that is balanced on event time. It includes SNFs with observations for all relevant variables between 2000 and 2017 and that were acquired by a county hospital between 2005 and 2012. See main text for more details.
Table 2.A.3: Effects of 2003 Reform on Various SNF State-Level Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Markdown FMAP</th>
<th>Effective FMAP</th>
<th>Federal Exp</th>
<th>State Exp</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPL × Post</td>
<td>0.1904***</td>
<td>-15.7243***</td>
<td>-149.80***</td>
<td>-111.50***</td>
</tr>
<tr>
<td></td>
<td>(0.0453)</td>
<td>(3.6935)</td>
<td>(46.76)</td>
<td>(34.25)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.6361</td>
<td>0.7131</td>
<td>0.971</td>
<td>0.975</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SNF Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Direct Care Staff Ratio</th>
<th>RN Ratio</th>
<th>LPN Ratio</th>
<th>CNA Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPL × Post</td>
<td>-0.0041</td>
<td>-0.0106</td>
<td>-0.0021</td>
<td>-0.0084</td>
</tr>
<tr>
<td></td>
<td>(0.0576)</td>
<td>(0.0289)</td>
<td>(0.0260)</td>
<td>(0.0423)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.929</td>
<td>0.918</td>
<td>0.898</td>
<td>0.901</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>SNF Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Source: Long-Term Care: Facts on Care in the U.S. (2020); Centers for Medicare & Medicaid Services (2020a), authors’ calculation. Each column in each panel stands for one standard DD model with UPL flagging states with a pre-reform UPL scheme in place, and Post taking the value 1 for years after 2002. Federal and state expenditures are measured in million $. ***, **, and * = statistically different from zero at the 1%, 5%, and 10% level. Standard errors are clustered at the state level and reported in parentheses. All models have 248 state-year observations from 31 states. See main text for more details.
Table 2.A.4: Difference-in-Differences Models with Gardner 2sDD Correction

<table>
<thead>
<tr>
<th></th>
<th>NA or PA</th>
<th>Occupancy Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$D_c \times T_t$</td>
<td>0.1834***</td>
<td>0.1654***</td>
</tr>
<tr>
<td></td>
<td>(0.0530)</td>
<td>(0.0575)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.385</td>
<td>0.385</td>
</tr>
<tr>
<td>SNF Fixed Effects</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>SNF Controls</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Gardener (2021) 2SDD</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Medicaid Days</th>
<th>Alzheimer Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$D_c \times T_t$</td>
<td>1270*</td>
<td>1373**</td>
</tr>
<tr>
<td></td>
<td>(652)</td>
<td>(685)</td>
</tr>
<tr>
<td>Mean</td>
<td>18,591</td>
<td>18,591</td>
</tr>
<tr>
<td>SNF Fixed Effects</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>SNF Controls</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Gardener (2021) 2SDD</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

Source: Long-Term Care: Facts on Care in the U.S. (2020), authors' own calculation and illustration. Each column in each panel stands for one DD model similar to equation (2.11) but with just one post-reform dummy equaling one after $t=0$ and a control dummy for the implementation period at $t=0$. ***, **, and * = statistically different from zero at the 1%, 5%, and 10% level. Standard errors are clustered at the SNF level and in parentheses. All models have 2203 SNF-year observations. “NP or PA” measures whether the SNF has a nurse practitioner (NP) or a Physician Assistant (PA). See main text for details. Columns (3), (4), (7) and (8) implement two-stage DD models following Gardner (2021) and Butts and Gardner (2022).
### Table 2.A.5: Mechanisms—Alzheimer Patients in SNFs

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. All SNFs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_c \times T_t$</td>
<td>5.5700***</td>
<td>4.3992**</td>
<td>594</td>
<td>381</td>
<td>769*</td>
<td>665</td>
<td>-0.0079</td>
<td>-0.0133</td>
<td>3.9556**</td>
<td>3.2386*</td>
<td>0.1273</td>
<td>-0.1252</td>
</tr>
<tr>
<td></td>
<td>(1.8981)</td>
<td>(1.9331)</td>
<td>(506)</td>
<td>(470)</td>
<td>(433)</td>
<td>(420)</td>
<td>(0.0106)</td>
<td>(0.0114)</td>
<td>(1.6997)</td>
<td>(1.8080)</td>
<td>(0.3187)</td>
<td>(0.3356)</td>
</tr>
<tr>
<td>Mean</td>
<td>26.15</td>
<td>26.15</td>
<td>10956</td>
<td>10956</td>
<td>10047</td>
<td>10047</td>
<td>0.312</td>
<td>0.312</td>
<td>18.35</td>
<td>18.35</td>
<td>2.969</td>
<td>2.969</td>
</tr>
<tr>
<td><strong>B. Lower Quality SNFs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_c \times T_t$</td>
<td>5.9997**</td>
<td>4.1374</td>
<td>1,180**</td>
<td>890</td>
<td>1,135**</td>
<td>899</td>
<td>-0.0200</td>
<td>-0.0248</td>
<td>4.4423*</td>
<td>3.0658</td>
<td>-0.1584</td>
<td>-0.4737</td>
</tr>
<tr>
<td></td>
<td>(2.6670)</td>
<td>(2.7540)</td>
<td>(586)</td>
<td>(573)</td>
<td>(548)</td>
<td>(565)</td>
<td>(0.0190)</td>
<td>(0.0198)</td>
<td>(2.4321)</td>
<td>(2.5875)</td>
<td>(0.4141)</td>
<td>(0.4479)</td>
</tr>
<tr>
<td>Mean</td>
<td>20.69</td>
<td>20.69</td>
<td>7277</td>
<td>7277</td>
<td>6672</td>
<td>6672</td>
<td>0.325</td>
<td>0.325</td>
<td>15.48</td>
<td>15.48</td>
<td>2.076</td>
<td>2.076</td>
</tr>
<tr>
<td><strong>C. Higher Quality SNFs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_c \times T_t$</td>
<td>5.7663**</td>
<td>5.5825**</td>
<td>277</td>
<td>269</td>
<td>459</td>
<td>598</td>
<td>0.0007</td>
<td>-0.0049</td>
<td>3.7486</td>
<td>3.8684</td>
<td>0.5166</td>
<td>0.3566</td>
</tr>
<tr>
<td></td>
<td>(2.5353)</td>
<td>(2.5434)</td>
<td>(762)</td>
<td>(706)</td>
<td>(634)</td>
<td>(608)</td>
<td>(0.0125)</td>
<td>(0.0129)</td>
<td>(2.2741)</td>
<td>(2.4299)</td>
<td>(0.4423)</td>
<td>(0.4489)</td>
</tr>
<tr>
<td>Mean</td>
<td>30.11</td>
<td>30.11</td>
<td>13623</td>
<td>13623</td>
<td>12495</td>
<td>12495</td>
<td>0.303</td>
<td>0.303</td>
<td>20.43</td>
<td>20.43</td>
<td>3.616</td>
<td>3.616</td>
</tr>
<tr>
<td>SNF Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Gardner (2021) SDD</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

Source: MDS (2015), authors' own calculation and illustration. Each column in each panel stands for one DD model similar to equation (2.11) but with just one post-reform dummy equaling one after t=0 and a control dummy for the implementation period at t=0. ***, **, and * = statistically different from zero at the 1%, 5%, and 10% level. Standard errors are clustered at the assigned SNF level and in parentheses. Panel A considers all assigned SNFs, Panel B considers only SNFs without an Alzheimer unit and without a Nurse Practitioner (NP) or Physician Assistant (PA) prior to conversion. Panel C considers the remaining SNFs. All models in Panel A have 1882, in Panel B 791, and in Panel C 1091 observations. Columns (2), (4), (6), (8), (10) and (12) implement two-stage DD models following Gardner (2021) and Butts and Gardner (2022).
Table 2.A.6: Mechanisms—Hospital Discharges of Dementia Patients

<table>
<thead>
<tr>
<th></th>
<th>Focal SNF</th>
<th>Non-Ass. SNF</th>
<th>Any SNF</th>
<th>Any Lower Care</th>
<th>1-Year Mort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>A. All SNFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_c \times T_t$</td>
<td>0.0220***</td>
<td>0.0175**</td>
<td>-0.0158</td>
<td>-0.0281**</td>
<td>-0.0137</td>
</tr>
<tr>
<td></td>
<td>(0.0077)</td>
<td>(0.0084)</td>
<td>(0.0180)</td>
<td>(0.0138)</td>
<td>(0.0161)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0828</td>
<td>0.0828</td>
<td>0.326</td>
<td>0.326</td>
<td>0.328</td>
</tr>
<tr>
<td>B. Low. Qual. SNFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_c \times T_t$</td>
<td>0.0335**</td>
<td>0.0384***</td>
<td>0.0110</td>
<td>-0.0034</td>
<td>-0.0443**</td>
</tr>
<tr>
<td></td>
<td>(0.0128)</td>
<td>(0.0109)</td>
<td>(0.0164)</td>
<td>(0.0154)</td>
<td>(0.0226)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0719</td>
<td>0.0719</td>
<td>0.322</td>
<td>0.322</td>
<td>0.328</td>
</tr>
<tr>
<td>C. High Qual. SNFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_c \times T_t$</td>
<td>0.0170*</td>
<td>-0.0056</td>
<td>-0.0388</td>
<td>-0.0367</td>
<td>0.0105</td>
</tr>
<tr>
<td></td>
<td>(0.0099)</td>
<td>(0.0153)</td>
<td>(0.0266)</td>
<td>(0.0230)</td>
<td>(0.0210)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0871</td>
<td>0.0871</td>
<td>0.328</td>
<td>0.328</td>
<td>0.330</td>
</tr>
<tr>
<td>ZIP Code FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Age, race, gender</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Gardner 2SDD</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>

Source: MEDPAR (2015), authors’ own calculation and illustration. Each column in each panel stands for one DD model similar to equation (2.11) but with just one post-reform dummy equaling one after $t=0$ and a control dummy for the implementation period at $t=0$. ***, **, and * = statistically different from zero at the 1%, 5%, and 10% level. Standard errors are clustered at the assigned SNF level and in parentheses. Panel A considers all assigned SNFs, Panel B considers only SNFs without an Alzheimer unit and without a Nurse Practitioner (NP) or Physician Assistant (PA) prior to conversion. Panel C considers the remaining SNFs. All models in Panel A have 44,298, in Panel B 12,350, and in Panel C 31,948 observations. Columns (2), (4), (6), (8), and (10) implement two-stage DD models following Gardner (2021) and Butts and Gardner (2022).
2.B Principal Agent Conflict: County Hospitals and Indiana

As detailed in the main text, Indiana’s creative financing model allowed county hospitals to acquire SNFs. In this section, we allow for the possibility that the incentives between the hospital and the state are not fully aligned; for example, the hospital does not fully internalize the cost of financing for the state. Specifically, we assume that the hospital receives the nominal price $P$ from the state, which the hospital takes as given and chooses $\theta$ and $Q$ optimally according to

$$\max_{\theta, Q} B(Q, \theta) + (P - \theta) \times Q.$$  \hspace{1cm} (2.B.1)

We assume that the hospital internalizes the benefit of SNF health care provision $B(Q, \theta)$. It transfers its SNF profits $(P - \theta) \times Q$ in full to its hospital operations. The first order conditions are

$$B_{\theta}(\theta, Q)/Q - 1 = 0$$  \hspace{1cm} (2.B.2)

$$B_{Q}(\theta, Q)\left(1 + \frac{P}{\theta}\right) - 1 = 0.$$  \hspace{1cm} (2.B.3)

The quality first order condition is identical to the one in the main text. However, hospitals face even stronger incentives to increase quantity, as they do not internalize the state’s cost of funding $(1 - FMAP) \times P \times Q$. This reinforces the public moral hazard channel.

Before turning to the state’s optimal choice of $P$, we characterize the relationship between the hospital’s choice of $\theta, Q$ and $P$. Using the implicit function theorem, we find:

$$\frac{\partial Q}{\partial P} = -\frac{B_{\theta\theta}}{B_{QQ} \cdot B_{\theta\theta} - (B_{Q\theta} - 1)^2} = \Xi > 0$$  \hspace{1cm} (2.B.4)

$$\frac{\partial \theta}{\partial P} = -\frac{1 - B_{Q\theta}}{B_{QQ} \cdot B_{\theta\theta} - (B_{Q\theta} - 1)^2}.$$  \hspace{1cm} (2.B.5)
where the second order conditions imply that $B_{QQ} < 0$, $B_{\theta\theta} < 0$ and $B_{QQ}B_{\theta\theta}-(B_{Q\theta}-1)^2 > 0$, which denotes the determinant of the Hessian.

**State Problem:** States choose $P$ optimally according to:

$$\Pi(P) = \max_{P} \left( \max_{\theta, Q} \left( B(Q, \theta) + (P - \theta) \times Q \right) - (1 - FMAP) \cdot P \cdot Q \right). \quad (2.B.6)$$

The first order condition is

$$\frac{\partial \Pi}{\partial P} = FMAP \cdot Q - (1 - FMAP) \cdot P \cdot \frac{\partial Q}{\partial P}, \quad (2.B.7)$$

which implies

$$\frac{\partial \Pi}{\partial P} > 0 \quad \text{if} \quad \epsilon_{Q(P)} < \frac{FMAP}{1 - FMAP} \quad (2.B.8)$$

$$\frac{\partial \Pi}{\partial P} < 0 \quad \text{if} \quad \epsilon_{Q(P)} > \frac{FMAP}{1 - FMAP}. \quad (2.B.9)$$

Here, $\epsilon_{Q(P)}$ denotes the elasticity of $Q$, chosen by the hospital, with respect to $P$. That means that the state will choose $P = \bar{P}$ if the volume of care does not respond too elastically to changes in the nominal price: $\epsilon_{Q(P)} < 1$.

Sufficient conditions for this to be the case are that $B_{\theta\theta}$ and $B_{QQ}$ are constant in $\theta$ and $Q$. Consequently $\Xi$ is constant in $\theta$ and $Q$. To see why these conditions are sufficient, note that

$$\epsilon_{Q(P)} = \Xi \cdot \frac{P}{Q}, \quad (2.B.10)$$

with $\epsilon_{Q(P)}|_{P=0} = 0$. We now have

$$\frac{\partial \epsilon_{Q(P)}}{\partial P} = \Xi \cdot \frac{\Xi}{Q} \cdot (1 - \epsilon_{Q(P)}), \quad (2.B.11)$$

which implies that $\epsilon_{Q(P)}$ is bounded from above by 1. Note also that $FMAP > \frac{1}{2}$ for
all states, including Indiana, which implies that \( \frac{F_{MAP}}{1 - F_{MAP}} > 1 \). Hence, we have \( \epsilon_{Q(P)} \leq 1 < \frac{F_{MAP}}{1 - F_{MAP}} \) and \( \frac{\partial \Pi}{\partial P} > 0 \) such that \( P = \bar{P} \).
2.C  Digitized Audit Reports and CMS-64 Forms

One contribution of this paper is the systematic digitization and analysis of 24 audit reports from 18 states in the pre-reform era from 1997 to 2003. All reports were requested by the Centers for Medicare & Medicaid Services (CMS) and carried out by the Office of Audit Services (OAS) of the Office of Inspector General (OIG). OIG has the mandate to protect the integrity of the Department of HHS programs and the well-being of their beneficiaries (Department for Health and Human Services, 2020). The reports provide information on the amounts expropriated and transferred back to the state.

We then combine the transfer amounts with information on total nominal Medicaid spending (Total MHEX) separately for states (State MHEX) and the federal government (Total MHEX - State MHEX), from CMS-64 forms. States submit CMS-64 forms to the CMS Data Center (Centers for Medicare & Medicaid Services, 2020a). Finally, we construct the markdown and the effective (volume) FMAP as

\[ \mu = \frac{P - \theta}{P} = \frac{\text{Transfers}}{\text{Total MHEX}} \]  
\[ \text{Effective FMAP} = \text{FMAP} \cdot \frac{P}{\theta} = \text{FMAP} \cdot \frac{\text{Total MHEX}}{\text{Total MHEX} - \text{Transfers}} = \frac{\text{FMAP}}{1 - \mu} \]  

2.C.1 Evidence from Audit Reports

Figure 2.C.1a presents first estimates of the (negative) public markdown \( \frac{P - \theta}{P} \) based on the audit report information for a select set of states. Consistent with hypothesis H1, nominal prices are marked down in all audited state-years, ranging from a 6% mark-down in Alabama (2000) to a 38% mark-down in Pennsylvania (1999). In other words, the states expropriated

\[^{10}\text{The data are publicly available through the Medicaid Budget and Expenditure System/State Children’s Health Insurance Program Budget and Expenditure System (MBES/CBES).}\]
Figure 2.C.1: Public Markdowns and the Effective FMAP—Evidence from Audit Reports

Notes: Figure 2.C.1a presents the negative markdown, \(-\frac{\theta}{\theta_{\text{nom}}}\), which indicates the (negative) share of nominal SNF Medicaid funds diverted towards other uses. Figure 2.C.1b presents the nominal FMAP and the effective quantity FMAP: \(FMAP = \frac{\theta}{\theta_{\text{nom}}}\). The presented information was assembled from several audit reports: Department for Health and Human Services: Office of Inspector General (2001b,c,e, 2004c,b, 2005f,b,c,e,d, 2014); Centers for Medicare & Medicaid Services (2020a) using equations (2.C.1) and (2.C.2).
between 6 and 38% of nominal Medicaid funds and transferred those towards programs of higher perceived priority. The average markdown is 17.9%. In total, these states transferred $4.3bn back to the state budget ($360m per state and year) out of $17.2bn in total nominal spending. This corresponds to an average markdown of even 25% when weighted by nominal spending.

Figure 2.C.1b presents the analogues effective quantity FMAP estimates, $FMAP \times \frac{P}{\theta}$. When considering the redirection of funds, the effective quantity FMAP rates increase by between 4.3 percentage points in Alabama (2000) and 33.2 percentage points in Pennsylvania (1999). Figure 2.C.1b illustrates that multiple states systematically expropriated a significant share of nominal Medicaid spending.
2.D Hospital Financial Statements

To provide an estimate of the markdown on the acquired nursing homes, we use Marison County’s comprehensive annual HHC financial reports available between 2008 and 2019. As mentioned in Section 2.2.3, starting in 2003, HHC became the pioneer of nursing homes acquisitions in Indiana. In 2008 (2009), HHC owned the license of 29 (40) nursing homes and received more than 80% of UPL-SPs paid to SNFs in the state. By 2019, they held the license of 78 (out of the 450) nursing homes that had been acquired by a county hospital. We limit our analysis to the HHC of Marion county, as the other hospitals receiving UPL-SPs did not disclose the size of these transfers in their financial statements, nor how much of it was used for nursing home patients.

The HHC financial statements provide information on all these elements, which allows us to estimate the markdown of their Medicaid revenues. We approximate the regular Medicaid payments on their nursing home patients (Reg MHEX) from the revenue that HHC received from their long-term care division and the share of Medicaid patients. They disclose the size of UPL-SPs to nursing homes under “Medicaid Special Revenue” for their long-term division (Sup MHEX), while they disclose transfers as “Returns to the General Fund.”

We calculate the markdown and effective FMAP using:

\[
\mu = \frac{P - \theta}{P} = \frac{\text{Transfer}}{\text{RegMHEX} + \text{SupMHEX}} \quad (2.D.1)
\]

\[
\text{Effective FMAP} = \frac{\text{RegMHEX} + \text{SupMHEX}}{\text{RegMHEX} + \text{SupMHEX} - \text{Transfer}} = \frac{\text{FMAP}}{1 - \mu} \quad (2.D.2)
\]

---

11Between 2008 and 2019, none of the funds allocated to the General Fund were transferred back for use by the LTC division of HHC. The largest recipient of these funds (approximately 65%) is the inpatient hospital division under Wishard Health Services (before 2013) and Eskenazi Health Services (starting in 2013). The remainder the transferred funds was allocated to the Capital Projects and Debt Service Funds.
Estimates of equations (2.D.1) and (2.D.2) are provided in Figures 2.D.1a and 2.D.1b.
Figure 2.D.1: Public Markdowns and the Effective FMAP: Evidence from the HHC of Marion County

(a) Public Markdowns

\[ \text{FMAP} = \frac{F_{\text{MAP}}}{P} \]

(b) Effective vs. Nominal FMAP

Notes: Figure 2.D.1a presents the negative markdown, \(- \frac{P_{\theta}}{P}\), which indicates the (negative) share of nominal SNF Medicaid funds diverted towards other uses. Figure 2.D.1b presents the nominal FMAP and the effective quantity FMAP: \( FMAP \times \frac{P_{\theta}}{P} \). The estimates refer to SNFs operated by HHC of Marion County and build on the HHC’s financial reports as well as equations (2.D.1) and (2.D.2).
CHAPTER 3

Common Ownership and Distorted Portfolio Allocation: Evidence from Colombian Pension Funds

This chapter studies the distortions of common ownership on the portfolio allocation of financial entities. We use evidence from Colombian pension funds, which is a highly concentrated market, the largest entities are integrated with large financial conglomerates and manage funds that amount to roughly 25% of GDP and 87% of the market capitalization of the Colombian stock exchange, which potentially gives them market power in Colombian financial markets. We use a rich database on the daily portfolio positions of all pension funds, which is collected by the Financial Superintendence of Colombia. We conduct exercises that exploit variation from mergers and acquisitions of pension funds and other firms with assets in the Colombian financial system. The preliminary results suggest that pension funds disproportionately favor investments in related firms as merge firms more than double their share of commonly owned assets in their portfolios. This reduced-form evidence lays the building blocks to a structural analysis that allows decomposing the observed changes between ownership, market power, and changes in expectations.

3.1 Introduction

This paper studies whether financial institutions disproportionately favor assets that are issued by firms that are commonly owned, namely, part of the same conglomerate. We use evidence from the Colombian pension funds. This is an ideal setting to explore this as pension funds manage assets that amount to 25% of GDP, the largest pension funds are part of conglomerates whose firms issue assets that are traded in financial markets, and there are very rich datasets
covering the decisions of pension funds and savers. Moreover, we exploit changes in ownership from two mergers, which allows us to quantify the disproportionate investment in commonly owned assets.

Our main data source is Format 351 from the Financial Superintendence of Colombia that has daily information on the portfolio of all financial institutions at the asset and risk-adjusted portfolio level. This dataset also provides information on asset characteristics, like market value and price of acquisition, and on the characteristics of financial institutions and their portfolios. In this version we use the public version that only has information for pension funds on a monthly basis between 2005 and 2021. However, for extensions of this chapter we plan on using the proprietary, daily data. Additionally, we plan on complementing the portfolio data with transition matrices of savers between pension funds, and with detailed data on balance sheets and ownership of financial and non-financial corporations.

We use a difference-in-differences model to compare the share on commonly-owned assets of pension funds that merged with those that did not. That is, we exploit the variation in ownership coming from two mergers where two of the largest pension funds that are part of financial conglomerates acquire other pension funds. The first took place in December 2012 in which Proteccion, part of conglomerate Grupo Sura, acquired ING, while the second happened in December 2013 when Porvenir, part of conglomerate Grupo Aval, merged with Horizonte. By conducting the estimation on sub-samples that only include assets owned by the largest conglomerates in Colombia we test whether after the merger pension funds disproportionately favored assets issued by firms in their conglomerate.

The results show that effectively pension funds after the merger disproportionately favor commonly owned assets. The Porvenir-Horizonte merger resulted disproportionate increase of 175% in assets by Grupo Aval, without presenting any disproportionate changes in the holding of assets from other groups. Similarly, the Proteccion-ING merger resulted in a increase of 78% exclusively in the assets of Grupo Sura. By further decomposing the sample, we observe that changes occur mostly in low-liquidity stocks and in participations in joint and private capital
funds. Additionally, the changes in the share of commonly owned assets happened within the moderate- and high-risk portfolios, where pension funds have more regulatory leeway to adjust their investments.

We see this reduced-form analysis as the building block of the a structural analysis that allows to decompose the mechanisms explaining how ownership changes result in distortions to the portfolio allocation decisions of pension funds. We hypothesize that market power in the financial markets plays an important role in the decision of pension funds to favor commonly owned assets by distorting their price to favor the conglomerate. This hypothesis is informed by the fact that most of the changes in investment decisions occur in low-liquidity markets (on stocks and private funds) where there is more space to distort prices. Furthermore, this opens the question of the effects of this interaction between common ownership and market power on efficiency, risk-adjusted profitability, and financial stability, which we hope to address with complementary analyses and data.

We see this paper contributing to two main strands of literature. The first is related to the effect of common ownership of financial and real sector firms on their production and investment decisions. There is an extensive literature that argues that when financial entities hold participation in multiple firms, even when they are minority or lack control, in the same market there is reduced competition (Azar et al., 2018; Rubin, 2006; Rotemberg, 1984). However, little is known about how common ownership will impact the portfolio allocation decisions of financial firms in a conglomerate. We fill this gap by showing some stylized facts on how financial entities have incentives to disproportionately invest in assets from the conglomerate. We expect that this preliminary analysis serves as a building block to further disentangle the mechanism behind this behavior.

The second is related to the investment distortions that exist in the Colombian pension fund system. Jara et al. (2005) and Jara (2006) document that the portfolio of pension funds in Colombia is sub-optimal relative to the efficient return-risk portfolio frontier. They argue that the main factor explaining this result are the regulatory limits on investments on foreign
assets. However, in the counterfactual they argue that a large part of the inefficiency cannot be explained by these restrictions. G. (2008) argues that the poor institutional design of the agent-principal problem between pension funds and savers contributes to this inefficiency. We contribute to this literature by introducing common ownership as an additional source of distortion relative to the risk-adjusted return maximizing motive, which is highly relevant in a financial system like the Colombian which is highly concentrated and there are strong ties between firms within the same conglomerate.

3.2 Pension Funds in Colombia

In 1993 Colombia adopted a mixed pension fund system in which savers could choose between a defined-benefit system (Regimen de Prima Media) managed by a public institution (Colpensiones) and an subsidized individual savings system (Regimen de Ahorro Individual Subsidiado, and henceforth private system) managed by private pension funds (OECD et al., 2014). All workers in Colombia are required by law to save in one of these systems depending on their income. Workers with lower salaries contribute 4% of their monthly wage and their employers 12% of the monthly wage, while higher salaried workers contribute 4.5% and their employers 13.5%. In practice, however, only approximately 40% of the working age population regularly saves in pension funds. Out of the people who save, 71% of the affiliates save with the private system, which in 2021 amounted to roughly 17 million people (Azuero, 2020).

The private pension fund system manages a sizable share of financial assets and is highly concentrated. The private pension funds manage assets that by 2021 amounted roughly to 25.0% of GDP and 87% of the market capitalization of the stock exchange in Colombia, suggesting they could have sizable effects on prices with their investment decisions. Additionally, the management of these assets is highly concentrated on a few pension funds. In 2005, where our sample starts, there were six private pension funds: Porvenir (27% market share), Proteccion (22%), Horizonte (20%), ING (15%), Colfondos (15%), and Skandia (1%). In December
2012 Proteccion acquired ING and in December 2013 Porvenir acquired Horizonte. So, by
the end of our sample period in 2021 there were only four pension funds: Porvenir (59%),
Proteccion (29%), Colfondos (11%), and Skandia (1%).

The Colombian financial system as a whole is widely concentrated and dominated by
conglomerates that own the two largest pension funds. Porvenir is part of Grupo Aval, which
is composed of four banks that issued 25% of total credit, the largest trust fund by assets
managed, and multiple other financial and non-financial institutions that issue assets that are
traded in financial markets. Proteccion is part of grupo Sura, which is also comprised of the
largest bank in Colombia (with 26% of total credit), other financial institutions, and non-
financial corporations including the largest retailer (until 2010), wholesaler of construction
materials, and foodstuff manufacturer.

3.3 Data

The Colombian pension fund system provides an environment with very rich datasets, which
makes it suitable for the study of the distortions that arise in the presence of market power and
common ownership. This section describes the data used in this paper. Additionally it presents
the data available for extensions that will allow to further disentangle the effects of market
power and ownership from potential confounding factors, such as changes in expectations,
information flow between firms, risk aversion, and saver’s preferences.

3.3.1 Daily Portfolio Position

We use data from Format 351 compiled by the Financial Superintendence of Colombia, which
includes information on the investment portfolio of all financial institutions in Colombia. An
observation in this dataset is an asset paired with the portfolio of the financial institution that
holds the asset. This means that a single financial institution might hold the same asset in
different portfolios, which are divided according to their risk profile. The data is available on
a daily frequency starting in January 1, 2005 until March 31, 2022. It includes information on the following characteristics of the asset: market value on the cutoff date, the accounting code of the investments, value at which it was purchased, date of acquisition, type of asset, currency of denomination, name of the issuer, and the maturity date (if it is a fixed income asset). Additionally, it includes information on the name of the financial institution and the risk-profile of the portfolio.

This paper uses the public version of this data which only includes the portfolios of pension funds and the last day of the month between May 2006 and May 2021\(^1\). It includes six pension funds, which each has four portfolios: lowest risk (programmed retirement)\(^2\), low risk (conservative fund), moderate risk, and high risk. A day includes an average of 14,727 asset-portfolio observations for a total of 2,651,533 observations. Extensions of this paper will use the proprietary version of this dataset, which includes the portfolios of all institutions in the Colombian financial system and daily data.

### 3.3.2 Balance Sheet Data of Pension Funds

We use data of the balance sheets of pension funds to validate that the portfolio data is accurate. A potential caveat of the portfolio data is that it is built from self-reporting by individual entities, which means it is susceptible to measurement error or inaccurate reporting. To address this issue we compare the portfolio data with the end of the month balance sheet information that financial entities must submit to the Financial Superintendence of Colombia and are subject to additional auditing and supervision. The balance sheet data is at the asset type (i.e., 6-digit accounting unit) and portfolio of pension fund, which is a higher level of aggregation than the portfolio data. The comparison reveals that absolute difference between both datasets is less than 0.01% of the assets held by pension funds.

\(^1\)The data can be downloaded from https://cuestionpublica.com/aplicacion-portafolio-inversion-fondos-privados-pensiones/

\(^2\)By regulation it must only hold high liquidity assets to meet the redemptions of new retirees.
3.3.3 Ownership of Financial Institutions

The Financial Superintendence of Colombia provides a historical archive with all relevant changes to financial institutions in Colombia between 2000 and 2022. These changes include mergers, acquisitions, changes in name, divisions, bankruptcies, conversions, and liquidations. This dataset also provides information on the parties involved, the date in which entities filed for authorization, the date in which the merger was approved, and the date in which the merger was finalized. For the purpose of this preliminary analysis, we focus on the mergers between pension funds. However, future extensions of this analysis will also exploit the variation coming from changes in ownership and control from other financial institutions that issue assets in which pension hold investments.

We also use information on conglomerates by the Financial Superintendence of Colombia to determine common ownership. We focus on the largest three financial conglomerates (Grupo Aval, Grupo Sura Bancolombia, and Grupo Bolivar)—all of which have a pension fund, except for Grupo Bolivar—and on the other financial conglomerate that runs a pension fund (Grupo Skandia). We define that a firm is owned by one of these conglomerates if it is defined as such by the Financial Superintendence of Colombia, or if one of the entities in the group (excluding pension funds) owns more than 25% of equity or if it is one of the three top stock holders in the firm.

3.3.4 Ownership of Non-Financial Institutions

Non-financial institutions are not required to report changes to the Financial Superintendence of Colombia, so we complement the ownership data with the reports from the Emis records between 2000 and 2022. We use the same criteria that a firm is part of the financial conglomerate if one of the entities in the group (excluding pension funds) owns more than 25% of equity or if it is one of the three top stock holders in the firm. We check the information for the 332 domestic non-financial issuers that have assets in the portfolio of pension funds.
Additionally, we record the date in which there is a change in ownership and the sector in which the firm operates.

3.3.5 Transition Matrices of Pension Fund Savers

Extensions of this analysis will include the behavior of the savers and their reaction to changes in the returns and risk of pension funds. We argue that this margin has limited effect on the changes that we observe, as savers have a very high switching from bureaucratic limitations and information constraints. Nonetheless, in a structural demand analysis we have the monthly transition matrices of savers between portfolios within a pension fund, between pension funds, and out of the formal employment sector. This would allow to build market shares and their changes to estimate switching costs and quantify the preferences of savers of the characteristics of the portfolios of pension funds.

3.3.6 News Archives

We plan to use the news archives from the main financial media in Colombia (Primera Pagina, La Republica, Portafolio, El Tiempo, and El Espectador) to complement the data collected for this analysis. The idea is to use news reports on pension funds and issuers to find the dates in which the financial markets learned about possible changes of ownership. This would help to disentangle how much of the observed changes in the portfolio allocation on assets from commonly owned issuers responds to changes in expectations and how much is a direct effect of ownership. Additionally, these reports will give us information on unexpected shocks to issuers that could impact the value of the assets, and assess whether commonly pension funds respond differently to these shocks and if they use the market power to buffer part of these shocks.
3.4 Empirical Design

The reduced-form analysis in this paper intends to document the changes in the portfolio after changes in ownership on pension funds. To do so, we use the variation coming from two mergers between pension funds. The first took place in December 2012, when Proteccion, which is part of the conglomerate Sura-Bancolombia, acquired ING. The second was in December 2013 when Porvenir (part of conglomerate Aval) acquired Horizonte. For both cases, we build a two synthetic pension fund that aggregates the portfolios at the equivalent risk levels before the mergers. We compare the portfolios to that of non-merged pension funds using the following difference-in-differences specification:

\[ s_{aitj} = \beta_1 \text{Merged}_j + \beta_2 \text{Post}_t + \beta_3 \times \text{Merged}_j \times \text{Post}_t + \delta_i + \delta_j + \delta_\tau + \delta_t + \epsilon_{aitj}, \]  

(3.1)

where the dependent variable, \( s_{aitj} \), is the share that asset \( a \) from issuer \( i \) in the portfolio \( \tau \) managed by pension fund \( j \) in month \( t \). We include fixed effects for issuer (\( \delta_i \)), pension fund (\( \delta_j \)), risk-level of the portfolio (\( \delta_\tau \), and month (\( \delta_t \)). \( \text{Merged}_j \) is an indicator variable that equals one if pension fund \( j \) was part of the merger, and \( \text{Post}_t \) is an indicator variable that equals 1 starting on the month in which the merger takes place. The variable \( \epsilon_{ct} \) is a random unobserved variable. Standard errors are clustered at the pension fund level (\( j \)). \( \beta_3 \) is the parameter of interest that shows the disproportionate increase in the portfolios of merged firms on assets after the merger takes place.

The specification presented in Equation 3.1 does not necessarily allows us to establish a causal link of changes in ownership leading to a distortion in the the portfolio allocation of assets. The reason is that the merger might be altering other factors beyond ownership that could affect the investment decisions of pension funds or other investors, such as synergies or improved expectations about the performance of a conglomerate as a whole, information
flows between firms within the conglomerate, changes in the market power of the pension funds, or strategic responses from the savers. Instead, we see this reduced-form specification as a preliminary analysis to establish some stylized facts on investment decisions after changes in ownership, which would then serve as building blocks to a structural analysis that allows disentangling ownership from the confounders mentioned above.

3.5 Reduced-Form Evidence on Portfolio Allocation Decisions after Changes in Ownership

To study the changes in the portfolio allocation decisions we estimate Equation 3.1 for both mergers separately and vary the sample of the analysis to study whether the allocation decisions change differently accordingly to ownership of the assets. Besides the full sample, we restrict the analysis to the assets owned by Grupo Aval, Grupo Sura, Grupo Bolivar, and Grupo Skandia. The idea is to test whether the increased assets coming from the absorbed pension fund are disproportionately invested in assets that are issued by firms within the same conglomerate.

The results show that pension funds disproportionately increase in their portfolios the share of assets that are commonly owned. After Horizonte merged with Porvenir, the average share of individual assets that are owned by grupo Aval more than doubled from their joined share before the merger. That is, an increase of 0.07 percentage points (p.p.) from a pre-merger average of 0.04% (Table 3.1, Column 2). Meanwhile, the changes on the holding of assets from all other conglomerates show no statistically significant changes after the merger (Table 3.1, Columns 3-5). Similarly, after ING merged with Proteccion the average share of commonly owned assets increase by 78% (0.12 p.p. increase from 0.16%). In contrast, their share in assets from Aval, Bolivar and Skandia did not show statistically significant changes (Table 3.2, Columns 2-5).

To delve into the potential mechanisms behind the disproportionate increase in the share of
Table 3.1: Change in Portfolio Shares after the Merger between Povenir and Horizonte (Grupo Aval)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Post</strong></td>
<td>-0.1796</td>
<td>0.03284</td>
<td>0.6073</td>
<td>-0.1917</td>
<td>0.08865</td>
</tr>
<tr>
<td></td>
<td>(0.1205)</td>
<td>(0.04810)</td>
<td>(0.3984)</td>
<td>(1.8594)</td>
<td>(0.1424)</td>
</tr>
<tr>
<td><strong>Merged</strong></td>
<td>-0.2468***</td>
<td>-0.05541***</td>
<td>-0.08849**</td>
<td>-0.3399*</td>
<td>0.07199</td>
</tr>
<tr>
<td></td>
<td>(0.03503)</td>
<td>(0.004701)</td>
<td>(0.03719)</td>
<td>(0.2007)</td>
<td>(0.05611)</td>
</tr>
<tr>
<td><strong>Post x Merged</strong></td>
<td>0.2427</td>
<td>0.06825***</td>
<td>0.06675</td>
<td>0.6036</td>
<td>-0.06601</td>
</tr>
<tr>
<td></td>
<td>(0.3808)</td>
<td>(0.005214)</td>
<td>(0.04173)</td>
<td>(0.6165)</td>
<td>(0.04456)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
<th>R-squared</th>
<th>Sample</th>
<th>Mean Pre-merger</th>
<th>Issuer FE</th>
<th>Pension Fund FE</th>
<th>Portfolio FE</th>
<th>Month FE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>192,369</td>
<td>0.088</td>
<td>All</td>
<td>0.295</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>53,372</td>
<td>0.097</td>
<td>Aval</td>
<td>0.042</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>48,029</td>
<td>0.021</td>
<td>Sura</td>
<td>0.140</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>7,533</td>
<td>0.122</td>
<td>Bolivar</td>
<td>0.529</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>907</td>
<td>0.482</td>
<td>Skandia</td>
<td>0.001</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Financial Superintendence of Colombia. OLS estimation of equation 3.1 using portfolio share as the outcome. Merged is equal to 1 if the pension fund managing the portfolio is Horizonte or Porvenir. Treat is 1 after December 2013 when both entities merged. Standard errors are clustered at the pension fund level.
Table 3.2: Change in Portfolio Shares after the Merger between Proteccion and ING (Grupo Sura)

<table>
<thead>
<tr>
<th>Portfolio Share</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>-0.09728</td>
<td>0.002718</td>
<td>0.5112</td>
<td>-0.9174</td>
<td>0.03619</td>
</tr>
<tr>
<td></td>
<td>(0.1205)</td>
<td>(0.04831)</td>
<td>(0.4002)</td>
<td>(1.8676)</td>
<td>(0.1388)</td>
</tr>
<tr>
<td>Merged</td>
<td>0.08401**</td>
<td>-0.04780***</td>
<td>-0.08264*</td>
<td>-0.7564***</td>
<td>0.01523</td>
</tr>
<tr>
<td></td>
<td>(0.04128)</td>
<td>(0.004632)</td>
<td>(0.04510)</td>
<td>(0.2155)</td>
<td>(0.07319)</td>
</tr>
<tr>
<td>Post x Merged</td>
<td>-0.08906</td>
<td>0.04306</td>
<td>0.1283***</td>
<td>0.9358</td>
<td>0.08025</td>
</tr>
<tr>
<td></td>
<td>(0.08289)</td>
<td>(0.04987)</td>
<td>(0.04845)</td>
<td>(0.8234)</td>
<td>(0.07325)</td>
</tr>
</tbody>
</table>

| Observations    | 192,369  | 53,372   | 48,029   | 7,533    | 907      |
| R-squared       | 0.088    | 0.095    | 0.021    | 0.123    | 0.481    |
| Sample          | All      | Aval     | Sura     | Bolivar  | Skandia  |
| Mean Pre-merger | 0.324    | 0.040    | 0.164    | 0.615    | 0.001    |
| Issuer FE       | YES      | YES      | YES      | YES      | YES      |
| Pension Fund FE | YES      | YES      | YES      | YES      | YES      |
| Portfolio FE    | YES      | YES      | YES      | YES      | YES      |
| Month FE        | YES      | YES      | YES      | YES      | YES      |

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Financial Superintendence of Colombia. OLS estimation of equation 3.1 using portfolio share as the outcome. Merged is equal to 1 if the pension fund managing the portfolio is ING or Proteccion. Treat is 1 after December 2012 when both entities merged. Standard errors are clustered at the pension fund level.
Table 3.3: Change in Portfolio Shares after the Merger between Povenir and Horizonte (Grupo Aval), by Type of Asset

<table>
<thead>
<tr>
<th></th>
<th>Panel A: High-liquidity Stocks</th>
<th>Panel B: Low-liquidity Stocks</th>
<th>Panel C: Fixed Income</th>
<th>Panel D: Joint Funds and Private Capital Funds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td><strong>Post</strong></td>
<td>0.3023</td>
<td>-0.08673**</td>
<td>-0.2224</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3725)</td>
<td>(0.04403)</td>
<td>(0.4535)</td>
<td></td>
</tr>
<tr>
<td><strong>Merged</strong></td>
<td>-0.2923***</td>
<td>-0.02087**</td>
<td>-0.03581</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05809)</td>
<td>(0.009755)</td>
<td>(0.07503)</td>
<td></td>
</tr>
<tr>
<td><strong>Post × Merged</strong></td>
<td>0.2085</td>
<td>0.01609</td>
<td>0.03308</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.6504)</td>
<td>(0.01023)</td>
<td>(0.08476)</td>
<td></td>
</tr>
<tr>
<td><strong>Panel D: Joint Funds and Private Capital Funds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Post</strong></td>
<td>0.4193</td>
<td>0.05693</td>
<td>6.4012**</td>
<td>0.2505</td>
</tr>
<tr>
<td></td>
<td>(0.5828)</td>
<td>(0.1151)</td>
<td>(2.5918)</td>
<td>(0.8111)</td>
</tr>
<tr>
<td><strong>Merged</strong></td>
<td>-0.1918***</td>
<td>0.005223</td>
<td>-0.3154</td>
<td>0.9839***</td>
</tr>
<tr>
<td></td>
<td>(0.05222)</td>
<td>(0.01280)</td>
<td>(0.2945)</td>
<td>(0.07033)</td>
</tr>
<tr>
<td><strong>Post × Merged</strong></td>
<td>-0.1114**</td>
<td>0.03268**</td>
<td>-0.3631</td>
<td>-0.9713***</td>
</tr>
<tr>
<td></td>
<td>(0.05505)</td>
<td>(0.01456)</td>
<td>(0.3488)</td>
<td>(0.08134)</td>
</tr>
</tbody>
</table>

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Financial Superintendency of Colombia. OLS estimation of equation 3.1 using portfolio share as the outcome. Merged is equal to 1 if the pension fund managing the portfolio is Horizonte or Povenir. Treat is 1 after December 2013 when both entities merged. The type of assets is based on the definition of the original dataset from the Financial Superintendency of Colombia. Standard errors are clustered at the pension fund level.
Table 3.4: Change in Portfolio Shares after the Merger between Proteccion and ING (Grupo Sura), by Type of Asset

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: High-liquidity Stocks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>0.3864</td>
<td>-0.08508*</td>
<td>-0.2188</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3725)</td>
<td>(0.04410)</td>
<td>(0.4534)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>-0.04516</td>
<td>-0.009441</td>
<td>-0.04216</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.06507)</td>
<td>(0.008033)</td>
<td>(0.08517)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post × Merged</td>
<td>-0.06924</td>
<td>0.004458</td>
<td>0.02506</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.07021)</td>
<td>(0.008578)</td>
<td>(0.09291)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Low-liquidity Stocks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>-0.01324</td>
<td>-0.008618</td>
<td>1.816e-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02897)</td>
<td>(0.03537)</td>
<td>(0.00149)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>-0.003394</td>
<td>-0.01552***</td>
<td>-0.001372*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004313)</td>
<td>(0.004562)</td>
<td>(0.00071)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post × Merged</td>
<td>0.009202</td>
<td>0.02323</td>
<td>0.001269***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007862)</td>
<td>(0.025044)</td>
<td>(0.00039)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel C: Fixed Income</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>-2.6151***</td>
<td>-0.001110</td>
<td>-0.1603</td>
<td>0.1601**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.7243)</td>
<td>(0.1699)</td>
<td>(0.3820)</td>
<td>(0.07468)</td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>-0.08584</td>
<td>-0.1154***</td>
<td>-0.06261***</td>
<td>0.03079</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.06069)</td>
<td>(0.007330)</td>
<td>(0.02137)</td>
<td>(0.02268)</td>
<td></td>
</tr>
<tr>
<td>Post × Merged</td>
<td>0.1866</td>
<td>0.1074</td>
<td>0.1101</td>
<td>-0.06114***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.06502)</td>
<td>(0.07960)</td>
<td>(0.12216)</td>
<td>(0.02341)</td>
<td></td>
</tr>
<tr>
<td><strong>Panel D: Joint Funds and Private Capital Funds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>0.4853</td>
<td>0.1056</td>
<td>5.8723**</td>
<td>-0.2592</td>
<td>0.03619</td>
</tr>
<tr>
<td></td>
<td>(0.5844)</td>
<td>(0.1156)</td>
<td>(2.6022)</td>
<td>(0.8266)</td>
<td>(0.1388)</td>
</tr>
<tr>
<td>Merged</td>
<td>-0.1580***</td>
<td>0.01286</td>
<td>-0.3397</td>
<td>-0.3578***</td>
<td>0.01523</td>
</tr>
<tr>
<td></td>
<td>(0.05188)</td>
<td>(0.01395)</td>
<td>(0.2724)</td>
<td>(0.07486)</td>
<td>(0.07319)</td>
</tr>
<tr>
<td>Post × Merged</td>
<td>-0.08709</td>
<td>-0.05596***</td>
<td>0.6381**</td>
<td>0.4432***</td>
<td>0.08025</td>
</tr>
<tr>
<td></td>
<td>(0.05408)</td>
<td>(0.01502)</td>
<td>(0.3061)</td>
<td>(0.07779)</td>
<td>(0.07325)</td>
</tr>
</tbody>
</table>

Sample: All, Aval, Sura, Bolivar, Skandia
Mean Pre-merger: 0.324, 0.040, 0.164, 0.615, 0.001
Issuer FE: YES, YES, YES, YES, YES
Pension Fund FE: YES, YES, YES, YES, YES
Portfolio FE: YES, YES, YES, YES, YES
Month FE: YES, YES, YES, YES, YES

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Financial Superintendence of Colombia. OLS estimation of equation 3.1 using portfolio share as the outcome. Merged is equal to 1 if the pension fund managing the portfolio is ING or Proteccion. Treat is 1 after December 2012 when both entities merged. The type of assets is based on the definition of the original dataset from the Financial Superintendence of Colombia. Standard errors are clustered at the pension fund level.
commonly owned assets we conduct two additional exercises in which we estimate Equation 3.1 in further sub-samples. The first exercise divides assets by type according to the information of the Financial Superintendence of Colombia into: high-liquidity stocks, low-liquidity stocks, fixed income, and joint and private capital funds. The second exercise divides the data by type of portfolio according to their risk profile based on the tiers required by regulation.

The analysis by assets reveals a differentiated behavior by type of asset. In the merger between Porvenir and Horizonte, there is an increase in the share of assets issued by Grupo Aval for low-liquidity stocks (105%), fixed income (22%), and for joint and private capital funds (94%) (Table 3.3, Panels B-D, Column 2). The increase in the share in these assets is exclusively present for assets issued by Aval. Furthermore, for joint and private capital funds there seems to be substitution away from assets from other issuers, mostly from Grupo Bolivar, into those issued by Aval (Table 3.3, Panel D, Column 1, 2, and 4). In the merger between Proteccion and ING, the increase in the assets own by Grupo Sura occur on low-liquidity stocks (62%) and in the joint and private capital funds (121%) (Table 3.3, Panels B and D, Column 3).

Dividing the sample by portfolio type, we observe that for both mergers the changes in the share of commonly owned assets after the merger happened only on the more risky portfolios, which are the ones that give more leeway to pension funds in their investment decisions. In the merger between Porvenir and Horizonte, the disproportionate increase of the share of commonly owned assets (i.e., those owned by Grupo Aval) is observed only for high- and moderate-risk portfolios, which rose by 22% and 106%, respectively (Table 3.5, Panels A and B, Column 2). Similarly, the merger between Proteccion and ING resulted in disproportionate increases of the assets issued by Grupo Sura of 128% in the high-risk portfolio and of 97% in the moderate-risk portfolio (Table 3.6, Panels A and B, Column 3). As in the previous exercises, the share of assets issued by other conglomerates did not disproportionately vary after the merger.
Table 3.5: Change in Portfolio Shares after the Merger between Povenir and Horizonte (Grupo Aval), by Type of Portfolio

<table>
<thead>
<tr>
<th>Panel A: High-Risk Portfolio</th>
<th>Portfolio Share (percentage)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>0.1303***</td>
<td>0.01422**</td>
<td>0.7066***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03860)</td>
<td>(0.006986)</td>
<td>(0.2234)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>0.01597</td>
<td>3.157e-04</td>
<td>0.05201</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01362)</td>
<td>(0.002378)</td>
<td>(0.07147)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post × Merged</td>
<td>0.004616</td>
<td>0.007211***</td>
<td>-0.04096</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01410)</td>
<td>(0.002456)</td>
<td>(0.07584)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B: Moderate-Risk Portfolio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>-0.6702</td>
<td>0.008593</td>
<td>1.3621**</td>
<td>-0.4201</td>
<td>0.06991</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.4270)</td>
<td>(0.06863)</td>
<td>(0.6255)</td>
<td>(2.4492)</td>
<td>(0.1655)</td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>-0.3618***</td>
<td>-0.06253***</td>
<td>-0.1214*</td>
<td>-0.2630</td>
<td>-0.01016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.04930)</td>
<td>(0.007887)</td>
<td>(0.07249)</td>
<td>(0.2860)</td>
<td>(0.06568)</td>
<td></td>
</tr>
<tr>
<td>Post × Merged</td>
<td>0.3438</td>
<td>0.09649***</td>
<td>0.05302</td>
<td>0.6951</td>
<td>-0.003329</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.5804)</td>
<td>(0.009468)</td>
<td>(0.08919)</td>
<td>(0.6220)</td>
<td>(0.05387)</td>
<td></td>
</tr>
<tr>
<td>Panel C: Conservative Portfolio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>0.007017</td>
<td>0.01215</td>
<td>-0.002602</td>
<td>0.1394</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03104)</td>
<td>(0.02137)</td>
<td>(0.03092)</td>
<td>(0.2179)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>-0.01981*</td>
<td>-0.01935**</td>
<td>-0.02567***</td>
<td>-0.04137</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01026)</td>
<td>(0.007911)</td>
<td>(0.009448)</td>
<td>(0.05947)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post × Merged</td>
<td>0.07507</td>
<td>0.04151</td>
<td>0.07854</td>
<td>0.3130</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.08083)</td>
<td>(0.08366)</td>
<td>(0.07010)</td>
<td>(0.36196)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel D: Portfolio for Programmed Retirement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>-0.1017***</td>
<td>0.01487</td>
<td>0.09138***</td>
<td>0.04003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02954)</td>
<td>(0.01788)</td>
<td>(0.03412)</td>
<td>(0.07574)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>-0.07772***</td>
<td>-0.02701***</td>
<td>-0.04266***</td>
<td>-0.004709</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009703)</td>
<td>(0.006066)</td>
<td>(0.01022)</td>
<td>(0.02612)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post × Merged</td>
<td>0.06897</td>
<td>0.004475</td>
<td>0.02673</td>
<td>-0.005687</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05021)</td>
<td>(0.006421)</td>
<td>(0.02089)</td>
<td>(0.02594)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sample | All | Aval | Sura | Bolivar | Skandia |
Issuer FE | YES | YES | YES | YES | YES |
Pension Fund FE | YES | YES | YES | YES | YES |
Month FE | YES | YES | YES | YES | YES |

Notes: ***p<0.01, **p<0.05, *p<0.1. Data from the Financial Superintendence of Colombia. OLS estimation of equation 3.1 using portfolio share as the outcome. Merged is equal to 1 if the pension fund managing the portfolio is Horizonte or Porvenir. Treat is 1 after December 2013 when both entities merged. The type of portfolio is based on the regulation from the Financial Superintendence of Colombia. Standard errors are clustered at the pension fund level.
Table 3.6: Change in Portfolio Shares after the Merger between Proteccion and ING (Grupo Sura), by Type of Portfolio

<table>
<thead>
<tr>
<th>Portfolio Share (percentage)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: High-Risk Portfolio</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>0.1243***</td>
<td>0.01688**</td>
<td>0.6588***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03876)</td>
<td>(0.007015)</td>
<td>(0.2240)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>0.009655</td>
<td>0.01003***</td>
<td>-0.02189</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01622)</td>
<td>(0.002617)</td>
<td>(0.09275)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post × Merged</td>
<td>0.02546</td>
<td>-0.004843*</td>
<td>0.1721***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01646)</td>
<td>(0.002661)</td>
<td>(0.04633)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Moderate-Risk Portfolio</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>-0.7497*</td>
<td>-0.001391</td>
<td>1.2537**</td>
<td>-1.0688</td>
<td>0.06603</td>
</tr>
<tr>
<td></td>
<td>(0.4290)</td>
<td>(0.06904)</td>
<td>(0.6287)</td>
<td>(2.4602)</td>
<td>(0.1602)</td>
</tr>
<tr>
<td>Merged</td>
<td>-0.2340***</td>
<td>-0.04779***</td>
<td>-0.08254</td>
<td>-0.6617**</td>
<td>0.007979</td>
</tr>
<tr>
<td></td>
<td>(0.05244)</td>
<td>(0.007682)</td>
<td>(0.08356)</td>
<td>(0.3076)</td>
<td>(0.08533)</td>
</tr>
<tr>
<td>Post × Merged</td>
<td>0.1777</td>
<td>0.02855</td>
<td>0.1712***</td>
<td>0.9631</td>
<td>0.01944</td>
</tr>
<tr>
<td></td>
<td>(0.15942)</td>
<td>(0.08838)</td>
<td>(0.06685)</td>
<td>(0.8240)</td>
<td>(0.08580)</td>
</tr>
<tr>
<td><strong>Panel C: Conservative Portfolio</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>0.02248</td>
<td>0.02088</td>
<td>0.02249</td>
<td>0.2227</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03109)</td>
<td>(0.02150)</td>
<td>(0.03094)</td>
<td>(0.2308)</td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>0.01166</td>
<td>-0.003736</td>
<td>-0.004384</td>
<td>0.06435</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01237)</td>
<td>(0.007604)</td>
<td>(0.01307)</td>
<td>(0.06779)</td>
<td></td>
</tr>
<tr>
<td>Post × Merged</td>
<td>0.01862</td>
<td>0.001041</td>
<td>-0.00087</td>
<td>-0.07928</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01269)</td>
<td>(0.007913)</td>
<td>(0.01348)</td>
<td>(0.06982)</td>
<td></td>
</tr>
<tr>
<td><strong>Panel D: Portfolio for Programmed Retirement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>-0.05458*</td>
<td>0.01253</td>
<td>0.09558***</td>
<td>0.03449</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02959)</td>
<td>(0.01798)</td>
<td>(0.03409)</td>
<td>(0.08665)</td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>0.1133***</td>
<td>-0.007578</td>
<td>-0.03612***</td>
<td>-0.01025</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01111)</td>
<td>(0.006296)</td>
<td>(0.01305)</td>
<td>(0.04278)</td>
<td></td>
</tr>
<tr>
<td>Post × Merged</td>
<td>-0.1020***</td>
<td>0.009158</td>
<td>0.01879</td>
<td>0.005413</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01133)</td>
<td>(0.006487)</td>
<td>(0.01346)</td>
<td>(0.04260)</td>
<td></td>
</tr>
</tbody>
</table>

**Sample:** All, Aval, Sura, Bolivar, Skandia

**Issuer FE:** YES, YES, YES, YES, YES

**Pension Fund FE:** YES, YES, YES, YES, YES

**Month FE:** YES, YES, YES, YES, YES

**Notes:** ***p<0.01, **p<0.05, *p<0.1. Data from the Financial Superintendence of Colombia. OLS estimation of equation 3.1 using portfolio share as the outcome. Merged is equal to 1 if the pension fund managing the portfolio is ING or Proteccion. Treat is 1 after December 2012 when both entities merged. The type of portfolio is based on the regulation from the Financial Superintendence of Colombia. Standard errors are clustered at the pension fund level.
3.6 Conclusion and Discussion

This chapter studies the distortions of common ownership on the portfolio allocation of financial entities. Using evidence from pension funds in Colombia, we exploit the variation in ownership coming from two mergers where two of the largest pension funds that are part of financial conglomerates acquire other pension funds. We use a difference-in-differences specification to test whether after the merger the pension funds disproportionately increased their holdings of assets issued by other firms in their financial conglomerate. The results show that this is effectively the case as the Porvenir-Horizonte merger resulted disproportionate increase of 175% in assets by Grupo Aval and the Proteccion-ING merger resulted in a disproportionate increase of 78% in assets by Grupo Sura. The changes occur mostly in low-liquidity stocks and in participations in joint and private capital funds. Additionally, the changes in the share of commonly owned assets happened within the moderate- and high-risk portfolios.

Although we cannot claim that the changes we observe after the merger in the share of commonly owned assets are caused exclusively by the change in ownership, our results suggest that common ownership is an important factor explaining the observed behavior. The reason is that the main potential confounding factor is not supported by our results. If the increased investment in assets results from a positive externality of the merger leading into a strengthened conglomerate, we should see the demand from all agents for these assets to increase. However, the results show that only the merged firm disproportionately increased the participation in these assets.

Furthermore, we see this reduced-form analysis as the building block of the a structural analysis that allows to decompose the mechanisms explaining how ownership changes result in distortions to the portfolio allocation decisions of pension funds. We hypothesize that market power in the financial markets plays an important role in the decision of pension funds to favor commonly owned assets by distorting their price to favor the conglomerate. This hypothesis is informed by the fact that most of the changes in investment decisions occur
in low-liquidity markets (on stocks and private funds) where there is more space to distort prices. Furthermore, this opens the question of the effects of this interaction between common ownership and market power on efficiency, risk-adjusted profitability, and financial stability, which we hope to address with complementary analyses and data.
Bibliography


Department for Health and Human Services: Office of Inspector General (2001a). *Review of Medicaid Enhanced Payments to Public Hospital To Providers and the Use of Intergovernmental Transfers by the Alabama State Medicaid Agency (A-04-00-02 169).*

Department for Health and Human Services: Office of Inspector General (2001b). *Review of Medicaid Enhanced Payments to Public Providers and the Use of Intergovernmental Transfers by the Alabama State Medicaid Agency (A-04-00-02165).*
Department for Health and Human Services: Office of Inspector General (2001c). *Review of Medicaid Enhanced Payments to Public Providers and the Use of Intergovernmental Transfers by the State of Nebraska (A-07-00-02076).*


Department for Health and Human Services: Office of Inspector General (2001g). *Review of the Commonwealth of Pennsylvania’s Use of Intergovernmental Transfers to Finance Medicaid Supplementation Payments to County Nursing Facilities (A-03-00-00203).*


Federal Register (2018). Federal financial participation in state assistance expenditures; federal matching shares for Medicaid, the children’s health insurance program, and aid to needy aged, blind, or disabled persons for october 1, 2019 through september 30, 2020. Federal Register, 83(229):”61157–61160”.

Federal Register (2019). Medicaid program; medicaid fiscal accountability regulation,


First Regular Session 112th General Assembly (2001). *Senate Enrolled Act No. 309.*


Poray, M. C., Gray, A. W., and Boehlje, M. (2002). Evaluation of Alternative Coordination Systems Between Producers and Packers in the Pork Value Chain. Staff papers, Purdue University, Department of Agricultural Economics.


Spanko, A. (2020). *CMS to Retire Old Nursing Home Compare Site on December*


Transportation Division (1964). *Estimated Costs for the Motor Vehicle Transportation of*


U.S. Department of Health and Human Services (2001). Medicaid program; revision to medicaid upper payment limit requirements to hospital services, nursing facility services, intermediate care facility services for the mentally retarded, and clinic services. *Federal Register*, 66(9):3148.


