

ESSAYS IN PUBLIC FINANCE

by

CHRISTIAN D IMBODEN

A DISSERTATION

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

June 2019

DISSERTATION APPROVAL PAGE

Student: Christian D Imboden

Title: Essays in Public Finance

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

Caroline Weber	Chair
Van Kolpin	Co-Chair
Jeremy Piger	Core Member
David Guenther	Institutional Representative

and

Janet Woodruff-Borden	Vice Provost and Dean of the Graduate School
-----------------------	--

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

Degree awarded June 2019

© 2019 Christian D Imboden

DISSERTATION ABSTRACT

Christian D Imboden

Doctor of Philosophy

Department of Economics

June 2019

Title: Essays in Public Finance

This dissertation deals with important issues in the field of public finance, namely, how we raise government revenue via taxation and how we spend it in the form of public goods. The first substantive chapter examines the incidence of corporate income taxes on the owners of corporate capital, the shareholders. By allowing stock markets to value the future impacts of corporate income tax changes, I am able to estimate their incidence on shareholders using changes in stock prices around changes in state-level corporate income tax rates. Estimates are generally statistically insignificant for tax decreases, but stock prices respond to tax increases with an approximately ten percent decline for each percent of tax increase. In the next chapter, co-authors and I examine income reporting using tax and survey data. As survey income data are frequently used by economists, it is imperative that incomes are measured as accurately as possible. We find that there are systematic differences in how individual respondents report their wage income to the Current Population Survey versus the Internal Revenue Service that are related to demographic characteristics, including age and educational attainment. However, there is great heterogeneity in misreporting within groups.

In the final substantive chapter, I examine strategic interactions between different levels of government. I find that there are theoretical cases where a subordinate level of government would benefit from sabotaging the plans of a dominant level of government. I also find that there are theoretical cases where competition between levels of government can be welfare-improving for citizens of the local government.

This dissertation includes previously unpublished co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Christian D Imboden

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
State University of New York at Albany, Albany, NY
New York University, New York, NY
University of Michigan, Ann Arbor, MI

DEGREES AWARDED:

Doctor of Philosophy, Economics, 2019, University of Oregon
Master of Science, Economics, 2015, University of Oregon
Master of Science, Accounting, 2004, State University of New York at Albany
Master of Fine Arts, Musical Theatre Writing, 2002, New York University
Bachelor of Fine Arts, Jazz and Contemporary Improvisation, 2000,
University of Michigan

AREAS OF SPECIAL INTEREST:

Public Economics
Labor Economics

GRANTS, AWARDS AND HONORS:

Gerlof Homan Award in International Economics, University of Oregon, 2016
Graduate Teaching Fellowship, University of Oregon, 2014-2019

ACKNOWLEDGEMENTS

I would like to thank the members of my committee for their helpful input. I would also like to thank Glen Waddell for early comments, and thank other faculty and students in the Department of Economics for their suggestions about earlier iterations of these chapters. I offer a special thanks to John Voorheis at the US Census Bureau for his assistance. The statistical summaries reported in this dissertation have been cleared by the Census Bureaus Disclosure Review Board, release authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407. Wharton Research Data Services (WRDS) was used in preparing this dissertation. This service and the data available thereon constitute valuable intellectual property and trade secrets of WRDS and/or its third-party suppliers.

To Ruth Johnson Colvin, who showed me that you are never too old to change your
life and the lives of others.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
II. DO STOCK PRICES RESPOND TO CHANGES IN CORPORATE INCOME TAX RATES?	4
Introduction	4
Related Literature	9
Data and Institutional Details	14
Methodology and Results	25
Conclusion	44
III. MEASURING SYSTEMATIC WAGE MISREPORTING BY DEMOGRAPHIC GROUPS	46
Introduction	46
Data	50
Identification and Model Specification	57
Results	59
Discussion	83
Conclusion	89

Chapter	Page
IV. GAMES OF FISCAL MANAGEMENT	91
Introduction and Related Literature	91
Games of Sabotage	98
Bidding Bureaucrats	111
Conclusion	116
V. CONCLUSION	117
APPENDIX:	119
Corporate Income Tax Data Collection	119
Estimating the Growth Rate and Rate of Return	120
Comparing Tax Law Changes	122
Event Studies	124
Cleaning of Matched CPS/Administrative Data	136
Variable Construction	136
Payoff Determination in Sabotage Games	138
REFERENCES CITED	144

LIST OF FIGURES

Figure	Page
1. Kernel Density Plots of Wage Reporting Differentials by Race and Ethnicity	66
2. Kernel Density Plots of Wage Reporting Differentials by Education Level	67
3. Wage Reporting Differentials Across the Wage distribution, by Education Level and Race	69
4. Basic Sequential Game, Generic Payoffs	101
5. Basic Sequential Game, Specific Payoffs	101
6. Reversed Order Sequential Game, Specific Payoffs	102
7. Simultaneous Games, Standard and Modified Payoffs	103
8. Fully Sequential Game with Lobbying	106
9. Game with Lobbying Followed by Simultaneous Play, Modified Payoffs	107
10. Game with Lobbying and Simultaneous Play in Normal Form	107
11. Basic Niskanen Model with One Level of Government	113
12. Niskanen Model with Competing Levels of Government	114
A1. Event Studies, Tax Decreases, Firms with Log Treatments Greater Than 0.005 in Magnitude	128
A2. Event Studies, Tax Increases, Firms with Log Treatments Greater Than 0.005 in Magnitude	129
A3. Event Study Placebo Tests, Tax Decreases, Using Random Firms with No Operations in State of Tax Law Change	132
A4. Event Study Placebo Tests, Tax Increases, Using Random Firms with No Operations in State of Tax Law Change	133
A5. Event Study Placebo Tests, Tax Decreases, Shifting Event Dates	

Figure	Page
Backwards in Time by Four Years	134
A6. Event Study Placebo Tests, Tax Increases, Shifting Event Dates	
Backwards in Time by Four Years	135
A7. Maximization Problems of the Empire and the Village	139
A8. Effects of Village Sabotage on the Utility Possibilities Set	142

LIST OF TABLES

Table	Page
1. Descriptive Statistics, Top Marginal State Corporate Tax Rates, Levels	15
2. Dates of Law Changes Relating to Top Corporate Tax Rates on Banks	19
3. Mean Number of Days between Event Dates Used in Event Studies, Tax Decreases (Standard Deviations in Parentheses)	20
4. Mean Number of Days between Event Dates Used in Event Studies, Tax Increases (Standard Deviations in Parentheses)	21
5. Content of Law Changes of Top Marginal Corporate Tax Rates on Banks	22
6. First-Difference Regression Results	30
7. CAPM-Style Regression Results	33
8. Placebo Tests, Treatment Shifted Backwards in Time	35
9. Placebo Tests, Treatment Shifted Forwards in Time	36
10. Timing of Market Response	38
11. Descriptive Statistics	56
12. Extensive Margin Linear Probability Regressions	61
13. Individual Level Regressions	64
14. Individual Level Regressions, by Industry	70
15. Individual Level Regressions, by Occupation	71
16. Effects of Rounding	75
17. Individual Level Regressions, Respondents without Rounded Wages . .	77
18. Individual Level Regressions, Misreporting Not Likely Caused by Rounding	80

Table	Page
19. Tax Unit Regressions	82
A1. Sample Evolution	136

CHAPTER I

INTRODUCTION

The design, use, and implementation of public finance systems are of key concern for all involved in our modern economy. Tax economists are concerned with how tax system design affects who ultimately bears the burden of taxation. Governments use data collected from tax systems in order to examine the health of our economy and to direct public spending. A complex system of government entities implements various government policies, though it is not always clear which entity should manage which role. These three interlocking aspects of public finance are the subject of this dissertation.

In the first substantive chapter, I examine the incidence of the corporate income tax on the owners of corporate capital, i.e. the shareholders. I examine the response of the stock prices of regional banking stocks to changes in US state-level corporate income tax rates. I build a novel dataset featuring the legislative history and details of state law changes involving the corporate income tax rates on financial corporations. I find that stock prices and corporate income tax rates are inversely related. For tax decreases, the relationship is statistically insignificant, but for tax increases, the effects are significant and large. Stock prices show approximately a ten percent decrease for every percent increase in corporate income tax rates. These effects are felt mostly during the period following the announcement of a law change and the introduction of that law change into the state legislature. The asymmetrical results are common in the literature about corporate income taxes.

The second substantive chapter includes previously unpublished co-authored material. It is co-authored with John Voorheis of the US Census Bureau and Caroline Weber of the University of Washington. I wrote the majority of the chapter, with John Voorheis providing important work on the Data section and with institutional details, and Caroline Weber providing edits. I wrote the majority of the computer code used to create this chapter, with John Voorheis contributing code for the visualizations of income misreporting within groups across income levels, and Caroline Weber providing valuable supervision and code writing lessons and examples. I conceived this chapter, which is an offshoot of another ongoing project. In this chapter, we examine differences in wage reporting between individuals' Internal Revenue Service and other administrative records versus their responses to the Current Population Survey. We measure the percentage gaps between wages reported for survey purposes versus tax purposes. We find that these gaps vary by demographic attributes, including age, educational attainment, and racial and ethnic groupings. A large proportion of these gaps are due to the rounding of numbers, but rounding does not drive results. We propose econometric corrections for mismeasured wage data and suggest changes to survey design.

In the final substantive chapter, I examine the possible strategic interactions between multiple levels of government via a series of abstract games. In the first set of games, I investigate what might happen if a subordinate level of government has the opportunity to sabotage a policy being implemented by a dominant level of government. I find cases where the dominant level of government responds by allowing the local level to implement policy, as well as cases where the local level of government will indeed choose to sabotage the plans of the dominant level. In the second series of games, I investigate what might happen if bureaucrats representing

these levels of government were to compete for the opportunity to implement policy by making bids. I find that the mere existence of multiple levels of government can be welfare improving for the citizens in these circumstances.

CHAPTER II

DO STOCK PRICES RESPOND TO CHANGES IN CORPORATE INCOME TAX RATES?

Introduction

For the last few decades, Federal and state lawmakers in the United States have paid significant attention to improving the efficacy of the corporate income tax. Gravelle (2017) enumerates several public policy examinations at the Federal level, including the 2005 Advisory Panel on Tax Reform, several opinion pieces by policy makers, a 2007 Treasury Department background paper, and numerous bills introduced in Congress. A similar level of examination has occurred at the state level.¹ Though there are numerous issues complicating the debate over the corporate tax, in this paper, I examine the impacts of corporate tax changes on shareholders.

The incidence of the corporate tax is notoriously slippery. Depending on model selection, choice of functional forms within models, and model parameterization, theory gives wildly different results for the incidence of the corporate tax, ranging from more than one hundred percent to less than zero percent of the burden falling on shareholders (Harberger (1962), Auerbach (2006), Harberger (2006)). Empirical results are also far from conclusive, with a wide variety of incidence rates. Estimates of the incidence on labor range from near zero to up to two thousand percent.² A minority of corporate tax incidence studies

¹Most recently in Iowa, where state corporate tax rates were overhauled (Bloomberg Tax, 2014).

²For discussions see Desai et al. (2007) and Dwenger et al. (2011).

estimate the incidence on capital (examples include Cragg et al. (1967), who estimate that capital bears close to the full burden of the tax, and Desai et al. (2007), who estimate that capital bears between twenty-five and fifty-five percent of the tax).

In this paper, I try to shed some light on the existing contradictory empirical results by providing evidence from stock markets. The major contribution of this paper is to, for the first time, causally estimate the incidence of corporate income taxes on shareholders, by using a large set of comparable corporate income tax rate changes and a rich set of stock price panel data. Taking the stance that stock markets efficiently integrate publicly available information into stock prices, stock price changes that accompany corporate tax rate changes should inform us about the present value of the burden falling on shareholders.

Estimating stock price changes caused by changes in tax rates at the federal level can be difficult, as the federal tax code often changes all at once in infrequent, major overhauls, so the effects of a particular tax change among many simultaneous changes are difficult to tease out. Cross-country comparisons are challenging because of the political and cultural heterogeneity of the underlying countries. What is needed is a set of reasonably similar jurisdictions, a set of corporate tax rate changes that vary across time and jurisdictions, and a set of stocks of comparable companies where the jurisdiction-by-jurisdiction apportionment of income can be ascertained.

I examine the effects of changes in state-level corporate tax rates on the value of stocks listed on the major United States stock exchanges. For most US equities, determining the exact state-by-state breakdown of corporate income is at best a noisy task, as tax returns of publicly traded corporations are not publicly available

information, and proprietary sources of corporate data may not have sufficient data to properly allocate income across states. Fortunately, there is a subset of stocks that permit accurate state-by-state allocation: regional bank stocks. These stocks fulfill a number of desirable characteristics for this exercise: the companies are often completely located within one or a handful of states, corporate taxes represent a large proportion of their net profits, their branch units are comparable within companies, and operations of different companies are similar. I calculate the percentage of each firm's income allocated to each state for tax purposes. Using the set of state-level corporate tax rate changes from 1994-2017, I examine the change in stock valuation of these companies around the public unveilings of corporate tax rate changes, using a series of first-difference regressions.

First-difference regressions are necessary because stock prices follow a random walk (Samuelson (1965) and Malkiel (1973)). The resulting regressions have a similarity to the types of regressions found in arbitrage pricing theory ("arb models"). I augment these regressions with similar regressions that remove the risk-free component from stock returns, and in some cases, use historical measures of diversifiable risk to isolate the "abnormal" component of stock returns. These latter regressions have the flavor of the Capital Asset Pricing Model (CAPM), developed by Sharpe (1964) and Lintner (1965). Both arb- and CAPM- style analyses yield similar results: corporate tax rates and stock prices bear an inverse relationship, and the sizes of the associated stock returns are greater in magnitude than what would be expected if only future tax expenses were to change, *ceteris paribus*.

Further inspection reveals that this effect is almost entirely driven by tax increases. When the full sample of stock returns and tax changes is split into tax decreases versus increases, I find that tax decreases of about one percent are

associated with stock price increases of about one percent, though the estimates are statistically insignificant. On the other hand, when accounting for outliers, tax increases of about one percent lead to an average decline in stock price of about ten percent.

Overall, these results are large, and can possibly be explained by a combination of the short-run incidence of the tax (that is, the initial impact on profits before production factors can be reallocated more efficiently) as well as the longer-run effects of interstate competition between banks, as well as competition between C corporations (those companies directly affected by the new tax rates) and S corporations (those companies that do not pay the corporate tax). If firms in different states and with different tax treatments face different cost structures due to different corporate income tax rates, then those firms with lower costs can offer their products at lower cost to consumers, increasing their market share and decreasing the market shares of higher cost firms. A number of possible theories can help to explain the asymmetry between stock price responses to tax increases versus decreases. Taxes affect the capital structure of firms asymmetrically, as the government shares in corporate profits but not losses, and the ability for a firm to borrow is likely more impeded by an increase in expenses than it is helped by a decrease in expenses. However, the size of the effects suggests more is going on. One possible suggestion is that investors, believing that business has some control over the actions of the state government, view tax increases as a sign that the state government is becoming less business-friendly, so the tax increase may be a harbinger of the loss of regulatory capture of the state legislature. Finally, tax decreases may lead to more possible choices for corporate boards compared to tax increases. With tax increases, expenses of some firms rise while expenses of other

firms do not, causing the higher taxed firms to cut profits in order to compete on loan prices. On the other hand, firms facing tax decreases face more options. They could use their newfound advantage to compete on price, or management incentives could be so aligned as to motivate management to keep some of the lower-tax windfall for themselves rather than passing it on to shareholders.

I further break the timing windows of the tax changes down into smaller “subwindows,” in order to determine when, in the process of becoming law, the tax changes impact stock prices most. I find that most of the change in stock price occurs between when legislators first announced the upcoming change and when they introduce the bill to the state legislature. The majority of the remainder of the change in stock price occurs between when the bill is introduced and the bill is passed. Little changes between when the bill is passed when the bill is signed by the governor, consistent with the finding that governors almost never vetoed corporate income tax rate changes between 1994-2017.

I test the robustness of my main findings with a battery of placebo tests. In general, these tests find that the large, significant declines in stock prices caused by tax increases are not more prevalent during time periods when these tax changes did not occur.

These results must be taken with some caveats, as even the most basic corporate tax incidence models show that differing production functions yield meaningfully different results (Harberger, 1962), and banks may not be representative of publicly traded corporations in other industries. For example, these companies are highly levered relative to the stock market as a whole.³

³An analysis of the debt ratios of the Dow Jones Industrial Average components (sans financial components) versus an equally sized sample of regional bank stocks, shows that the bank stocks have total debt-to-equity ratios that are, on average, approximately six times higher than the Dow components.

Therefore, lawmakers considering corporate income tax rate changes need to be cognizant of the types of firms adopting the C corporation form, if they are to have any ability to predict the impact of these changes. However, banking is an industry tied to most other industries, so these results may be more externally valid than those of many single-industry studies.

The next section provides a review of related literature. After that, I describe data used in this study while providing institutional details relevant to the data. I then describe the methodologies used and present results for the “arb-style” and “CAPM-style” regressions, including regressions within subwindows and placebo regressions. I discuss reasons for the timing, magnitude, and asymmetry of results. Finally, I conclude, commenting on avenues for future research.

Related Literature

This research follows from findings in the public finance and financial economics literatures. First, it draws from and speaks to a venerable but ever controversial literature on the incidence of the corporate tax. As mentioned, conclusions about the incidence of the corporate tax are far from settled. Even in Harberger’s seminal paper, small changes in modeling assumptions drastically change results, from all of the burden falling on capital to most of the burden falling on labor. It should be noted that Harberger’s baseline result has the entire burden falling on *all* capital, not just corporate capital. However, this baseline result is based on assumptions about the capital intensity of the corporate sector. The more general model in his paper allows for the corporate sector to be more labor intensive, which will lead to more of the burden falling on workers. A more modern review of the theoretical literature can be found in Auerbach (2006). Of

note, Harberger's closed economy results reverse in an open economy model, with all of the burden falling on labor. An important common feature of most models is that they describe long run phenomena and ignore short run effects.

On the empirical side, most research focuses on the incidence of the corporate tax on labor. This is at least partly due to the fact that owners of corporate capital tend to be on the upper end of the income distribution, whereas the typical laborer is not. Thus, determining incidence has great importance in terms of the progressivity of the corporate tax. Examples of these labor-focused studies include Dwenger et al. (2011), who find an elasticity of wage rates to tax rates of -2.37, Arulampalam et al. (2012), who find an elasticity of -0.92, and Hassett and Mathur (2010), who find an elasticity of -0.5 to -0.6. Early empirical studies on the incidence on capital include Krzyzaniak and Musgrave (1963), who find that capital can benefit from a tax, and Cragg, Harberger, and Mieszkowski's rebuttal, which reverses the 1963 results due to previously unaddressed econometric shortcomings. A more recent study is Desai et al. (2007), who estimate the relative burden on labor and capital by assuming that they sum to unity, and find that twenty-five to fifty-five percent of the burden falls on capital. Gordon (1985) finds small benefits of the corporate tax on investment. In addition, Gravelle (2017) gives a thorough summary of other studies relating to the corporate income tax (Laffer curve, investment, etc.) as well as describing common econometric issues that arise in studying the corporate income tax.

This paper adds to a newer, burgeoning body of literature that relies on changes to state corporate tax rates in the United States as a source of variation. An early example is Feldstein and Poterba (1980), who find that an omission of state and local taxes understates the rate of return to capital. More recent

examples include Felix and Hines (2009), who show that unionized workers benefit by capturing approximately half of the benefit of lower state corporate tax rates, Giroud and Rauh (2015), who find that employment and number of establishments both have state corporate tax elasticities of approximately -0.4, and Heider and Ljungqvist (2015), who find that the use of leverage has a state corporate tax elasticity of about 0.4. In addition, Ljungqvist and Smolyansky (2014) find the asymmetric result that tax increases decrease employment and firm income but tax decreases do little, Ljungqvist et al. (2017) find another asymmetry in that tax increases reduce firm risk-taking while tax cuts do little, and Suarez-Serrato and Zidar (2016) find that the narrowing of states' corporate tax bases over time reduce states' ability to raise revenue through rate increases. Important themes in this literature include the use of tax changes, not levels, as a source of variation (implored for by Auerbach (2006)), leveraging the vast heterogeneity of these changes over space and time, and noting asymmetric results stemming from tax increases versus decreases.

Important for this research, Giroud and Rauh (2015) find that most state corporate tax changes are exogenous with respect to the income of individual firms, following a narrative method of categorizing tax changes as more or less exogenous according to Romer and Romer (2010). This finding is supported by Ljungqvist and Smolyansky (2014), who note that corporate tax revenues typically account for only a small portion of state revenues. The corporate tax may be of second order consideration for closing budget deficits, thus these changes may tend to be more exogenous.

Previous studies differ from this paper in that they focus on non-financial capital. Many states have different corporate tax rates on financial and non-

financial institutions. To my knowledge, this is the first paper to focus mainly on the effects of changes in state corporate tax rates on financial institutions. As a result, some of the data look slightly different, and some previous conclusions, such as the inference that state corporate tax rates follow a random walk (Ljungqvist and Smolyansky, 2014) must be revisited. Previous studies also look at firm behavior as rate changes come into effect, whereas this study looks at the market's response to rate changes as the laws are announced and made, well before they come into effect (or sometimes well after, in the case of retroactive law changes).

This paper begins with no prior assumption that markets will notice the impacts of corporate income tax rates on firm prices. Much work has been done in recent years investigating whether people respond to tax incentives, especially if the taxes are shrouded or complex. For taxes to be incorporated into purchasing and selling decisions (including investing decisions), these taxes must be salient. While Rosen (1976) finds similar relationships between wage rates and working hours and tax rates and working hours, showing that these taxes are salient, Chetty et al. (2009) find that non-posted sales taxes do not lead to decreases in quantities as much as similarly sized price changes, implying that non-posted sales taxes are not salient.⁴ If corporate income taxes change and stock prices do not, this could mean that these taxes had no effect on future profits (e.g., there is no incidence on shareholders), or that such taxes do indeed have an effect, but that they are not salient to investors. Determining the proper value of investments is the job of larger investors, but taxes may not be salient if they are minuscule or hard to keep track of.

⁴Another similar paper is Finkelstein (2009), who finds that drivers do not fully respond to automatically deducted (and therefore less salient) road tolls.

The efficient markets hypothesis (EMH), related to the work of Samuelson (1965), developed by Fama (Fama (1965), Fama et al. (1969), Fama (1970)), and made famous in the popular press by Malkiel (1973) helps to square the ambiguity of the tax salience literature with investor behavior. Fama (1970) developed the notion of the semi-strong form of the EMH, which states that stock prices reflect all publicly available information. If corporate income taxes are salient, then investors should properly evaluate their impacts when determining the values of equities. Lo (2005) enumerates how much of widely accepted modern financial economics is derived from the EMH. Although non-behavioral critiques of the EMH exist (see, for example, Buffett (1984)), the majority of EMH criticisms come from the behavioral finance camp (a summary of these criticisms can be found in Lo (2005)). Malkiel (2003) rebuts these criticisms by noting that many of the most famous behavioral exceptions, such as the January effect, disappear nearly as soon as they are discovered; in other words, non-salient but material determinants of price soon become salient, at least in efficient equities markets. Lo (2005) has an interesting approach to this debate, reconciling the EMH and behavioral anomalies by combining the two camps into an imperfect but rational evolutionary process in which rational stock trading strategies are learned over time and respond to changing market conditions. This paper takes the stance that the major US stock markets are imperfectly efficient in the semi-strong sense, but efficient enough to capture the impacts of publicly available information in a reasonable amount of time (say, a few days).

Finally, this paper relates to an ongoing debate about the transparency of publicly traded companies' tax information. Currently, the tax returns of publicly traded companies in the United States are not made publicly available. This paper

makes use of company data where corporations' business establishment locations are identifiable using publicly available information; in most cases they are not. This point is debated, for example in Lenter et al. (2003). While the debate is complicated (for example, full transparency may lead to companies publishing lower quality information), this paper suggests that persons who have detailed information about the state-by-state allocations of corporations' incomes may have a trading advantage over other investors.

Data and Institutional Details

The goal of this paper is to show how changes in state corporate tax rates effect stock returns; thus, data must consist of a sample of stock returns and a schedule of state corporate tax law changes. For years 1994-2017, I look at changes to states' top marginal corporate tax rates, as these types of law changes are most comparable across states ("states" refers to the 50 states plus the District of Columbia).⁵ Table 1, which combines data culled from *The Book of the States* as well as state websites, provides information on the levels of the top marginal tax rates on C corporations for years 1994-2017 for the 51 "states" in the sample. Of note, states exhibit wide heterogeneity in rates throughout the sample, and over time, there is a trend towards lower tax rates. From this point onward, I ignore

⁵These top rate changes include changes to surtaxes, which alter the top rates *effectively* paid by corporations. I only look at top rates, even in the case of states with multiple brackets, for simplicity. First, most states only have one rate. States with multiple brackets have top brackets beginning at incomes so low relative to the typical income of a firm in my sample so as to not significantly affect my findings. Possible exceptions are the \$250,000 top bracket in Kentucky and the \$1,000,000 brackets of New Mexico and South Dakota, which affect less than one percent of profits in the sample of firms.

states with a corporate tax base that is not based on net income (for example, states that tax gross receipts instead).⁶

TABLE 1.
Descriptive Statistics, Top Marginal State Corporate Tax Rates,
Levels

Tax variable	1994 (beginning of sample)	2017 (end of sample)
States taxing net income of non-financial corporations	44	44
States taxing net income of financial corporations	45	45
States w/ 0 top rates, non-financial corps.	4	4
States w/ 0 top rates, financial corps.	3	3
States w/ tax based on variables other than net income	3	3
States w/ multiple corp. tax brackets	13	14
Highest bracket (STATE)	\$1M (NM)	\$1M (NM, OR)
No. states w/ different rates for financial and non-financial corporations	15	14
Highest top marginal rate, non-financial corporations (STATE)	12% (IA)	12% (IA)
Highest top marginal rate, financial corporations (STATE)	12.54% (MA)	10.84% (CA)
Lowest non-zero top marginal rate, non-financial corporations (STATE)	4% (KS)	3% (NC)
Lowest non-zero top marginal rate, financial corporations (STATE)	1% (ME, SD)	0.25% (SD)

Corporate taxes are assessed on the taxable net incomes of Subchapter C corporations (Subchapter S corporations and other pass-through entities are not

⁶States excluded for this reason are Michigan, Ohio, and Texas.

subject to the tax). The taxable net incomes of companies in the stock sample are substantially similar to those companies' accounting profits for book purposes.⁷

Given that, under the EMH, stock prices reflect all publicly available information, stock prices should, at the very latest, change shortly after a tax change is made certain, i.e. when the proposed tax change is signed into law. It is important to note that changes in tax rates do not map one-to-one with changes in state laws; one law may enact multiple rate changes over time.⁸ Thus, the events used in this study are dates relating to changes in *laws*, which do not always coincide with the dates of *rate changes*.

Many states have different corporate tax rates for financial institutions, and since I examine the returns to regional bank stocks specifically (explained momentarily), my schedule of law changes looks slightly different than those schedules used in the aforementioned literature. I identify these law changes first by identifying changes in the top marginal rates paid by financial institutions by comparing different years' rate schedules found in *The Book of the States* for years 1994-2017 (Council of State Governments, 2017).⁹ Where surtaxes are employed, top marginal rates and surtax rates are combined into one top effective marginal rate. From these rate changes, I map the large set of top marginal rate changes to a smaller set of corresponding law changes, by searching LexisNexis and the websites of state legislatures. For each applicable law change, I collect five key dates: the

⁷Although there are a number of "M-1 adjustments" made to reconcile net income for tax purposes and net income for book purposes, this study merely assumes that such adjustments net to zero. Differences in the accounting presentations of tax expenses for tax purposes versus book purposes can lead to adjustments such as deferred tax assets and liabilities. I collected firm-year level data on deferred tax assets and liabilities, where available. Inclusion of net deferred tax asset data does not meaningfully change results.

⁸A complicated example is the corporate tax law passed in Indiana in 2013, which phased in nine successively lower rates over the course of ten years.

⁹For further information about the collection of rate change data, see the Appendix.

date that the earliest talk of an impending corporate rate change was in the news, the date the legislation was introduced, the date the legislation was passed by the state legislature, the date the bill became law, and the date(s) the rate change(s) in the law went into effect (laws may feature multiple rate changes over multiple years).¹⁰ If investors view the legislative process as a process of an impending law becoming more certain, then, under the EMH, stock prices leading up to the signing of a law should reflect investors' evolving notions of the probability of the law's passage. Of these dates, the fuzziest is the first date, hereafter called the "first news date." Unlike the other dates, there is some subjectivity in selecting this date, and choosing a date that is too late may result in missing out on a period of time when investors assessed the likelihood of an impending law change as ever increasing (thus, by picking too late of a date I may miss the relevant stock market reaction).¹¹

For each law change affecting top marginal rates, additional data collected includes whether or not the change only affected financial institutions, and whether or not the rate changed via a surtax rate change or a regular rate change. Additionally, based on narratives pieced together from LexisNexis news articles, I code each law change using the four categories listed in Romer and Romer (2010) in order to assess the exogeneity of each law change: one for law changes designed to increase output, two for changes designed to change variables related to output, three for dealing with inherited budget deficits, and four for philosophical or ideological reasons such as fairness. I add a fifth category, not found in their

¹⁰These first four dates always follow in chronological order, but the fifth may not, in the case of retroactive law changes. Nineteen of the forty-nine law changes studied were retroactive.

¹¹Just because an impending law change is not touted in the news does not mean that lawmakers and investors are not already discussing it in less public circles.

paper, for exogenous law changes that were due to an outside body determining that the tax code must change.¹² Reasons one and two describe more endogenous tax changes while the remainder are considered to be exogenous. Each law change is given a two letter, two digit abbreviation for ease of reference, using the state abbreviation and the year the first top rate change in the law came into effect (e.g. MD08).

Since many law changes featured multiple rate changes over time, each law change is distilled down to one overall top rate change, expressed in logarithmic change in present value of future earnings of firms. For details on this procedure, see the Appendix. This log change can be thought of as the overall percent change in corporate tax rates.¹³ Table 2 shows the dates of events relating to all state law changes relating to changes in top marginal corporate income tax rates on financial institutions for 1994-2017. Tables 3 and 4 distill this information by displaying the average number of days between key dates in the legislative history for a subset of tax decreases and increases that are used in the event studies in the next section. Table 5 provides summary information for relating to the content of each law change, including the top rate in effect before each change, top rate or rates after the change, and the overall magnitude of the change.

In order to assess the impact of state corporate tax rate changes on stock prices, one must be able to determine how much of the underlying companies' incomes are allocated to each state for tax purposes. This task is complicated by states' tax nexus laws, which dictate how income is apportioned to the various

¹²For example, this could occur when a portion of a state's tax code was struck down by the Supreme Court.

¹³Technically, it is constructed not as the change in tax rate τ , but rather the change in $(1 - \tau)$, multiplied by negative one, adjusted for the timing of the law change.

TABLE 2.
 Dates of Law Changes Relating to Top Corporate Tax Rates on
 Banks

Abbrev.	First news date	Bill introduced date	Bill passed date	Bill became law date	Law effective date
AL02	3/1/99	11/15/99	11/23/99	3/21/00	1/1/01
AZ94	1/1/94	3/7/94	3/30/94	4/4/94	12/31/93
AZ99	11/4/97	5/8/98	5/8/98	5/20/98	1/1/98
AZ00	11/4/97	2/4/98	5/6/98	5/19/98	12/31/99
AZ01	1/11/99	4/7/99	4/7/99	4/15/99	1/1/01
AZ14	11/19/10	2/14/11	2/16/11	2/17/11	1/1/14
CA97	1/4/94	4/10/96	7/8/96	7/15/96	1/1/97
CO99	12/16/98	1/13/99	5/3/99	6/4/99	1/1/99
CO00	7/13/99	1/5/00	5/1/00	5/3/00	1/1/00
CT95	10/28/94	5/27/95	5/31/95	6/1/95	1/1/95
CT03	12/1/02	2/3/03	2/18/03	3/6/03	1/1/03
CT04	4/16/03	7/30/03	7/31/03	8/16/03	1/1/04
CT06	2/9/05	3/16/05	6/7/05	6/30/05	1/1/06
CT09	2/9/09	8/31/09	8/31/09	9/8/09	1/1/09
CT12	2/14/11	5/2/11	5/3/11	5/4/11	1/1/12
CT14	5/13/13	6/1/13	6/3/13	6/18/13	1/1/14
CT16	3/2/15	6/2/15	6/3/15	6/30/15	1/1/16
DC95	2/27/91	1/14/94	8/1/94	8/2/94	1/1/95
DC15	2/15/14	2/15/14	6/24/14	6/25/14	1/1/15
ID01	1/21/00	3/26/01	3/29/01	4/11/01	1/1/01
ID13	12/1/11	2/17/12	3/29/12	4/5/12	1/1/12
IL11	1/4/10	1/6/10	1/12/11	1/13/11	1/1/11
IN14	10/14/12	2/18/13	4/1/13	4/26/13	1/1/14
KS98	10/3/97	1/10/98	3/1/98	3/18/98	1/1/98
KY06	2/12/04	2/2/05	3/8/05	3/18/05	1/1/05
MA95	2/2/95	5/10/95	7/17/95	7/27/95	1/1/95
MA10	12/18/07	6/13/08	7/3/08	7/3/08	1/1/10
MD08	7/26/07	10/29/07	11/13/07	11/19/07	1/1/08
NC97	12/9/94	7/8/96	8/2/96	8/2/96	1/1/97
NC09	10/20/08	2/17/09	8/5/09	8/7/09	7/1/09
NC14	4/28/12	4/17/13	7/17/13	7/23/13	1/1/14
ND17	6/17/16	8/1/16	8/3/16	8/5/16	1/1/17
NH00	1/7/99	3/4/99	4/22/99	4/29/99	7/1/99
NH16	10/29/14	2/18/15	6/24/15	9/16/15	1/1/16
NJ06	1/17/06	6/26/06	7/8/06	7/8/06	1/1/06
NM14	1/3/13	2/14/13	3/16/13	4/4/13	1/1/14
NY00	12/17/97	1/20/98	4/14/98	4/28/98	7/1/99
NY07	4/27/05	1/11/06	3/31/06	3/31/06	1/1/07
NY16	12/10/13	1/6/14	3/31/14	3/31/14	1/1/16
OR10	11/20/08	3/12/09	6/11/09	7/20/09	1/1/10
PA94	1/15/93	3/2/93	6/14/94	6/16/94	1/1/94
PA95	11/15/94	1/23/95	6/15/95	6/30/95	1/1/95
SD01	1/13/99	1/25/00	2/18/00	3/3/00	1/1/01
TN03	3/30/99	1/31/02	7/3/02	7/4/02	7/15/02
VT97	2/5/97	3/13/97	6/12/97	6/26/97	1/1/97
VT07	1/6/04	4/21/04	5/19/04	6/7/04	1/1/07
WV07	10/30/06	10/30/06	11/13/06	11/14/06	1/1/07
WV09	2/26/07	2/15/08	3/8/08	4/1/08	1/1/09

TABLE 3.
Mean Number of Days between Event Dates Used in Event
Studies,
Tax Decreases (Standard Deviations in Parentheses)

	Legislation Introduced	Legislation Passed	Legislation Signed	Legislation Effective
First News Date	191.65 (157.92)	228.00 (158.66)	242.06 (159.83)	567.71 (310.78)
Legislation Introduced		60.79 (105.24)	74.00 (107.99)	345.74 (343.82)
Legislation Passed			12.60 (18.71)	280.25 (352.71)
Legislation Signed				251.29 (352.79)

states based on varying ratios of sales, payroll, and property. For most publicly traded companies, this is opaque, as publicly traded companies do not have to break down their sales, payroll, and property factors by state in their public filings, and popular proprietary databases do not contain all three factors. I use a sample of regional and community bank stocks in the US because they overcome this impediment. Unlike most public firms, who raise equity capital in order to expand perhaps nationally or even internationally, companies in the sample of regional bank stocks tend to only have operations in one or a few states, suggesting that part of these firms' business strategy is to stay small and develop a community-oriented reputation. Alternatively, the fact that these firms tend to be limited to a small number of states may be a vestige of old state banking laws that required banks to overcome legal obstacles to operate in multiple states. The modal firm in my sample only operates in one state.

The business of these smaller banks is simple: they accept deposits which are used to make loans, primarily to homeowners and small businesses. The vast majority of these companies break down the number of bank branches that are

TABLE 4.
Mean Number of Days between Event Dates Used in Event
Studies,
Tax Increases (Standard Deviations in Parentheses)

	Legislation Introduced	Legislation Passed	Legislation Signed	Legislation Effective
First News Date	247.67 (397.48)	362.17 (419.80)	386.83 (414.19)	422.67 (450.57)
Legislation Introduced		114.50 (136.65)	139.17 (123.80)	175.00 (182.72)
Legislation Passed			24.67 (46.46)	60.50 (187.55)
Legislation Signed				35.83 (149.10)

in each state in their annual Forms 10-K, which are filed with the Securities and Exchange Commission (SEC) and made publicly available via the SEC's online Electronic Data Gathering, Analysis, and Retrieval system (EDGAR). EDGAR only maintains filings back to 1994. I use the descriptions of bank branch locations to allocate income for each company across multiple states, in proportion to the number of branches.¹⁴ Corporate taxes represent a large portion of these companies' profits, giving investors a good reason to pay attention to the applicable tax rates.¹⁵

The sample of regional bank stocks was created by combining a list of regional bank stocks from InvestSnips.com and by searching the Center for Research in Security Prices (CRSP) database (accessed via the Wharton Research Data Service (WRDS)) for companies with Standard Industrial Classification (SIC)

¹⁴This assumes that each branch within a company uses the same amount of payroll, property, and sales (interest revenue from loans) at each branch. This assumption is unnecessary in the case of companies with operations in only one state. For further information about the allocation of income across states, see the appendix.

¹⁵For example, Umpqua Bank, a typical firm in my sample, paid twenty-two percent of pre-tax accounting profit in Federal corporate income taxes and six percent of pre-tax accounting profit in state corporate income taxes in 2017, for a total of twenty-eight percent.

TABLE 5.
Content of Law Changes of Top Marginal Corporate Tax Rates on
Banks

Abbrev.	Top rate before	Top rate(s) after	Overall log size
AL02	6	6.5	0.0051
AZ94	9.3	9	-0.0034
AZ99	9	8	-0.0113
AZ00	8	7.968	-0.0003
AZ01	7.968	6.968	-0.0096
AZ14	6.968	6.5, 6, 5.5, 4.9	-0.0164
CA97	11.3	10.84	-0.0050
CO99	5	4.75	-0.0027
CO00	4.75	4.63	-0.0013
CT95	11.5	11.25, 10.75, 10.5, 9.5, 8.5, 7.5	-0.0370
CT03	7.5	9, 8.25, 7.5	0.0016
CT04	9	9.375, 7.5	0.0008
CT06	7.5	9, 7.5	0.0010
CT09	7.5	8.25, 7.5	0.0016
CT12	8.25	9, 7.5	0.0020
CT14	9	9, 7.5	0.0020
CT16	9	9, 8.25, 7.5	0.0025
DC95	10.25	9.975	-0.0030
DC15	9.975	9.4, 9, 8.5, 8.25	-0.0165
ID01	8	7.6	-0.0044
ID13	7.6	7.4	-0.0022
IL11	7.3	9.5, 7.75, 7.3	0.0075
IN14	8.5	8, 7.5, 7, 6.5, 6.25, 6, 5.5, 5, 4.9	-0.0287
KS98	6.375	4.375	-0.0215
KY06	8.25	7, 6	-0.0233
MA95	12.54	12.13, 11.72, 11.32, 10.91, 10.5	-0.0209
MA10	10.5	10, 9.5, 9	-0.0141
MD08	7	8.25	0.0135
NC97	7.75	7.5, 7.25, 7, 6.9	-0.0082
NC09	6.9	7.107, 6.9	0.0003
NC14	6.9	6, 5, 4, 3	-0.0359
ND17	7	4.31	-0.0276
NH00	7	8	0.0103
NH16	8.5	8.2, 7.9	-0.0060
NJ06	9	9.36, 9	0.0010
NM14	7.6	7.3, 6.9, 6.6, 6.2, 5.9	-0.0152
NY00	9	8.5, 8, 7.5	-0.0130
NY07	7.5	7.1	-0.0041
NY16	7.1	6.5	-0.0057
OR10	6.6	7.9, 7.6, 6.6	0.0031
PA94	12.25	11.99, 10.99, 10.75, 9.99	-0.0236
PA95	10.99	9.99	-0.0013
SD01	1	0.25	-0.0072
TN03	6	6.5	0.0052
VT97	8.25	9.75	0.0171
VT07	9.75	8.5	-0.0116
WV07	9	8.75	-0.0027
WV09	8.75	8.5, 7.75, 7, 6.5	-0.0184

codes relating to regional banking.¹⁶ If banks lobby for lower tax rates, stock returns could be endogenous to tax rate changes. To minimize these concerns, I drop companies belonging to the Financial Services Roundtable lobby and companies with a market capitalization of over ten billion dollars.¹⁷ In order to make sure markets are efficient enough to capture publicly available information in a timely manner, I drop banks that do not trade on major US exchanges and banks with market capitalizations below ten million dollars, as the markets for the stocks of these smaller banks may be too thin to obtain meaningful pricing data.¹⁸ I also drop companies that are “too national,” i.e. that have operations in more than ten states. Daily stock returns for the final sample of 639 firms are downloaded from the CRSP database.

Additional firm and firm-year data are downloaded from CRSP and Compustat. From Compustat, I have the major balance sheet and income statement items by firm-year (total assets, liabilities and equity, pre-tax and after-tax net income, as well as net deferred tax assets), plus the headquarters state of each firm. These data are used to create common ratios (such as the debt to equity ratio of each firm), as well as lagged values of the financial statement items (and ratios derived from them).¹⁹ From CRSP, I obtain the returns to the Standard and Poor’s 500 market index (S&P 500) firms’ market capitalization, daily trading volume and shares outstanding (from which the proportion of shares traded daily

¹⁶InvestSnips.com is a website specializing in creating themed lists of stocks.

¹⁷The Financial Services Roundtable is the largest banking lobby in the US.

¹⁸For additional information about dropping stocks from this sample, see the appendix.

¹⁹Given that tax law changes can occur virtually anytime throughout the year, it is not clear at what point in a year investors may be relying on brand new accounting and financial data or somewhat stale data. I find that the choice to use contemporaneous versus prior year data for these financial statement items and ratios does not seem to make a marked difference (potentially owing to the similarity of a firm’s financial structure from one year to the next).

is computed), and well as historical measures of individual stock risk, namely each firm's "beta," or the naive regression coefficient of each firm's relationship with the broader stock market. These betas are used to develop a historical expectation of each stock's co-movements with the market.

From the Federal Reserve Economic Database (FRED), I download measures of the federal funds rate, as changes in said rate likely have impacts on the profitability of banks' excess reserves. The number of housing starts is also obtained from FRED, as regional banks in particular are likely responsive to changes in the housing market, and the federal funds rate can be viewed as a measure of risk-free returns. Finally, from Kenneth French's website, I download the Fama-French components of the CAPM (namely, the risk free return and market return (which is based on a broader set of stocks than the S&P 500)), as well as the additional SMB ("small minus big" market capitalization, the excess return of smaller companies versus larger companies) and HML ("high minus low," the excess return of higher book-to-market ratio stocks versus low book-to-market stocks) factors, which have been shown to explain most of the shortcomings of the predictive power of the CAPM (Fama and French, 1993).

Firms are treated based on the size of the tax change that affects their operations. If a firm has no operations in a state of a law change, it receives a zero treatment for that law change. If a firm only operates in one state, and that state has a law change, the firm's treatment for that observation is the full size of the change as shown on Table 5, which is positive for tax increases and negative for tax decreases. If a firm has some but not all operations in the state of a law change, its treatment for that observation is the full treatment from Table 5 multiplied by the

proportion of firm operations in the state in question.²⁰ One observation contains a firm's stock return over a given period of time, the treatment corresponding to that time frame, and other variables of interest including other market and time-varying data (as enumerated previously) corresponding to that time period or the beginning of that time period, time invariant firm characteristics, and change-specific characteristics.

Methodology and Results

In this section, I describe some of the preliminary tests needed to study the relationship between stock prices and expected tax rates, and I detail the design and results of the main methodologies used to determine the causal impacts of corporate income tax rate changes on stock prices. Stock prices are known to follow a random walk, and I check for random walk properties in the series of state corporate income tax rates on financial C corporations. I determine that first-differencing variables is appropriate. First, I run a set of first-difference regressions, regressing logarithmic stock returns on log tax rate changes and other controls. In addition to utilizing ordinary least squares, I also use robust regressions, because coefficient estimates produced by these regression types are less influenced by outlier observations. The estimated equations bear a resemblance to arbitrage pricing models. I then isolate the part of stock returns that comprises the risk premium, and also the part that comprises the risk-adjusted abnormal return, in order to run first-difference regressions that have a CAPM flavor. To try to isolate the timing of the stock changes relative to the history of the tax change laws, I then divide up the overall time window used in the main regressions into smaller

²⁰For example, if a tax decrease's overall size is determined to be a log change of -0.008, and a firm has fifty percent of operations in the state in question, its treatment is -0.004.

subwindows and run the regressions on within these subwindows (hereafter referred to as within-subwindow regressions). Finally, I run placebo regressions as if the tax law changes occurred a number of years prior to or after the actual law changes.

Preliminary Tests

Although it may be tempting to regress variables related to capital on corporate tax rates (Krzyzaniak and Musgrave, 1963), such regressions are bound to be spurious, due to the random walk properties of these time series variables. Stock prices have long been argued to follow a random walk (hence the title of Malkiel's famous opus), and past studies have shown that state corporate tax rates follow a random walk (Ljungqvist et al., 2017). Because other studies use samples of non-financial firms, which sometimes have different corporate tax rates than financial firms, I run Dickey-Fuller tests on the top marginal corporate tax rates for the forty-five of fifty-one states that tax the income of financial corporations (Dickey and Fuller, 1979).²¹ I fail to reject the null hypothesis of a unit root for all states tested except California and Connecticut, which reject at the one and five percent levels, respectively. If a trend is included, Connecticut results become statistically insignificant, and California results are only significant at the ten percent level.²² Taken as a whole, these results suggest that top rates of state

²¹To be quite technical, corporate income tax rates can be thought of as a bounded variable $\in [0, 1]$ (they do not have to be so confined, as corporate rates could be effectively negative). By the strictest definition, this means that they cannot be thought of as a true random walk (neither can stock prices follow a random walk, as stock prices cannot be negative). However, an accepted procedure is to run Dickey-Fuller tests nonetheless. Little has been done to deal with this issue. An exception is Cavaliere and Xu (2014). A more advanced model, appropriate to the observed behavior of tax rates, would treat a zero rate as an absorbing state and allow regime switching between the zero state and non-negative rates that follow a random walk.

²²The results for Connecticut may be driven by the fact that the Connecticut legislature repeatedly renewed a large surtax which was originally intended to exist for only a couple of

income taxes on financial corporations also follow a random walk. By definition of a random walk, these results help to alleviate concerns that investors may expect tax rates to return to some average level after the implementation of a rate change.

First-Difference Regressions

The random walk properties of stock prices and state corporate income tax rates make regressions of logarithmic stock returns on logarithmic tax changes potentially spurious. I move past this issue by first-differencing, yielding the model

$$\Delta \log(P_{i,t}) = \delta Treatment_{i,t} + \Delta X'_{i,t} \gamma + \Delta \epsilon_{i,t}, \quad (2.1)$$

where $\Delta \log(P_{i,t})$ are logarithmic stock returns for firm i over this timespan t , $Treatment_{i,t}$ is the firm-specific change in corporate tax rates as previously defined, $\Delta X'_{i,t}$ a matrix of controls, and $\Delta \epsilon_{i,t}$ a compound error term. The matrix of controls can be changed flexibly and can include changes in firm-specific variables such as financial statement ratios, changes in industry-wide variables such as the federal funds rate, and potentially even firm-specific trends. This model looks remarkably like an arbitrage pricing model, and, as such, I shall refer to these as “arbitrage-style” or “arb-style” models. However, excepting the firm-specific trend coefficients, the coefficients are not firm-specific.

In the first-difference equation (2.1), the index t is shorthand for an observation corresponding to a particular timespan, which further corresponds to a particular law. Unless otherwise noted, the window used for each law is from the first announcement date of a law minus five days until the date that the law was

years. California also had an unusual tax rate system for banks in the early to mid 1990s, where a committee would set tax rates periodically (thus rates were not known far in advance).

signed plus five days. There are forty-eight law changes, so the estimation process effectively “stacks” forty-eight sets of first-difference observations on top of each other. Because each tax law change has a unique history, there is a unique before and after date for each change, as well as a unique number of days between the beginning and ending dates of that law change. The fact that the time windows are allowed to vary in length does stretch the framework of the usual first-difference approach. The window of announcement date minus five days until signage date plus five days is chosen for the following reason: markets respond to expectations, and the announcement of a tax law change is the earliest inkling of that change, so expectations should not change before that point; once a law is signed into law, it is virtually certain, so markets should account for the law change by that point.²³ The five day cushion on either side of the beginning and end dates is designed to ensure that I do not miss the effect of treatment by picking too late of an initial start date (for example, if newspaper articles are slow to report lawmakers’ announcements) or too early of an initial end date (for example, if newspapers are slow to report the signing of tax laws). The null hypothesis that the treatment effect is zero raises the question of *whether* markets are efficient with respect to news of state-level corporate income tax changes, and the generously wide window is quite flexible to the question of *when* investors respond, and allows me to pick up the impacts on stock prices of beliefs about upcoming changes that may evolve over time (i.e., given the same starting and ending price, I will estimate the same treatment effect whether investors changed the stock price all in one day or slowly,

²³The argument that a lawmaker’s announcement should be the first inkling of an upcoming tax change is contentious: however, this stance rests on the random walk-like properties of the series of state corporate income tax rates (by definition of a random walk, the expectation is that rates tomorrow will be today’s rates) as well as the fact that most such tax changes are categorized as exogenous relative to state macroeconomic trends, as categorized by Giroud and Rauh (2015) and this author (for financial rate changes).

day after day). A downside of such a long window is that, supposing treatment does not matter at either end of the window, I risk picking up excessive noise and losing significance. Although the panel used to create the first-difference data is not perfectly balanced, the modal firm remains in the panel throughout.

I estimate the model in (2.1), and the regression results are displayed in Table 6. The top panel shows *Treatment* coefficient estimates when all observations from all tax changes are included except for changes in North Carolina, while the middle and bottom panels only include those observations for tax decreases and increases, respectively. North Carolina changes are conservatively omitted for reasons more fittingly explained in section 2.4. All regressions use a basic set of industry controls. Moving from left to right, I begin to include law change-specific trends (to account for possible unique market conditions that occurred during any particular law change), and eventually firm-specific trends. Columns alternate between ordinary least squares regressions and robust regressions.²⁴

Estimates are negative, implying that stock prices move inversely with corporate income tax rates. For all law changes pooled together, *Treatment* coefficient estimates range from around negative one-half to negative three and one-half, and are lose significance as more controls are added. However, moving down the panels, we can see that the results are very heterogeneous for tax decreases versus tax increases. Estimates for tax decreases are insignificant and tend to be between zero and negative one using OLS. The outlier-robust estimates imply

²⁴For validation, I compare the first column of robust regression estimates with a column using median absolute deviation regressions (also called least absolute deviation, quantile or MAD regressions), and the results are very similar, suggesting that these outlier-sensitive methods achieve a similar effect. As the MAD technique requires excessive computing power for some specifications relative to the benefit bestowed by these additional checks, I do not display the MAD results.

TABLE 6.
First-Difference Regression Results

	(1) Least Squares	(2) Robust Regression	(3) Least Squares	(4) Robust Regression	(5) Least Squares	(6) Robust Regression
All Changes:						
Treatment	−3.40** (1.51)	−2.28** (0.97)	−1.66 (1.25)	−1.37 (0.84)	−0.51 (1.34)	−0.87 (0.86)
R^2	0.0953		0.3463		0.4399	
Obs.	11,787	11,787	11,787	11,787	11,787	11,786
Decreases Only:						
Treatment	−0.81 (1.23)	−1.00 (0.95)	−0.19 (0.99)	−0.62 (0.84)	−0.04 (1.11)	−0.75 (0.86)
R^2	0.2497		0.4256		0.5144	
Obs.	7,027	7,027	7,027	7,026	7,027	7,024
Increases Only:						
Treatment	−25.75*** (7.79)	−17.92*** (3.68)	−18.55* (10.18)	−10.23*** (3.13)	−18.73* (10.49)	−10.68*** (3.51)
R^2	0.0998		0.2664		0.4621	
Obs.	4,760	4,760	4,760	4,760	4,760	4,758
Industry Controls	Y	Y	Y	Y	Y	Y
Law Change	Y	Y	Y	Y	Y	Y
Specific Trends	N	N	Y	Y	Y	Y
Firm Specific	N	N	N	N	Y	Y
Trends	N	N	N	N	Y	Y
Standard	Clustered	See	Clustered	See	Clustered	See
Errors	by State	Footnote	by State	Footnote	by State	Footnote

This table presents results of first-difference regressions of log returns of bank stocks on treatment sizes, where treatment is the log size of a tax change in a state multiplied by the proportion of a firm's operations in the state undergoing the tax change (thus, coefficient estimates on the Treatment variable can be thought of as elasticities of stock prices on tax law changes). Regressions are taken over the time period of five days prior to the first news date to five days after the bill was signed into law. The top panel shows regression results when all tax law changes are pooled together; the bottom two panels show results only for tax decreases and tax increases, respectively. Robust regressions have no convenient R^2 analogue, though all tests shown yield F-tests with p values less than 0.0002. Robust regressions are based on an initial screening for outliers with Cook's distance of greater than one, followed by Li's method of following Huber iterations with biweight iterations in order to determine observation weights (minimizing the influence of outliers), before finally running a weighted regression (Li, 1985). Standard errors for robust regressions are calculated by using the correction suggested by Street et al. (1988). Industry controls include measures of number of banks in the US, number of housing starts, and the federal funds rate. Coefficient estimates marked with one, two, and three stars are significant and the one, five, and ten percent levels, respectively.

stock prices changes that are, on average, smaller than a simple predicted change in future tax expenses would suggest.

Estimates for tax increases only tend to be extremely large and significant (despite employing many fewer observations). They range from an elasticity of about negative twenty-five using OLS and the fewest controls, but begin to settle to about negative ten when more controls are included and outliers are accounted for. A decline in stock price of about ten percent for a one percent increase in corporate tax rates seems severe.

Many models in the financial literature, particularly the CAPM, isolate the portions of stock returns that are in excess of the risk-free portion of returns (i.e. the return one could earn investing in one-month Treasury bills). The CAPM also leverages the relationship that stocks will have with broader movements in the stock market as a whole, adjusting for risk. I slightly modify the model in (2.1) to be more unified with the CAPM, by including these concepts. This yields

$$\log(1 + \Delta P_{i,t}) - \log(1 + RF_t) =$$

$$\delta Treatment_{i,t} + \alpha + \beta(\log(1 + \Delta P_{Mkt,t}) - \log(1 + RF_t)) + \Delta X'_{i,t} \gamma + \Delta \epsilon_{i,t}, \quad (2.2)$$

where RF_t is the return on one-month Treasuries during timespan t , and $\Delta P_{Mkt,t}$ is the value-weighted market return of all common stocks on the NYSE, NASDAQ, or AMEX/NYSE American stock exchanges. Consistent with other uses of the CAPM in the literature, other controls are often added.²⁵ Firms may have individual intercepts α_i , or these subscripts may be suppressed. The β s represent financial

²⁵See, for example, Fama and French (1993). While the CAPM result is derived from theory and the application of other control variables is not, such variables are included in order to make up from some of the empirical shortcomings discovered when the CAPM is applied to actual stock returns.

β s. These β s may be firm-specific (in which case they are a moving average of the previous five years' β s), or these subscripts may be suppressed. I run these with both firm-specific subscripts and without; I present the results with subscripts suppressed, as the inclusion of firm-specific subscripts does not meaningfully change results. All in all, these regressions are another set of first-difference regressions, with steps taken to conform to the CAPM framework.

These CAPM-style regression results are displayed in Table 7. These results are almost identical to those of Table 6, suggesting that the results of the original, “arb-style” first-difference regressions are not driven by the inclusion of the risk-free portion of stock returns on the left hand side, and they are not driven due to ignoring the correlation of individual stock returns with the market. Once again, outlier-robust estimates for tax decreases hover around negative ten.

One of the identifying assumptions underlying these sets of regressions is the stance taken that a bank's net income can be proportionally allocated to the different states in which it has branches, based on the proportion of branches in each state (taking care to heed nexus throwback rules, as previously mentioned). Readers may have concerns that this is invalid, for instance, if the bank operates in one higher income state and one lower income state, or in agricultural versus urban areas. I test this identifying assumption by running the regressions using only companies that are located within one state. The results (not shown) are largely unchanged.

I will now test the validity of these estimates with placebo tests, in order to make sure that the regressions were not spurious. If placebo tests suggest that these large elasticities were due to the tax changes and not other factors, then the large size of these estimates bears some discussion.

TABLE 7.
CAPM-Style Regression Results

	(1) Least Squares	(2) Robust Regression	(3) Least Squares	(4) Robust Regression	(5) Least Squares	(6) Robust Regression
All Changes:						
Treatment	-3.72** (1.43)	-2.27** (0.97)	-1.67 (1.25)	-1.37 (0.84)	-0.51 (1.34)	-0.86 (0.85)
R^2	0.0825		0.3365		0.4314	
Obs.	11,787	11,787	11,787	11,787	11,787	11,783
Decreases Only:						
Treatment	-1.22 (1.48)	-1.16 (0.96)	-0.19 (0.99)	-0.63 (0.84)	-0.05 (1.12)	-0.74 (0.86)
R^2	0.2300		0.4193		0.5090	
Obs.	7,027	7,027	7,027	7,027	7,027	7,001
Increases Only:						
Treatment	-23.74*** (8.39)	-17.83*** (3.68)	-18.56* (10.18)	-10.17*** (3.13)	-18.71* (10.48)	-10.65*** (3.52)
R^2	0.0429		0.2416		0.4440	
Obs.	4,760	4,760	4,760	4,760	4,760	4,732
Industry Controls	Y	Y	Y	Y	Y	Y
Law Change						
Specific Trends	N	N	Y	Y	Y	Y
Firm Specific						
Trends	N	N	N	N	Y	Y
Standard	Clustered	See	Clustered	See	Clustered	See
Errors	by State	Footnote	by State	Footnote	by State	Footnote

This table presents results of regressions of log risk premia (log stock returns minus log risk free returns) of bank stocks on market premia and treatment sizes, where treatment is the log size of a tax change in a state multiplied by the proportion of a firm's operations in the state undergoing the tax change (thus, coefficient estimates on the Treatment variable can be thought of as elasticities of stock prices on tax law changes). Regressions are taken over the time period of five days prior to the first news date to five days after the bill was signed into law. The top panel shows regression results when all tax law changes are pooled together; the bottom two panels show results only for tax decreases and tax increases, respectively. Robust regressions have no convenient R^2 analogue, though all tests shown yield F-tests with p values less than 0.0001. Robust regressions are based on an initial screening for outliers with Cook's distance of greater than one, followed by Li's method of following Huber iterations with biweight iterations in order to determine observation weights (minimizing the influence of outliers), before finally running a weighted regression. Standard errors for robust regressions are calculated by using the correction suggested by Street, Carroll, and Ruppert (1988). Industry controls include measures of number of banks in the US, number of housing starts, and the federal funds rate. Coefficient estimates marked with one, two, and three stars are significant and the one, five, and ten percent levels, respectively.

Placebo Regressions

I rerun the regressions of the previous section by taking the actual treatment values, but applying them to stock returns on different dates. I shift the dates of the stock return timeframes t backwards one, two, three, four, and five years, and forwards one, two, three, four, and five years. Because tax events did not happen in these other time periods, there should be no relationship between these non-contemporaneous returns and the actual tax changes that occurred at the actual time of the tax law change.

Placebo regressions for the CAPM-style regressions are shown in Table 8 (for tax law changes shifted backwards in time) and Table 9 (for tax law changes shifted forwards in time). Placebo regressions for the “arb-style” regressions look similar. These placebo regressions exclude North Carolina changes, as one particular North Carolina change tends to erroneously drive significant placebo results. All North Carolina changes are dropped even though only one is problematic. To be consistent, the North Carolina changes are excluded therefore in the baseline regressions.

In general, the placebo regressions on these two tables tend to not show significance where they should not: of the sixty placebo regressions run, only seven are significant at the ten percent level or higher, roughly what one might expect. Of the significant results, most occur when the treatments are artificially shifted forward in time. It is more reassuring that significance is found less before these tax changes come into being; there is no reason why a tax change might not have a delayed effect (the most significant and large estimate occurs for tax decreases, one year after the true treatment). On the other hand, the placebo regressions, even when insignificant, occasionally show large magnitudes, suggesting that the baseline

TABLE 8.
Placebo Tests, Treatment Shifted Backwards in Time

	Minus Five Years	Minus Four Years	Minus Three Years	Minus Two Years	Minus One Year
All Changes:					
Least Squares	-2.48 (1.70)	-0.27 (1.74)	0.21 (2.28)	2.27 (2.53)	0.50 (1.99)
Robust Regression	-0.52 (1.23)	1.90* (1.12)	0.13 (1.06)	1.09 (1.04)	-0.67 (0.93)
Decreases Only:					
Least Squares	-1.81 (1.74)	0.99 (2.37)	0.18 (2.50)	2.41 (3.05)	-0.37 (1.96)
Robust Regression	-0.19 (1.44)	2.79** (1.32)	-0.04 (1.21)	1.11 (1.11)	-1.11 (0.92)
Increases Only:					
Least Squares	0.49 (3.51)	0.84 (3.31)	1.68 (1.77)	-4.97 (5.42)	-0.29 (2.84)
Robust Regression	-1.74 (3.41)	1.67 (3.22)	0.04 (3.26)	-0.74 (3.27)	0.37 (3.46)

This table presents results of placebo regressions of log returns of bank stocks (net of the risk-free component) on log treatment sizes. These placebo regressions use the actual treatment sizes, while changing the time frame of actual law changes to fictitious time frames. These time frames take the original time frames for each law change and shift them back in time by one, two, three, four, and five years. The top panel shows regression results when all tax law changes are pooled together; the bottom two panels show results only for tax decreases and tax increases, respectively. Robust regressions are based on an initial screening for outliers with Cook's distance of greater than one, followed by Li's method of following Huber iterations with biweight iterations in order to determine observation weights, before finally running a weighted regression. All regressions control for the number of trading days in each window as well as the change in the federal funds rate. Standard errors are shown in parentheses. Coefficient estimates marked with one, two, and three stars are significant and the one, five, and ten percent levels, respectively.

TABLE 9.
Placebo Tests, Treatment Shifted Forwards in Time

	Plus One Year	Plus Two Years	Plus Three Years	Plus Four Years	Plus Five Years
All Changes:					
Least Squares	-1.13 (2.11)	-0.15 (0.85)	-2.42** (1.15)	-2.80 (3.47)	0.52 (0.91)
Robust Regression	-0.54 (0.85)	-1.03 (0.82)	-1.74** (0.88)	-0.20 (0.95)	0.67 (0.93)
Decreases Only:					
Least Squares	-1.49 (2.49)	-0.75 (0.65)	-2.29** (1.06)	-2.80 (3.89)	0.23 (1.00)
Robust Regression	-0.65 (0.87)	-1.11 (0.84)	-1.43 (0.91)	-0.18 (1.03)	0.41 (1.02)
Increases Only:					
Least Squares	-6.68 (4.04)	5.76 (3.75)	0.89 (5.99)	-0.53 (4.06)	-0.20 (9.19)
Robust Regression	-8.45** (3.41)	2.49 (3.30)	-3.72 (3.52)	-6.63** (3.31)	-6.90 (5.47)

This table presents results of placebo regressions of log returns of bank stocks (net of the risk-free component) on log treatment sizes. These placebo regressions use the actual treatment sizes, while changing the time frame of actual law changes to fictitious time frames. These time frames take the original time frames for each law change and shift them forwards in time by one, two, three, four, and five years. The top panel shows regression results when all tax law changes are pooled together; the bottom two panels show results only for tax decreases and tax increases, respectively. Robust regressions are based on an initial screening for outliers with Cook's distance of greater than one, followed by Li's method of following Huber iterations with biweight iterations in order to determine observation weights, before finally running a weighted regression. All regressions control for the number of trading days in each window as well as the change in the federal funds rate. Standard errors are shown in parentheses. Coefficient estimates marked with one, two, and three stars are significant and the one, five, and ten percent levels, respectively.

estimates of Tables 6 and 7 could contain bias. Of the placebo tests, the largest magnitude coefficient is 8.45. If I subtract that estimate from the baseline estimates of elasticity for tax increases (which hover around ten percent), this produces a magnitude of negative two percent rather than the ten percent baseline estimate. Even if an elasticity of negative two is a lower bound on the true estimate, it is still a large number.²⁶

Timing

The baseline regressions suggest that investors do heed changes in state-level corporate income tax rates, and that the effects on stock prices are large. Until now, I have taken a fairly neutral stance on *when* during the course of the legislative process do investors adjust prices, only positing that they adjust sometime between when a law change is announced and when it is completed. To further nail down the timing of these price changes, I divide the main window t used in the baseline regressions into three “subwindows:” the first, from the announcement until the date that the bill is introduced into the state legislature; the second, from the introduction date until the date the bill is passed; and finally, from the date of passage until the bill is signed into law. I rerun the regressions within each of these three subwindows. Estimates using the CAPM-style regressions with full controls are shown in Table 10.

I find that, for both tax decreases and increases, the vast majority of the price change occurs in the first subwindow. This is statistically significant for all tax changes, decreases, and increases, when using robust regressions. For tax increases,

²⁶In addition to the placebos shown, I ran placebo tests, using contemporaneous stock returns to the actual tax changes, but with changing the actual treatments to fictitious treatments, using a distribution of random returns that simulates the actual returns (these are not shown). These placebo tests also show a null result.

TABLE 10.
Timing of Market Response

	First Announcement Date Until Bill Introduction Date	Bill Introduction Date Until Bill Passage Date	Bill Passage Date Until Bill Signing Date
All Changes:			
Least Squares	-0.39 (0.77)	-0.20 (0.68)	0.03 (0.20)
Robust Regression	-1.45*** (0.85)	-0.17 (0.31)	0.13 (0.22)
Decreases Only:			
Least Squares	-0.08 (0.75)	0.13 (0.50)	-0.10 (0.18)
Robust Regression	-1.17** (0.58)	0.17 (0.34)	-0.05 (0.22)
Increases Only:			
Least Squares	-9.97** (3.81)	-8.82 (10.00)	-0.46 (1.20)
Robust Regression	-8.06*** (3.41)	-3.27*** (1.24)	1.19 (1.07)

This table presents results of regressions of log returns of bank stocks (net of the risk-free component) on log treatment sizes, over various subwindows in the tax law changes' legislative histories. The first column shows the estimated elasticity of stock prices on tax rate changes during the time between when an upcoming bill is first announced until the bill is introduced, the second column shows the estimated elasticity of stock prices on tax rate changes during the time between when a bill is introduced until it is passed, and the third column shows the estimated elasticity during the time between when a bill is passed and the bill is signed by the governor. These regressions are otherwise the same as the CAPM-style regressions with full controls. This table shows how much of the total response occurred during each subwindow. The top panel shows regression results when all tax law changes are pooled together; the bottom two panels show results only for tax decreases and tax increases, respectively. Robust regressions are based on an initial screening for outliers with Cook's distance of greater than one, followed by Li's method of following Huber iterations with biweight iterations in order to determine observation weights, before finally running a weighted regression. All regressions control for the number of trading days in each window as well as the change in the federal funds rate. Standard errors are shown in parentheses. Coefficient estimates marked with one, two, and three stars are significant and the one, five, and ten percent levels, respectively.

a significant portion - between about thirty and forty-five percent, depending on the specification - occurs in the second subwindow, and this is significant using robust regression. No significant price change occurs in the final subwindow. The null result for the third subwindow makes sense, as governors signed bills in all but one instance (where the law was vetoed, returned to chamber, and was passed with a veto override). If investors know that these types of laws are virtually certain once passed, there should be little to no price change in this final subwindow.

The significant stock decline (at least using robust regression) in the second subwindow for tax increases suggests a potential opportunity for market arbitrage. If roughly three-tenths of the decline in stock price due to a tax increase comes after an impending bill is introduced, investors can seize advantage of this market opportunity. Perhaps this is difficult to do, as lawmakers' popularity is likely positively related to tax cuts, so lawmakers may try to announce upcoming tax increases in a low-key manner, making upcoming tax increases less salient.

Discussion

In this section, I discuss the size and asymmetry of my estimated elasticities. I also address concerns about the external validity of my results.

Were the banks to continue on their projected growth path (see the appendix for further details about measuring and estimating said growth), and the only thing to change were the size of their tax expenses (due to the changes in rates), and were these stocks properly valued before and after the law changes (within a reasonable amount of error), then I would estimate elasticities of around negative one, for both tax decreases and increases. That the estimates for tax increases are larger in magnitude than unity suggests that more is going on than just a

change in future tax expenses that is proportional to future income, assuming a current growth path. An analysis of the competitive nature of the regional banking industry suggests some possible reasons for the larger magnitudes.

Regional banking is a competitive industry. To most small business borrowers and mortgage borrowers, there is little to distinguish one loan from another, except for loan rates and terms. If we assume that homeowners are inflexible in the lengths of their loans (for example, they seek out thirty year loans), then they should distinguish between possible loan products by choosing the loan with the lowest rate. There may be elements of monopolistic competition in this industry based on location (that the banks stay small suggests there is a business advantage to staying community oriented), so borrowers may receive some kind of good feeling from borrowing locally. However, it is likely that most borrowers will seek out the lowest borrowing rate. If we take this industry to be perfectly competitive, then banks offering too high rates will make no loans; thus it is incumbent upon banks to offer the loans at the lowest rate possible. Under the perfectly competitive assumption, banks will offer loans at cost.

Taxes factor into the cost structure of these banks, so an increase in taxes will force banks to raise loan rates for all banks affected by the tax. If rates are decreased, banks will lower rates, lest their competition lower rates and take all of the market share. However, with state-level corporate tax changes, not all banks are affected - only banks in that state. In perfect competition, a higher rate in one state and a lower rate in another state will cause all of the borrowers in the high rate state to seek out lower rate loans in the lower tax state.

Of course, this does not happen exactly. The community orientation of these banks means that there is some cost - travel costs or perhaps psychic costs of

seeking out a bank not native to one's area - to switching banks. If these costs are heterogeneous, then a tax change in one state that causes loan rates to change may induce some people to switch banks, while others may not. In this case, a higher tax will not cause a complete erosion of market share, but a significant one.

Furthermore, state versus state competition is not the only channel for the erosion of market share for loans issued by regional banks. Not all banks are C corporations, and thus not all banks pay corporate income taxes. Beginning in 1997, banks could elect to become S corporations, which are treated like partnerships for tax purposes, meaning they are not taxed at the corporate level and income flows through to the individual shareholders, who are taxed on their S corporation income based on personal income tax rates. The absence of corporate taxation is a strong incentive for banks to adopt the S corporation form. However, converting from a publicly traded C corporation to an S corporation may be costly. S corporations are limited in the number of shareholders they may have, and publicly traded stocks are held by a wide variety of shareholders. Converting to an S corporation means that a publicly traded company will first have to buy back outstanding shares on the stock market, which will drive up the price of shares, making the process expensive. Nonetheless, the last two decades have seen a steady rise in the percentage of banks adopting the S corporation form, from roughly six percent in 1997 to roughly thirty-six percent in 2014 (De and Mehran, 2014). This occurred during a period when the number of banking institutions was steadily decreasing each year. The bulk of the increase in S corporation banks was made up by former C corporations converting into S corporations, and not by the emergence of new establishments.

The lack of corporate taxes on S corporations means that S corporation banks do not have to raise loan rates when corporate taxes increase, unlike C corporations. In addition to the competition between banks in different states, banks within states featuring corporate taxes face different cost structures depending on their organizational form. If a state's taxes increase, forcing C corporations to raise rates, S corporations can respond by holding their loan interest rates steady, which should erode some of the market share of the C corporation banks. Thus there are two sorts of competition by which publicly traded C corporation banks can lose market share in the face of higher taxes.

Why is the magnitude for tax increases so much larger than the essentially null result for tax decreases? An asymmetric result is par for the course in the literature on the effects of state corporate income taxes. For example, Ljungqvist et al. (2017) find that firms reduce risk-taking only in response to tax increases, Ljungqvist and Smolyansky (2014) find that employment and employment income generally only respond to tax increases, and Heider and Ljungqvist (2015) find that firms' leverage decisions respond only to tax increases. Taxation forces a natural asymmetry between corporate gains and losses, because the government shares in a firm's gains but not their losses. Additionally, debt covenants in loan agreements create nonlinearities that make losses potentially more impactful than gains. However, a unique story may be taking place here.

It is possible that markets are efficient in that they understand the proper direction of the price change, but inefficient in that they do not understand the magnitude. Markets can possibly (relatively) overreact to tax increases while (relatively) underreacting to tax decreases. However, a more compelling explanation is that markets may take a tax decrease as a signal of future things

to come in that state, not necessarily with tax rates, but with the business environment in general. Even if one takes the stance that regional banks do not hold much sway over their state legislatures, it is hard to imagine a world where business as a whole does not have some kind of control over the state government. There is likely some general sense that the business community possesses some degree of regulatory capture over the state government. Markets may interpret tax increases as a harbinger that the business community as a whole has lost regulatory capture over the state legislature, meaning that more bad news (from businesses' perspective) may be coming in the future. Thus, an elasticity of negative ten may have a component that is based on the erosion of market share and a component that represents future non-business-friendly events.

The placebo regressions speak to the internal validity of this research, however, readers may have concerns that these results are not externally valid. One nice feature of the banking industry as the subject of research is that banking is connected to most other industries, so the usual concerns about the idiosyncrasies of this industry are less pressing than they are for some other industries. Banks are unusual, in general, in that they have high debt to equity ratios. This sample of firms is also unusual in that the firms are not national, unlike most publicly traded firms. It remains a question whether the elasticities estimated are representative of other local industries, other banks, or firms in general. One nice validation of results comes from the recent corporate income tax change at the Federal level, where corporate rates were decreased from thirty-five percent to twenty-one percent in 2017. The Federal Deposit Insurance Corporation reported that, nationally, banks saw an increase in profits of over twenty-seven percent in the quarter after this large change. This corresponds to an elasticity (of profit, not stock price,

to tax rates) of roughly -1.25, in the ballpark of the elasticity estimates for tax decreases when accounting for outliers.

Conclusion

In this paper, I use the wisdom of the markets to estimate the change in capitalization value of publicly traded regional banks due to changes in state corporate tax rates. I do this by leveraging the heterogeneous set of changes in state corporate tax rates and by employing a variety of first-difference regressions. My results suggest that, on average, these changes in capitalization values are far greater than a 100% burden of the corporate tax on shareholders would suggest. Estimates of the elasticities of stock prices to corporate income tax rates are in the range of about negative one to zero for tax decreases and roughly negative ten for tax increases, when employing the most controls and accounting for outliers (where the negative sign implies an inverse relationship between tax rates and stock prices). Placebo regressions suggest that these relationships are not spurious but that they may contain bias.

The reason for the large magnitude is speculative, but may be driven by the intense competition in the banking industry. Higher taxes in some states may cause erosion of the market share of firms in those states by banks in lower tax states, which can sustain lower loan rates. In all states, S corporations, which are not subject to corporate tax, can compete with taxed firms. The asymmetry of results for tax decreases versus increases is typical for the state-level corporate income tax literature. This asymmetry may be driven by investor beliefs about regulatory capture.

The estimated elasticities raise the question of whether markets are correct in their assessments of tax changes. Future work will investigate the changes in future profits in the years following these changes. Given that investors may have potentially over- or underreacted, I will follow up by looking at price movements after these changes to see if any correction was made. Finally, I will investigate the management compensation schemes of these firms in years following the changes, to see if managers use the windfall of tax decreases to claim superior management and increase their compensation.

CHAPTER III

MEASURING SYSTEMATIC WAGE MISREPORTING BY DEMOGRAPHIC GROUPS

This dissertation includes previously unpublished co-authored material. It is co-authored with John Voorheis of the US Census Bureau and Caroline Weber of the University of Washington. I wrote the majority of the chapter, with John Voorheis providing important work on the Data section and with institutional details, and Caroline Weber providing edits. I wrote the majority of the computer code used to create this chapter, with John Voorheis contributing code for the visualizations of income misreporting within groups across income levels, and Caroline Weber providing valuable supervision and code writing lessons and examples. I conceived this chapter, which is an offshoot of another ongoing project. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. The statistical summaries reported in this paper have been cleared by the Census Bureau's Disclosure Review Board, release authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

Introduction

Demographic surveys, such as the Current Population Survey (CPS), are a valuable source of data for economists and statisticians. CPS income data are used by economists to estimate the scope and impacts of key labor market phenomena, such as returns to education and the Black-White wage gap, and to generate earnings statistics. When income data are measured with error, statistics can be

less accurate and will be less reliable, which can introduce bias into regression modeling.

Analyses of the methodology of economic observation reveal a number of ways measurement errors can occur, confounding results and inferences when left uncorrected (Morgenstern, 1963; Deaton, 1997). The more commonly acknowledged types of measurement errors assume that the mean of the measurement errors is zero (Bound et al., 2001). Previous studies using matched CPS and administrative data have investigated such zero-mean error structures (Bound and Krueger, 1991; Bollinger, 1998).

However, it is also possible that measurement error may be non-classical, i.e. the errors may have non-zero means. This type of error is often acknowledged as a possibility, although it has been difficult to quantify the degree to which this may present challenges to empirical research. However, by comparing CPS data to an alternate source such as administrative records, we can examine whether and to what degree non-zero mean measurement errors exist in survey data.

In this paper, we examine wage misreporting by matching responses from the CPS Annual Social and Economic Supplement (ASEC) to administrative records from three sources: individual-level Internal Revenue Service (IRS) Form W-2 and Social Security Administration (SSA) Detailed Earnings Record (DER) wage data, and tax unit-level IRS Form 1040 wage data. We utilize the accuracy of IRS wage data amongst wage-earners (individuals with self-employment income are dropped from the sample) to identify wage misreporting to the CPS. The CPS provides detailed sociodemographic information on respondents, including race, ethnicity and educational attainment. These matched datasets allow us to investigate the

roles these demographic characteristics play in CPS wage misreporting and explore heterogeneity of misreporting behavior within demographic groups.

We find evidence of heterogeneity in misreporting, finding consistent and highly significant estimated coefficients for several sociodemographic variables, across model specifications. Wage misreporting shows a strong relationship with age: individuals and tax units underreport wages to the CPS by 1.5 to 2.8 percent for each decade lived, on average. Estimated coefficients on education dummies show that more educated individuals and tax units underreport relatively less to the CPS relative to less educated individuals – college graduates underreport by 1.6 to 2.4 percent less than high school graduates. Within racial and ethnic indicator variables, the estimated reporting gap for Hispanics is consistently significant and robust across specifications, but there is heterogeneity within this group, driven largely by educational attainment. Coefficients for the black indicator variable only become significant and stable once outliers are excluded, also suggesting heterogeneous reporting behaviors within this group. Hispanics underreport by 3.5 to 5 percent more than non-Hispanic whites, while Blacks under-report by 0 to 1.4 percent more than non-Hispanic whites. These relative differences in misreporting behavior do not generally appear to be driven by potential differences in how these groups round income when they report it to the CPS.

These results are important for the computation of racial and ethnic income statistics and for the study of wage differentials due to racial and ethnic status. Studies that examine the history and impacts of racial and ethnic discrimination in labor markets are often built upon CPS income data (Card and Krueger, 1992; Trejo, 1997; Chay, 1998; Peoples and Talley, 2001; Juhn, 2003), Census data (Reimers, 1983; Borjas and Bronars, 1989), or other survey data (Cameron

and Heckman, 2001). In the case of simple ratios, our findings imply that racial and ethnic wage differentials are smaller than previously calculated. Our results also stress the importance of accounting for heterogeneity within groups when investigating racial and ethnic wage gaps.

Our results are also highly relevant to studies on the returns to education. Since Mincer (1958), economists have been estimating the private and social returns to educational attainment. Many of these use survey data (Card, 1993; Kane and Rouse, 1995; Cameron and Heckman, 2001). Our findings imply that simple ratios of education earnings differentials will overstate the returns to education, as college educated individuals have relatively less wage under-reporting.

Our findings add to body of knowledge showing how administrative records can be used to improve survey designs and use of survey data. Where non-zero mean errors arise, researchers can adjust incomes to create more accurate income data. Survey design analysis reveals how data can be made more accurate before it is used elsewhere.

In the next section, we describe the data used and the construction of variables. Then we describe the model and identifying assumptions. Later, we present our main regression results and plots of measurement error distributions, and provide evidence of systematic measurement errors. Then we discuss the possible mechanisms by which systematic measurement errors may occur, and the implications of these errors, particularly for studies of racial wage differentials or returns to schooling. Finally, we conclude and propose avenues for further research.

Data

This project creates a novel dataset by linking data from four major sources: the CPS ASEC, IRS Form 1040 tax returns, IRS Form W-2 wage reports, and the SSA DER.¹ The three administrative records data sources are available for different time periods: the 1040 tax returns are available for tax years 2000-2015, the form W-2 data are available for tax years 2005-2015, and the SSA DER extract is available for tax years 2000-2012.

The monthly basic CPS survey collects employment information from about 70,000 American households, and is the source of the BLS' published unemployment rates. The March ASEC supplement adds detailed household income data for the previous calendar year.² Households are surveyed for two consecutive years, rotating in sample for four months, out of sample for eight months, and then back in sample for the final four months. Recent survey non-response rates have been low relative to some other household surveys, in the range of approximately thirteen percent for the regular monthly CPS and fifteen percent for the ASEC (U.S. Census Bureau, 2015), although it has been increasing. This increase in non-response rates for the ASEC is likely partly due to the additional time investment households put in when completing the lengthier ASEC. The ASEC collects data on most Federally taxable income types, as well as types that are not recorded on Form 1040 (such as child support and welfare assistance).

¹While the two IRS datasets are available for the full population, the SSA DER data at the Census Bureau contains only an extract for CPS respondents. Since our analysis will be conducted using individuals in both the CPS ASEC and the administrative records, however, this distinction will not affect our results in this paper.

²Starting in 2002, the basic monthly CPS sample size was increased from around 60,000 households to around 70,000 households, and the ASEC supplement sample size increased to around 100,000 households.

These data are aggregated up to the household level. The CPS ASEC also collects demographic information such as age, gender, marital status, race and ethnicity, and educational history.

Our second major source of data comes directly from IRS Form 1040.³ Detailed income information for all taxable categories and a few non-taxable categories of income are reported on Form 1040. The IRS has delivered an extract of the universe of 1040 tax returns to the Census Bureau annually since 1998. These extracts contain the universe of tax units, but they do not contain all fields of Form 1040, limiting us to only observe certain line items or composites of line items. These are: wages, dividends, taxable and non-taxable interest income, social security income, rental and royalty income, and total money income (which is equivalent to “total income” on Form 1040). In this paper, we focus solely on wage income. Analyses using the other income concepts can be found in Imboden, Voorheis, Weber (Forthcoming). In addition, the 1040 data includes indicators for various schedules filed with the tax form.

Our third source of data is a subset of the universe of Detailed Earnings Records, which are collected by the Social Security Administration.⁴ These detailed records contain self-employed earnings subject to Medicare taxes (equivalent to all self-employment earnings reported to the Social Security Administration via IRS forms) and wages earned for all CPS respondents that are in the Social Security

³Here, 1040 refers to both the standard 1040 form as well as short form 1040A and easy form 1040EZ.

⁴Unlike the 1040 data, we do not have the universe of DER observations. Rather, the SSA delivered a subset of the DER to the Census Bureau annually beginning in 1991 for all individuals who ever responded to the CPS ASEC. This extract was created by the SSA using a list of SSNs sent by the Census. Before 2003, these SSNs were directly collected. After 2003, the Census sent the SSN associated with the PIK assigned by the PVS process.

system.⁵ Our fourth and related source of data comes from the universe of IRS Form W-2s.⁶ These data report all gross wage income paid to an individual by an employer in a tax year. Since an individual can work multiple jobs in a year, we aggregate unique forms to the individual-year level by summing all wages received by an individual across forms.

CPS ASEC records are matched to IRS and SSA records using the US Census Bureau's data linkage infrastructure. This data linkage infrastructure allows for the linking of individuals across survey and administrative records using anonymous identifiers called Protected Identity Keys (PIKs). PIKs are assigned to individuals in an administrative records, survey or census microdata file using the Person Identification Validation System (PVS). PVS is a probabilistic matching algorithm which uses personally identifiable information (PII) to link individuals to a reference file. Reference files used by PVS are modified versions of the SSA's Numerical Identification File called the Census Numident. The Census Numident is the universe of individuals who have received Social Security Numbers (SSN), and contains PII including the SSN itself, as well as age, date of birth, sex, race and address.⁷ PIKs are invariant across and map one-to-one with SSNs. Once PIKs have been assigned to a file, it is possible to match with any other administrative records or survey records which have been assigned PIKs. We match all CPS ASEC respondents from survey years 2001 through 2013 with IRS 1040 and SSA DER

⁵Virtually all tax units pay into the Social Security system, but some are exempt due to religious objections or waivers.

⁶Form W-2's are the underlying data source for the wages in the DER, for W2 and DER wage amounts are identical in overlapping years.

⁷There are multiple vintages of the Numident reference file, each of which has the best PII information for a given individual.

data for the previous tax year (i.e. matching the 2010 CPS ASEC to tax year 2009 IRS 1040s).

Appendix table A.6 summarizes the evolution of our sample. We drop any person records where PIKs are missing or invalid (about 10 percent of records). Using CPS demographic fields, we create mutually exclusive dummy variables for gender, marital status, education level (less than high school diploma completed, high school diploma completed with no college, some college completed with no degree, and bachelors degree completed or more) and for racial and ethnic groups (Black, Native American, Asian, White Hispanic, and non-Hispanic White).⁸ Form 1040 incomes are based not on individual income data, but rather income data for the tax unit, which may be based on one or two taxpayers, while the W-2 and DER data are individual-level files. Thus we construct two datasets at different levels of aggregation. We match IRS form W-2 and SSA DER administrative records to survey records using PIKs when analyzing W-2 or DER data. When analyzing the form 1040 data, we link individuals in the CPS to the form 1040 data by PIK, and then collapse the individual CPS records to the tax unit so that they can be properly compared to 1040 records.

We drop any data points where relevant income values in the ASEC were imputed or truncated.⁹ Because capital gains and losses were fully imputed throughout the sample, we drop all tax units that filed a Schedule D with their 1040. We drop any individuals and tax units that report positive self-employment

⁸In the current analysis, Blacks, Native Americans and Asians may be of either non-Hispanic or Hispanic ethnicity. In future work, we will explore whether using different group definitions (i.e. defining Hispanics of any race as a single group) affects our baseline results.

⁹CPS item non-responses are imputed using a “hot deck” methodology. For the protection of personal information, large income amounts reported to the CPS are truncated at various thresholds that vary by income type and across years.

income to the IRS (as measured by filing any of Schedules C, E, or F) or the CPS, because there may be confusion between some types of self-employment income and wages. We drop any tax units who are outside of working age or who are considered dependents for tax purposes. We start with 1,034,373 matched tax unit records, and are left with 348,507 records after sample restrictions.¹⁰

Our analysis will focus on the difference between reported wages, salary and bonuses in the CPS and administrative wage data. Despite the differences in income categories and definitions between the CPS, SSA, and IRS data, it is also possible to define other income concepts that are directly comparable between the CPS and the administrative sources, however we leave this to subsequent research. We define two types of wage misreporting – extensive margin misreporting, where no wages are reported in the CPS ASEC for individuals with non-zero wages in linked administrative records and vice versa, and intensive margin misreporting, which we measure using the wage reporting differential, or “gap.” On the extensive margin, the indicator for having wages of record R , $R \in \{\text{CPS}, \text{W-2}, \text{DER}\}$, is defined as

$$I_{ist,R} = \begin{cases} 1, & \text{if } y_{ist,R} > 0 \\ 0, & \text{otherwise,} \end{cases} \quad (3.1)$$

where $y_{ist,R}$ is wage income recorded by source R . On the intensive margin, the wage misreporting gap is defined as

$$G_{ist,Admin} = \log(y_{ist,CPS}) - \log(y_{ist,Admin}) \quad (3.2)$$

¹⁰Further data cleaning documentation is in the appendix.

where $G_{ist,Admin}$ is the wage reporting differential for tax unit/individual i in state s and year t , based on record $Admin \in \{W-2, DER\}$.

Selected descriptive statistics of the final sample are presented in Table 11. Most tax units report wages to both the CPS and IRS—about ninety-four percent report wages to the IRS, while ninety-two percent report wages on the CPS. More wages are reported to the CPS than the IRS (this is found in all comparable income types). However, the mean wage reporting differential is positive, meaning that IRS wages are greater than CPS wages for the average gap (this is also true for all other income types except interest plus dividends). Differences between means of levels and mean reporting differentials across the unrestricted sample could be driven by outliers.

Statistics may differ from those of the general US taxpaying population where CPS sampling methods and response rates differ from a sample of all taxpaying Americans, or where our data cleaning methods systematically removed members of certain groups from our original sample. Most notably, the CPS oversamples Hispanic Americans.¹¹ Consequently, the proportion of a racial/ethnic group in our data that differs most from that of the general US population of Hispanics, with more than a two percentage point difference in representation. Our sample also contains more females than the general population, perhaps because of higher response rates by females and/or higher incarceration rates among males.¹²

¹¹It is possible to reweight the CPS sample weights by the inverse probability of linkage to create weights to target to the US noninstitutional population.

¹²Prisoners and other institutionalized people are not in the CPS sample frame.

TABLE 11.
Descriptive Statistics

Demographic Variable	Observations	Mean	Standard Deviation	Median
Survey Year	641,000	2009	4.31	2009
Age	641,000	41.18	9.960	41
Female	641,000	0.5371	0.4986	1
Married	641,000	0.6309	0.4826	1
White Non-Hispanic	641,000	0.6538	0.4757	1
Black	641,000	0.1200	0.3249	0
White Hispanic	641,000	0.1497	0.3568	0
Asian	641,000	0.0623	0.2418	0
Native American	641,000	0.0141	0.1181	0
Less Than High School	641,000	0.1003	0.3004	0
High School Only	641,000	0.3101	0.4626	0
Some College	641,000	0.2984	0.4575	0
Bachelor's Degree	641,000	0.2911	0.4543	0
Income Variable	Observations	Mean	Standard Deviation	Median
CPS Wages	641,000	36,730	45,690	30,000
W-2 Wages	641,000	12,700	38,390	0
DER Wages	641,000	27,650	46,070	20,450
1040 Wages	641,000	58,960	69,180	45,320
Has CPS Wages	641,000	0.8670	0.3396	1
Has W-2 Wages	641,000	0.3120	0.4633	0
Has DER Wages	641,000	0.7145	0.4516	1
CPS W-2 Wage Gap	191,000	0.0287	0.5919	0.0112
CPS DER Wage Gap	437,000	0.0305	0.5825	0.0107

Source: CPS ASEC, IRS W-2, SSA DER and IRS 1040s, 2001-2016

This table presents descriptive statistics for the major explanatory variables and dependent variables used in the regressions. Figures are rounded to four significant digits. Medians are interpolated over at least fifty observations. All racial and ethnic categories and educational categories are mutually exclusive. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

Identification and Model Specification

Our central identifying assumption is that IRS wage data are sufficiently accurate that wage reporting differentials are attributable to measurement error in CPS recorded wages. Hereafter, “wages” refers specifically to wages as defined by the IRS. The IRS includes most employer to employee remuneration in this category, including salaries, tips, commissions, and bonuses.¹³ Note that only taxable wages are reported on IRS form 1040.¹⁴

In the US, wages are subject to employment taxes (Social Security, Medicare, and Federal unemployment insurance taxes) and income taxes, and are generally subject to employer withholding of Federal income and employment taxes. After year end, employers send duplicate wage earnings statements to their employees and to the IRS directly. The duplicate statements allow the IRS to verify the amount of wages earned by the tax unit. Discrepancies increase the probability that a tax unit will be audited, which should discourage misreporting (Allingham and Sandmo, 1972; Slemrod, 2007; Slemrod and Bakija, 2008). From audit-based estimates, only one percent of wage income in the US is underreported to the IRS (Internal Revenue Service, 2016). The IRS attributes the accuracy of reported wages, relative to other income types, to greater withholding and information requirements (Internal Revenue Service, 2016).

Using wages from the administrative records as a benchmark, we attribute differences between IRS/SSA and CPS reported wages to measurement error on the CPS. We interpret mismatch in reporting between CPS and administrative

¹³The CPS also explicitly includes salaries, tips, commissions, and bonuses in their definition of wages.

¹⁴Pre-tax payroll deductions, e.g. for employer sponsored health insurance or tax-advantaged retirement plans, are not included in taxable wages as a general rule.

records sources – e.g. reporting no wages in the CPS when a non-zero W-2 or DER wage report exists – as misreporting on the extensive margin. Similarly, we can interpret G_{ist} as the percent of wages over-reported (if G_{ist} has a positive sign) or underreported (if G_{ist} has a negative sign) to the CPS on the intensive margin. IRS and CPS wage data are comparable, as they measure the same type of income earned over the same time period. By design, they are reported close together in time: wages earned during one tax year are typically reported to the ASEC in March and to the IRS before mid-April of the following year (U.S. Census Bureau, 2006). We regress indicators for having wages in of one record type on a rich set of demographic information and fixed effects, conditional on individuals reporting wages to a complementary source. We also regress the differentials between IRS reported wages and CPS reported wages on a rich set of covariates and fixed effects, conditional on households reporting some wages to both the IRS and CPS. Coefficient estimates will not indicate the degree of misreporting *per se*, but rather will ascribe the degree of *relative* misreporting.

With the IRS data assumed as our benchmark, we are mostly focused on negative values of G_{ist} ; that is, when income is underreported to the CPS. However, we will also observe positive values of G_{ist} , which can occur due to random measurement error or individuals choosing to over-report wages to the CPS.

There also may be differences in misreporting at different parts of the wage distribution, shown by Brummet et al. (2018) to occur in the Consumer Expenditure Survey. We will consider multiple explanations for the data patterns we observe later in this chapter.

Demographic characteristics, tax unit characteristics, and state and year effects are likely to partially explain wage misreporting to the CPS. We

assume that these demographic characteristics are exogenous to wage reporting differentials. CPS designs vary across years, as do the timing between survey dates and tax deadlines.

Our extensive and intensive margin models take the form

$$Y_{ist} = \alpha + \delta Demogs_{it} + \phi FE_{st} + \epsilon_{ist}, \quad (3.3)$$

where Y_{ist} is an indicator for whether non-zero wages were reported in the CPS or administrative data for the extensive margin models, and equal to the previously defined wage gap G_{ist} in the intensive margin models. $Demogs_{it}$ is a vector of demographic dummy variables, FE_{st} is a vector of state and year fixed effects, and ϵ_{ist} is an error term. Demographic characteristics are determined by either the individual or the characteristics of the primary earner in the tax unit, and include age, gender, marital status, racial and ethnic group, and education level. Card and Krueger (1992) show the interactions between racial characteristics and the effects of schooling, and Cameron and Heckman (2001) show the interactions of gender, racial and ethnic characteristics, and schooling. Thus we also include interactions between education and race/ethnicity and interactions between gender and race. State fixed effects include fixed effects for US territories. We cluster standard errors at the primary sampling unit (PSU) level for all regressions.

Results

Extensive Margin Analysis

We begin by examining the extensive margin of reporting, by examining the degree to which demographic characteristics are associated with 1) reporting

positive wages on the CPS ASEC, given administrative records reports of positive wages in the W-2 or DER data or 2) whether positive wages exist in administrative records, given a report of positive wages in the CPS ASEC. That is, we estimate regressions of form in equation 3.3, where Y_{ist} is an indicator variable equal to one for positive wage reports (in either the CPS ASEC or administrative records) or zero otherwise. In this set up, positive coefficients are interpreted as a relative decrease in extensive margin misreporting, while negative coefficients are interpreted as a relative increase in extensive margin misreporting.

Table 12 summarizes the results of these extensive margin regressions. In this table, the first two columns capture extensive margin misreporting using the DER data as a reference, while the final two columns use the W-2 data. The first and third columns report the results for regressions using an indicator for positive administrative records wages as a dependent variable for the subset of individuals with CPS wages. If individuals do not have administrative wages, but do have CPS wages, then this suggests individuals claimed wages on the CPS when, in fact, they had none. Columns two and four perform the opposite task, reporting regression results using an indicator for positive CPS wages as a dependent variable, for the subset of individuals with positive administrative records wages.

Across the four extensive margin misreporting concepts, there is substantial evidence of heterogeneity in misreporting along multiple demographic dimensions – virtually all of the coefficients for our demographic characteristics of interest are statistically significant. In general, these results point towards a gradient in misreporting along education lines – college educated individuals are less likely to misreport wages on the extensive margin, but individuals with less than a high school education are more likely to misreport. There is also important

TABLE 12.
Extensive Margin Linear Probability Regressions

Dependent Variable	(1) hasderwages	(2) hasCPSwages	(3) hasw2wages	(4) hasCPSwages
Married	-0.0136*** (0.0019)	-0.0006 (0.0023)	-0.0052*** (0.0007)	-0.0050* (0.0029)
Female	0.0048*** (0.0012)	-0.0360*** (0.0031)	-0.0001 (0.0007)	-0.0412*** (0.0040)
Black	-0.0005 (0.0015)	-0.0292*** (0.0027)	0.0008 (0.0008)	-0.0263*** (0.0039)
Asian	-0.0117*** (0.0021)	-0.0191*** (0.0030)	-0.0033*** (0.0012)	-0.0178*** (0.0035)
Native American	-0.0121 (0.0095)	-0.0302** (0.0122)	-0.0091 (0.0061)	-0.0357*** (0.0135)
White Hispanic	-0.0367*** (0.0026)	-0.0159*** (0.0025)	-0.0041*** (0.0008)	-0.0128*** (0.0037)
Less Than High School	-0.0397*** (0.0037)	-0.0325*** (0.0032)	-0.0090*** (0.0016)	-0.0332*** (0.0042)
Some College	0.0218*** (0.0025)	0.0148*** (0.0026)	0.0034*** (0.0009)	0.0171*** (0.0033)
Bachelor's Degree	0.0192*** (0.0020)	0.0294*** (0.0025)	0.0041*** (0.0008)	0.0292*** (0.0031)
Age	0.0011*** (0.0001)	0.0006*** (0.0001)	0.0001** (0.0000)	0.0006*** (0.0002)
Observations	381,000	296,000	381,000	168,000
Adjusted R^2	0.863	0.015	0.959	0.017
Conditional on	Has CPS wages	Has DER wages	Has CPS wages	Has W-2 wages
State Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Interaction Terms	N	N	N	N

Source: CPS ASEC, IRS W-2and SSA DER, 2001-2016

This table presents the results of linear probability model regressions of indicator variables of wage reporting on key demographic indicators at the individual level. These indicators take on a value of one when the individual reports wages to the respective source, e.g. hasCPSwages takes on a value of one when the individual reports positive wages to the CPS. Inclusion in each regression is conditional on having wages in the opposite source, e.g. only observations with positive CPS wages are included in the regression for hasderwages. The baseline tax unit is male, White, unmarried, and has a high school diploma but no college. Standard errors are clustered at the sampling unit level. For the reported coefficients, those marked with three stars are significant at the one percent level, those marked with two stars are significant at the five percent level, and those marked with one star are significant at the ten percent level. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

heterogeneity across racial and ethnic groups – all non-whites are more likely to misreport wages on the extensive margin relative to non-Hispanic whites.

There is some interesting heterogeneity across the types of extensive margin misreporting. For instance, Blacks have no statistically significant relative difference in having DER wages conditional on reporting CPS wages, but Blacks are 2.9 percent more likely to misreport having wages on the CPS conditional on having DER wages. Similarly, married individuals are not statistically more likely to misreport CPS wages than non-married individuals, but are 1.4 percent less likely to have DER wages than unmarried individuals if they report CPS wages. Hispanics, however, are both more likely to misreport CPS wages and less likely to have DER wages if they report CPS wages relative to non-Hispanic Whites.

On the other hand, there is a very consistent pattern across educational groups, suggesting a gradient where more educated people are more likely to report wages accurately on the extensive margin. People with less than a high school diploma are about 4 percent less likely to have DER wages if they report CPS wages and are 3.2 percent more likely to misreport having CPS wages relative to high school graduates. On the other hand, college graduates are 1.9 percent more likely to have DER wages if they report CPS wages and are 2.9 percent less likely to misreport CPS wages if they have DER wages, again relative to high school graduates.

Individual Level Analysis

We now turn to analysis of the demographic correlates of intensive margin misreporting. We begin by examining the relationship between demographic characteristics and the level of misreporting at the individual level by analyzing

regressions using the log difference between DER wages and CPS wages and between W-2 wages and CPS wages, respectively, as dependent variables. Note that these wage gaps can be positive or negative. We can interpret the coefficients in terms of *relative* misreporting. Thus a positive coefficient is interpreted as a decrease in under-reporting (or increase in over-reporting) *relative* to the reference category, while a negative coefficient can be interpreted as an decrease in over-reporting (or increase in under-reporting) *relative* to the reference category. These summary measures are useful for understanding, for example, how accurate or inaccurate existing estimates of various wage gaps are. However, it is not fully satisfying as we would like to know whether they different groups are different in their under reporting behavior, over-reporting behavior or both. To understand this, we will turn to some figures later on in this section.

Table 13 reports results from these intensive margin individual level regressions. The first two columns of this table report results from an unrestricted sample, while the final two columns report results from a trimmed sample which excludes the top and bottom five percent of the administrative wage distribution, since misreporting may be very different at the tails of the distribution than in the middle. The first and third columns report results using the wage gap between CPS and DER wages, while the second and fourth columns report results using the wage gap between CPS and W-2 wages. All regressions include interactions between education and race/ethnicity categories, however we report only interactions which are statistically significant in a majority of models, as well as terms which involve interaction with the Black variable, as these may be of interest for the black/white wage gaps.

TABLE 13.
Individual Level Regressions

Dependent Variable	(1) CPSderwagegap	(2) CPSw2wagegap	(3) CPSderwagegap	(4) CPSw2wagegap
Married	0.0042 (0.0061)	0.0050 (0.0079)	-0.0005 (0.0028)	-0.0000 (0.0033)
Female	0.0191** (0.0079)	0.0279** (0.0116)	0.0029 (0.0035)	0.0068 (0.0047)
Black	0.0048 (0.0143)	-0.0013 (0.0209)	-0.0123** (0.0052)	-0.0143** (0.0072)
Asian	-0.0110 (0.0240)	0.0115 (0.0276)	-0.0282** (0.0113)	-0.0118 (0.0137)
Native American	0.0062 (0.0398)	0.0456 (0.0579)	0.0065 (0.0218)	0.0288 (0.0254)
White Hispanic	-0.0504*** (0.0118)	-0.0459*** (0.0155)	-0.0353*** (0.0060)	-0.0386*** (0.0085)
Less Than High School	-0.0212 (0.0149)	-0.0162 (0.0186)	-0.0044 (0.0064)	-0.0024 (0.0087)
Some College	0.0128 (0.0095)	0.0215* (0.0115)	0.0137*** (0.0039)	0.0187*** (0.0050)
Bachelor's Degree	0.0165** (0.0073)	0.0212** (0.0098)	0.0171*** (0.0035)	0.0243*** (0.0047)
Age	-0.0027*** (0.0003)	-0.0028*** (0.0004)	-0.0015*** (0.0001)	-0.0017*** (0.0002)
Female x Black	0 - 0.0413*** (0.0144)	-0.0262 (0.0193)	-0.0039 (0.0061)	0.0005 (0.0084)
Less Than HS x Black	0.0241 (0.0205)	0.0131 (0.0282)	-0.0186* (0.0109)	-0.0193 (0.0136)
Some College x Black	0.0122 (0.0114)	0.0077 (0.0144)	0.0088 (0.0058)	0.0070 (0.0071)
Bachelor's Degree x Black	0.0451*** (0.0149)	0.0477*** (0.0196)	0.0054 (0.0088)	0.0090 (0.0096)
Less Than HS x White Hispanic	-0.0385** (0.0191)	-0.0666*** (0.0248)	-0.0397*** (0.0081)	-0.0514*** (0.0119)
Some College x White Hispanic	0.0351*** (0.0125)	0.0428*** (0.0156)	0.0109** (0.0050)	0.0071 (0.0091)
Bachelor's Degree x White Hispanic	0.0573*** (0.0129)	0.0538*** (0.0175)	0.0218*** (0.0069)	0.0183** (0.0085)
Observations	283,000	161,000	254,000	145,000
Adjusted R^2	0.008	0.010	0.012	0.015
State Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Race/Gender Interactions	Y	Y	Y	Y
Race/Education Interactions	Y	Y	Y	Y
Gender/Education Interactions	Y	Y	Y	Y
Percentile Range	0-100	0-100	5-95	5-95

Source: CPS ASEC, IRS W-2and SSA DER, 2001-2016

This table presents the results of regressions of the CPS versus administrative records wage gaps on key demographic indicators at the individual level. These gaps are defined as the difference between total log wages reported to the CPS less total log wages reported to the respective administrative record, e.g. CPSw2wagegap measures the gap between CPS reported wages and W-2 reported wages. The coefficients of only a few selected interaction terms are shown. The baseline tax unit is male, White, unmarried, and has a high school diploma but no college. Standard errors are clustered at the sampling unit level. For the reported coefficients, those marked with three stars are significant at the one percent level, those marked with two stars are significant at the five percent level, and those marked with one star are significant at the ten percent level. Columns 3 and 4 truncate the sample at the 5th and 95th percentiles of the administrative records earnings distribution. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

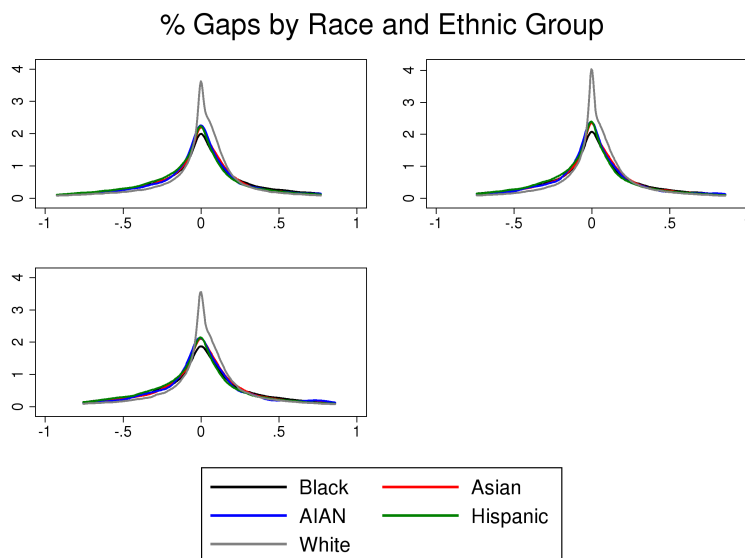
Here we observe less robust evidence of misreporting across race and ethnic groups relative to the extensive margin. In the unrestricted sample, there is no statistically significant difference in relative misreporting between, e.g. Blacks and Whites, although Hispanics under-report wages about five percentage points more than whites. There is more robust evidence of meaningful heterogeneity in reporting across racial groups when focusing on the middle of the distribution, as in columns 3 and 4, which exclude the top five and bottom five percent of the wage distribution. Blacks have 1.2 percentage point higher under-reporting rates in the trimmed DER-CPS sample, and 1.4 percentage points in the trimmed W2-CPS sample.

On the other hand, we observe much more robust evidence for an educational gradient in misreporting, particularly for Hispanics. Individuals with more education have lower under-reporting than individuals with less education, and these relative effects are monotonic over levels of education. Although not statistically significant, individuals with less than a High school degree under-report wages by about two percentage points more, relative to high school graduates, while college graduates under-report wages by about the same amount less. This gradient is even more stark when looking at Hispanics: a similar education gradient exists, but at much greater magnitudes. College educated Hispanics are under-report wages by as much as 6.7 percentage points less than Hispanic high school graduates, and Hispanics with less than a High School degree under-report by as much as 3.9 percentage points more than Hispanic high school graduates.

To further explore where this heterogeneity may be coming from, we examine the distribution of misreporting using several visualizations. First, we examine kernel density plots, which visualize the distribution of misreporting. Figure 1

shows this visualization, broken down by the race and ethnicity categories used in the regression analysis above, for the three sets of wage gaps (CPS wages minus DER, W-2 and 1040 wages respectively). Three things are immediately evident from these graphs. First, the density leans to the right, suggesting a tendency towards over-reporting. Second, whites have more mass near zero than non-whites, which is consistent with the regressions above. Third, the tails of the distribution are relatively “fat”: there instances of both large under- and over-reporting. A similar set of conclusions can be drawn from Figure 2, which shows kernel density plots by education level: there is a tendency towards over-reporting by all education groups, although college graduates have more mass near zero.

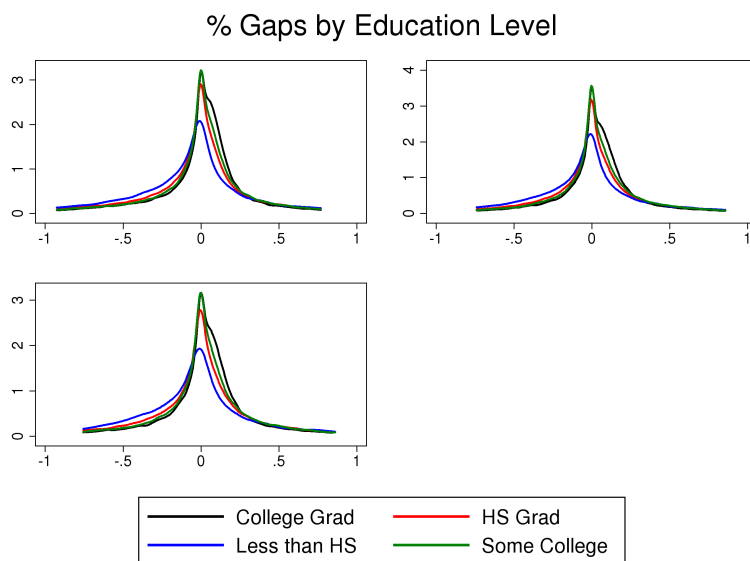
FIGURE 1.
Kernel Density Plots of Wage Reporting Differentials by Race and Ethnicity



Source: CPS ASEC, IRS W-2, SSA DER and IRS 1040s, 2001-2016

This figure shows kernel density plots of the fifth through ninety-fifth percentiles of wage reporting differentials for the mutually exclusive racial and ethnic groups, conditional on tax units reporting wages to both the IRS and CPS. Wage reporting differentials G_{ist} are shown along the horizontal axis, and density is along the vertical axis. The distribution of wage reporting differentials of tax units with a non-Hispanic white primary earner are plotted in grey, percentage gaps of tax units with a Hispanic white primary earner are plotted in green, percentage gaps of tax units with a black primary earner are plotted in black, percentage gaps of tax units with an Asian primary earner are plotted in red, and the percentage gaps of tax units with a Native American primary earner are plotted in blue. We use an Gaussian kernel. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

FIGURE 2.
Kernel Density Plots of Wage Reporting Differentials by
Education Level



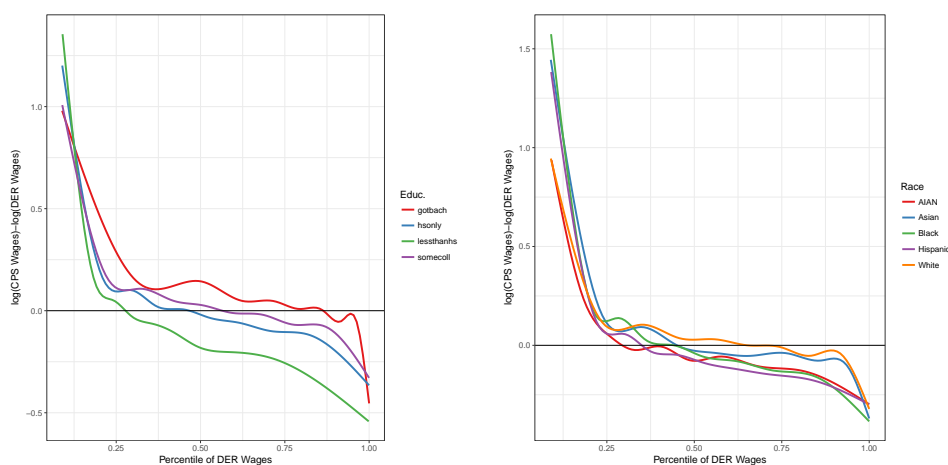
Source: CPS ASEC, IRS W-2, SSA DER and IRS 1040s, 2001-2016

This figure shows kernel density plots of the fifth through ninety-fifth percentiles of wage reporting differentials for the mutually exclusive education groups, conditional on tax units reporting wages to both the IRS and CPS. Wage reporting differentials G_{ist} are shown along the horizontal axis, and density is along the vertical axis. The distribution of wage reporting differentials of tax units with a primary earner who did not complete high school are plotted in blue, percentage gaps of tax units with a primary earner who only completed high school are plotted in red, percentage gaps of tax units with a primary earner who completed high school but did not complete a bachelor's degree are plotted in green, and the percentage gaps of tax units with a primary earner who completed a bachelor's degree are plotted in black. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

These kernel density plots are informative, but necessarily obscure heterogeneity across the wage distribution. A final visualization sheds light on this, and potentially rationalizes our previous set of results. Figure 3 shows estimates of the average wage gap by percentile of the administrative records wage distribution. The left panel breaks the visualization down by education level, while the right panel breaks it down by race. Here we can see that the kernel density plots were obscuring an important fact: over-reporting occurs primarily at the bottom of the wage distribution, while under-reporting primarily occurs at the top. This pattern occurs for all race and education groups, but there is heterogeneity in the degree of under-reporting at the top of the distribution and over-reporting at the bottom which is consistent with our regression results above: at the bottom, individuals with less than a high school degree have much larger over-reporting than individuals with a bachelor's degree, while at the top they have much larger under-reporting than college graduates. This suggests that the education gradient in misreporting should properly be thought of as a gradient in absolute misreporting.

It remains a very interesting question why individuals overreport at the bottom of the income distribution and underreport at the top. One competing explanation for the overreporting at the bottom of the income distribution is that there is shadow economic activity being reported to the CPS. This would generate patterns consistent with our findings. To consider this as an explanation, we must relax the assumption that the administrative records are the more correct source of wage data. In order to examine the validity of this competing explanation, we look into the industries and occupations of individuals that exhibit this behavior and see whether they are in occupations where shadow economic activity is likely.

FIGURE 3.
Wage Reporting Differentials Across the Wage distribution, by
Education Level and Race



Source: CPS ASEC, SSA DER, 2001-2016

The graphs shown above are generated by fitting a generalized additive model (GAM) to the bivariate wage percentile (horizontal axis) and wage misreporting (vertical axis) data. GAMs use smoothing splines to fit a smooth non-linear function to data. The left graph shows the relationships between the CPS/DER wage gaps and percentiles of DER wages, by education level, while the graph on the right shows the same relationships by racial and ethnic groups. These models are generated by utilizing the entire educational or racial/ethnic subset of the dataset to fit each model. Graphs of GAMs using the same subsets but instead using the CPS/W2 wage gaps by the same DER wage percentiles (not shown) look quite similar. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

Tables 14 and 15 repeat the individual level regressions, with additional dummy variables for industry and occupation, respectively. These indicators come from the CPS. In cases where the number of industry or occupation indicators changed across CPS years, a crosswalk is implemented to ensure consistency in coding. In addition to the independent variables shown in Tables 14 and 15, these regressions included the major demographic controls for age, gender, race, and education level (though no demographic interaction terms were included). It is paramount to note that the dependent variables measure the *relative* degree of misreporting, which is relative to a baseline respondent with both an industry and occupation in the “not in universe” category.

In Table 14, note that the signs of the coefficients on most of the industry indicators change when comparing the unrestricted sample versus the sample

TABLE 14.
Individual Level Regressions, by Industry

Dependent Variable	(1) CPSderwagegap	(2) CPSw2wagegap	(3) CPSderwagegap	(4) CPSw2wagegap
Agriculture	0.0374 (0.0277)	0.0170 (0.0359)	0.00834 (0.0161)	-0.0230 (0.0200)
Mining	-0.0573 (0.0397)	-0.0709 (0.0465)	-0.0497*** (0.0175)	-0.0697*** (0.0229)
Construction	0.0410* (0.0230)	0.0453 (0.0279)	-0.0216 * * (0.00881)	-0.0328*** (0.0107)
Manufacturing	0.0124 (0.0191)	0.00981 (0.0221)	-0.0191** (0.00744)	-0.0249*** (0.00892)
Trade	0.0558*** (0.0183)	0.0616*** (0.0230)	-0.00234 (0.00732)	-0.00924 (0.00931)
Transportation & Utilities	0.0196 (0.0223)	0.0254 (0.0286)	-0.0165* (0.00872)	-0.0229** (0.0105)
Information	0.00998 (0.0237)	-0.00140 (0.0328)	-0.00681 (0.0103)	-0.0212 (0.0134)
Finance, Insurance, & Real Estate	0.0239 (0.0217)	0.0253 (0.0262)	-0.00558 (0.00918)	-0.0110 (0.0115)
Professional	0.0594*** (0.0184)	0.0710*** (0.0245)	-0.00406 (0.00741)	-0.00430 (0.00975)
Education and Health Care	0.0404** (0.0197)	0.0392* (0.0237)	-0.00277 (0.00744)	-0.0129 (0.00924)
Leisure	0.0783*** (0.0215)	0.0940*** (0.0282)	0.0133* (0.00775)	0.0107 (0.0101)
Other Services	0.0771*** (0.0233)	0.0663** (0.0310)	0.00994 (0.00990)	0.000883 (0.0124)
Public Administration	0.0464** (0.0218)	0.0277 (0.0275)	0.00560 (0.00892)	-0.00547 (0.0119)
Armed Forces	0.239 (0.147)	0.357 (0.300)	0.0593 (0.0914)	0.0309 (0.0700)
Observations	283,000	161,000	254,000	145,000
Adjusted R^2	0.009	0.010	0.012	0.014
State Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Demographic Controls	Y	Y	Y	Y
Percentile Range	0-100	0-100	5-95	5-95

Source: CPS ASEC, IRS W-2and SSA DER, 2001-2016

This table presents results of regressions of CPS vs. administrative records wage gaps on CPS industry indicators at the individual level. These gaps are the difference between total log wages reported to the CPS less total log wages reported to the administrative record, e.g. CPSw2wagegap measures the gap between CPS reported wages and W-2 reported wages. Standard errors are clustered at the sampling unit level. Reported coefficients marked with three stars are significant at the one percent level, those marked with two stars are significant at the five percent level, and those marked with one star are significant at the ten percent level. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

TABLE 15.
Individual Level Regressions, by Occupation

Dependent Variable	(1) CPSderwagegap	(2) CPSw2wagegap	(3) CPSderwagegap	(4) CPSw2wagegap
Management	0.0771*** (0.0206)	0.0803*** (0.0262)	0.0133* (0.00787)	0.00820 (0.0103)
Professional	0.0446** (0.0196)	0.0403* (0.0240)	-0.00230 (0.00828)	-0.0112 (0.0105)
Service	0.0392** (0.0196)	0.0488** (0.0241)	-0.00457 (0.00720)	-0.00948 (0.00897)
Sales	0.0602*** (0.0186)	0.0724*** (0.0241)	0.00378 (0.00813)	-0.00211 (0.0102)
Office	0.0505*** (0.0195)	0.0501** (0.0237)	0.00302 (0.00749)	-0.00299 (0.00930)
Farming	-0.00861 (0.0292)	-0.0256 (0.0387)	-0.0145 (0.0160)	-0.0344* (0.0195)
Construction	0.0407* (0.0239)	0.0414 (0.0296)	-0.0224** (0.00886)	-0.0341*** (0.0110)
Installation	0.0375* (0.0217)	0.0349 (0.0274)	-0.0135 (0.0100)	-0.0338*** (0.0126)
Production	0.0176 (0.0196)	0.0125 (0.0231)	-0.0244*** (0.00793)	-0.0345*** (0.00944)
Transportation	0.0287 (0.0213)	0.0242 (0.0264)	-0.0125 (0.00839)	-0.0175 (0.0113)
Armed Forces	0.243* (0.146)	0.360 (0.299)	0.0613 (0.0911)	0.0331 (0.0700)
Observations	283,000	161,000	254,000	145,000
Adjusted R^2	0.008	0.010	0.012	0.014
State Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Race, Gender, Education, & Age Controls	Y	Y	Y	Y
Percentile Range	0-100	0-100	5-95	5-95

Source: CPS ASEC, IRS W-2and SSA DER, 2001-2016

This table presents the results of regressions of the CPS versus administrative records wage gaps on CPS occupation indicators at the individual level. These gaps are defined as the difference between total log wages reported to the CPS less total log wages reported to the respective administrative record, e.g. CPSw2wagegap measures the gap between CPS reported wages and W-2 reported wages. Standard errors are clustered at the sampling unit level. For the reported coefficients, those marked with three stars are significant at the one percent level, those marked with two stars are significant at the five percent level, and those marked with one star are significant at the ten percent level. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

trimmed at the fifth and ninety-fifth percentiles. This is consistent with a large amount of overreporting on the CPS in the upper end of the sample for most professions. This difference is most notable for the trade, professional, and leisure industries. Overall, respondents in the armed forces relatively overreport the most, but this is likely due to the fact that combat pay is non-taxable and would be excluded from taxable wages on Forms W-2. Future iterations of this work will drop respondents in the armed forces for this reason. In the unrestricted sample, respondents in the trade, professional, leisure, and other services industries relatively overreport the most, while those in the mining industry relatively underreport the most by a wide margin. In the trimmed sample, those in the leisure industry continue to relatively overreport and those in the mining industry continue to relatively underreport, however the coefficients on most other indicators attenuate towards zero. Interestingly, the coefficients for the construction, manufacturing, and transportation industries are significant and substantially negative in the trimmed sample. These industries tend to be more heavily unionized than most, though the education and health care industry, which is more unionized than the mining industry, does not share this degree of relative underreporting.

Table 15, showing regression results for occupation indicators, should also be interpreted with the same caveats as Table 14. Again, we see the same large amount of relative overreporting by respondents working in the armed forces, again, likely due to the non-taxability of combat pay. In addition, respondents with management and sales positions tend to relatively overreport to the CPS in the untrimmed sample. However, again, most of these larger coefficients disappear when we move to the trimmed sample, indicating that these large effects in

the upper tails are not indicative of the occupation as a whole. However, when looking at the trimmed sample, we see that respondents with jobs in construction and production show the greatest degree of relative underreporting to the CPS, consistent with the findings for the construction and manufacturing industries in Table 14.

Going forward, it would be very useful as we try to understand the role of the shadow economy in what appears to be “overreporting to the CPS,” to be able to see industry and occupation reporting across the distribution as we did earlier for racial and educational groups. This would allow us to isolate their contribution to the overreporting phenomenon we are trying to explain.

Sources of Misreporting: Rounding

We have shown the demographic traits associated with wage misreporting; now we search for the mechanisms by which individuals may misreport. In particular, we look at rounding. If individuals in different demographic groups round in systematic ways, then those rounding heuristics could translate into a large amount of misreporting. For example, if members of a demographic group tend to use a rounding heuristic where they drop all digits after the thousands place, then that group will tend to underreport wages to the CPS. If another group uses a different rounding heuristic, such as tending to round up to the nearest ten thousand dollars, the effects of rounding will go in the opposite direction. It is also the case that systematic rounding choices will have a larger percent gap effect at the bottom end of the distribution where rounding to the nearest 1,000 say, is a much larger change than at the top of the distribution. Thus it is important to determine how much wage misreporting is based on rounding as opposed to

deviations where the CPS respondent intentionally aims to inflate or deflate their wages to the CPS for other reasons.

We begin by constructing dummy variables for different degrees of wage rounding on the CPS. These are defined as

$$I_{ist,X} = \begin{cases} 1, & \text{if } y_{ist,CPS} \bmod (X * 1,000) = 0 \\ 0, & \text{otherwise,} \end{cases} \quad (3.4)$$

for $X \in \{1, 5, 10\}$ and where $y_{ist,CPS}$ is wage income as reported to the CPS. In other words, these dummies indicate if CPS reported wages are reported to the nearest one, five, or ten thousand dollars. Table 16 displays the results of simple regressions of the gap variables and absolute values of gap variables on these three rounding dummy variables. It is clear that rounding accounts for a hefty portion of misreporting. However, these simple regression coefficients are difficult to interpret for two main reasons: first, the fact that numbers are rounded on the CPS does not necessarily mean that they are incorrect, just that they are round (so, in cases where the W-2 or DER number is also round, these coefficients are possibly attenuated); and second, the log difference caused by, say, rounding to the nearest ten thousand dollars is likely much larger at lower income levels than higher income levels, due to the simple nature of the natural logarithm function. Therefore, these coefficient estimates may give some insight into the overall scope of the rounding problem, without pinpoint accuracy.

To further examine the effects on specific regression coefficients, we take two approaches. First, we re-run the individual level regressions using only CPS respondents who do *not* report wages that are rounded to the nearest thousand, and compare the results with the full-sample regressions. Second, we mimic the

TABLE 16.
Effects of Rounding

Dependent Variable	(1) CPSderwagegap	(2) CPSw2wagegap	(3) Abs(CPSderwagegap)	(4) Abs(CPSw2wagegap)
Rounded 1,000	0.111*** (0.00399)	0.104*** (0.00533)	-0.109*** (0.00420)	-0.112*** (0.00590)
Rounded 5,000	0.0401*** (0.00280)	0.0400*** (0.00392)	0.0210*** (0.00235)	0.0195*** (0.00310)
Rounded 10,000	0.0109*** (0.00318)	0.00817** (0.00392)	0.0142*** (0.00275)	0.0131*** (0.00341)
Observations	336,000	191,000	336,000	191,000
Adjusted R^2	0.009	0.008	0.006	0.006
State Fixed Effects	N	N	N	N
Year Fixed Effects	N	N	N	N
Race, Gender, Education, & Age Controls	N	N	N	N
Percentile Range	0-100	0-100	0-100	0-100

Source: CPS ASEC, IRS W-2and SSA DER, 2001-2016

This table presents the results of regressions of the CPS versus administrative records wage gaps on indicator variables for the degree of rounding of wages reported to the CPS. These indicators are set equal to one if the respondent's wages are rounded to the nearest one, five, or ten thousand. The gaps are defined as the difference between total log wages reported to the CPS less total log wages reported to the respective administrative record, e.g. CPSw2wagegap measures the gap between CPS reported wages and W-2 reported wages. "Abs" refers to the absolute value, e.g. Abs(CPSw2wagegap) measures the absolute value of the gap between CPS reported wages and W-2 reported wages. The coefficients of only a few selected interaction terms are shown. The baseline tax unit is male, White, unmarried, and has a high school diploma but no college. Standard errors are clustered at the sampling unit level. For the reported coefficients, those marked with three stars are significant at the one percent level, those marked with two stars are significant at the five percent level, and those marked with one star are significant at the ten percent level. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

rounding of CPS respondents by rounding administrative wages as well, and re-run the individual level regressions using the rounding adjusted wage gaps.

The first approach is shown in Table 17, which shows regression results for most demographic and a few selected industry variables. Note that only about eighteen percent of respondents in the sample report wages that are not rounded to a nearest thousand dollars of some sort.

Despite the loss in power, most coefficient estimates for the demographic variables are highly significant. Focusing on columns (3) and (4) of Table 17, we can see that most of the coefficients share the same signs as their corresponding coefficients on Table 13 (for demographic variables) and Table 14 (for industry variables). Coefficient estimates are similar to the full sample for racial and ethnic indicators. For educational indicators, estimates increase for *LessThanHighSchool* and *Bachelor'sDegree*, when industry controls are included. This suggests that rounding is causing some of the relative underreporting of less educated individuals, while mitigating some of the relative overreporting of more educated individuals. For age, the most stable coefficient thus far, estimates for non-rounders only are only about half the magnitude of previous estimates. This is consistent with the idea that individuals budget more conservatively by mentally rounding down as they age, as excluding rounders attenuates these estimates. Looking at industries, coefficient estimates here for the mining and construction industries are less significant and closer to zero, suggesting that rounding is driving previously reported estimates for these industries. Coefficient estimates here for agriculture and public administration are larger in magnitude and more significant than before, suggesting that rounding was previously creating a lot of noise in these estimates. However, these suggestive findings are not certain, as individuals who

TABLE 17.
Individual Level Regressions, Respondents without Rounded Wages

Dependent Variable	(1) CPSderwagegap	(2) CPSw2wagegap	(3) CPSderwagegap	(4) CPSw2wagegap
Black	-0.0501*** (0.0122)	-0.0379** (0.0165)	-0.0161*** (0.00438)	-0.0128** (0.00584)
Asian	-0.0707*** (0.0198)	-0.0701*** (0.0241)	-0.0278*** (0.00618)	-0.0309*** (0.0100)
Native American	-0.0722** (0.0291)	-0.0587* (0.0332)	-0.0147 (0.0109)	-0.0152 (0.0147)
White Hispanic	-0.0670*** (0.0118)	-0.0609*** (0.0171)	-0.0357*** (0.00449)	-0.0313*** (0.00649)
Less Than High School	-0.0472*** (0.0124)	-0.0573*** (0.0157)	-0.0158*** (0.00469)	-0.0225*** (0.00681)
Some College	0.0250*** (0.00810)	0.0313*** (0.00963)	0.0164*** (0.00303)	0.0178*** (0.00398)
Bachelor's Degree	0.0576*** (0.00935)	0.0608*** (0.0126)	0.0325*** (0.00309)	0.0345*** (0.00409)
Age	-0.00167*** (0.000320)	-0.00157*** (0.000465)	-0.000786*** (0.000122)	-0.000641*** (0.000167)
Agriculture	0.0894** (0.0352)	0.0971** (0.0467)	-0.0376** (0.0157)	-0.0442** (0.0206)
Construction	0.0611*** (0.0184)	0.0589** (0.0244)	0.000462 (0.00650)	-0.00191 (0.00919)
Mining	0.0354 (0.0494)	0.0411 (0.0613)	-0.0394** (0.0160)	-0.0347 (0.0240)
Public Administration	0.0962*** (0.0155)	0.0832*** (0.0211)	0.0155*** (0.00566)	0.0203*** (0.00774)
Observations	61,000	35,000	52,500	30,000
Adjusted R^2	0.008	0.009	0.011	0.011
State Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Race, Gender, Education, & Age Controls	Y	Y	Y	Y
Industry Controls	Y	Y	Y	Y
Percentile Range	0-100	0-100	5-95	5-95

Source: CPS ASEC, IRS W-2 and SSA DER, 2001-2016

This table presents the results of regressions of the CPS versus administrative records wage gaps at the individual level for only those CPS respondents whose wages were not rounded to the nearest thousand dollars. These gaps are defined as the difference between total log wages reported to the CPS less total log wages reported to the respective administrative record, e.g. CPSw2wagegap measures the gap between CPS reported wages and W-2 reported wages. Selected coefficients are shown. Standard errors are clustered at the sampling unit level. For the reported coefficients, those marked with three stars are significant at the one percent level, those marked with two stars are significant at the five percent level, and those marked with one star are significant at the ten percent level. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

round may tend to share unobserved characteristics and systematically differ from non-rounders, or *vice versa*.

For the next approach, we simulate the rounding approach taken by CPS respondents. We are able to observe whether they reported wages to the nearest one, five, or ten thousand dollars, or whether they did not round to a nearest thousand. If they reported wages on the CPS that are divisible by ten thousand, we assume that they meant to round to the nearest ten thousand and round the corresponding administrative record wages to those wages' nearest ten thousand. If they reported CPS wages that are divisible by five thousand but not ten thousand, we assume that they meant to round to the nearest five thousand and round their administrative record wages to those wages' nearest five thousand. If they reported CPS wages that are divisible by one thousand but not by five or ten thousand dollars, we assume that they meant to round to the nearest one thousand dollars and round their administrative record wages to those wages' nearest one thousand dollars. The remaining gap between reported CPS wages and administrative record wages, adjusted for rounding to these thousand dollar increments, cannot be the result of rounding, unless respondents rounded to a lower or higher amount (such as the nearest hundred dollars or million dollars), or unless they used a non-standard rounding algorithm, such as always rounding up or down. Regression results using these modified wage gaps are shown on Table 18. Columns (1) and (3) show results when we mimic the rounding, and columns (2) and (4) show results when we do not, for comparison. In general, coefficient estimates look similar to before, but with attenuated magnitudes. The same patterns of underreporting by historically disadvantaged racial and ethnic groups and the gradient with respect to education level hold as before. The decrease in the magnitude of the *Age* coefficient

is consistent with the idea that much of the underreporting as respondents get older is due to rounding. Selected industries are shown. The industry coefficients that changed substantially the most are in the mining and construction industries, both by greatly decreasing the magnitudes of coefficient estimates. In particular, a great deal (about a couple of percentage points) of the wage gaps for respondents in the mining industry is likely caused by rounding.

We may also be interested in how the s-shaped curves plotted in Figure 3 change when we address rounding, particularly because rounding will be a bigger deal as a percentage at the bottom of the income distribution. The results are not yet publicly available, but we can report that while the s-shaped curves flatten some, the overall s-shaped patterns remain.

Overall, we find in this section that while rounding can play some role in estimates of particular variables, we see no clear overarching effect of rounding on misreporting across all variables for all our different rounding sensitivity analysis.

Implications for Estimates of the Income Distribution

Public use CPS data are often used to prepare estimates of the income distribution as well as statistics derived from the income distribution, such as the Gini coefficient. A summary of relevant research using CPS data to this end can be found in (Burkhauser et al., 2012). The s-shaped pattern of wage reporting—that lower wage amounts tend to be overreported to the CPS while higher wage amounts tend to be underreported—suggests that CPS-based measures of inequality will understate the degree of inequality, *ceteris paribus*, due to this compression.

In recent years, researchers such as Burkhauser et al. have noticed that CPS-based measures of inequality tend to show inequality slowing in the 1990s, while

TABLE 18.
Individual Level Regressions, Misreporting Not Likely Caused by Rounding

Dependent Variable Researchers Rounded CPS Responses	(1)	(2)	(3)	(4)
	CPSderwagegap Y	CPSderwagegap N	CPSderwagegap Y	CPSderwagegap N
Black	-0.00394 (0.00428)	-0.00116 (0.00816)	-0.0112 * ** (0.00207)	-0.0151*** (0.00382)
Asian	-0.00932 (0.00717)	-0.0183* (0.00963)	-0.0188*** (0.00305)	-0.0240*** (0.00533)
Native American	-0.0218** (0.0110)	-0.0382 (0.0236)	-0.0155*** (0.00509)	0.00209 (0.0129)
White Hispanic	-0.0324*** (0.00485)	-0.0413*** (0.00734)	-0.0258*** (0.00244)	-0.0305*** (0.00324)
Less Than High School	-0.0545*** (0.00486)	-0.0515*** (0.00783)	-0.0298*** (0.00243)	-0.0307*** (0.00434)
Some College	0.0159*** (0.00271)	0.0360*** (0.00765)	0.0131*** (0.00124)	0.0189*** (0.00351)
Bachelor's Degree	0.0191*** (0.00312)	0.0401*** (0.00777)	0.0186*** (0.00158)	0.0241*** (0.00511)
Age	-0.00262*** (0.000133)	-0.00261*** (0.000293)	-0.00107*** (0.0000552)	-0.00145*** (0.000121)
Mining	-0.0968*** (0.0124)	-0.573*** (0.0397)	-0.0404*** (0.00645)	-0.0497*** (0.0175)
Construction	-0.0144** (0.00666)	0.0410* (0.0230)	-0.0156*** (0.00298)	-0.0216*** (0.00881)
Constant	-0.0655 (0.415)	-0.116 (0.407)	0.0142*** (341.4)	0.350*** (0.0124)
Observations	336,000	283,000	301,000	254,000
Adjusted R^2	0.007	0.008	0.009	0.010
State Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Race, Gender, Education, & Age Controls	Y	Y	Y	Y
Industry Controls	Y	Y	Y	Y
Percentile Range	0-100	0-100	5-95	5-95

Source: CPS ASEC, IRS W-2 and SSA DER, 2001-2016

This table presents results of regressions of the CPS vs. DER wage gaps at the individual level after adjusting administrative records by rounding. If CPS reported wages are reported to the nearest ten thousand dollars, we round DER wages to the nearest ten thousand dollars before regressing, and so on. The wage gaps are defined as the difference between total log wages reported to the CPS less total log wages reported on SSA DER records. Selected coefficients are shown. Columns (2) and (4) repeat coefficient estimates shown on Table 14, do not incorporate researcher-created rounding, and are shown for comparison. Standard errors are clustered at the sampling unit level. For the reported coefficients, those marked with three stars are significant at the one percent level, those marked with two stars are significant at the five percent level, and those marked with one star are significant at the ten percent level. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

IRS-based measures tend to find gains to the richest Americans rising rapidly throughout this time period (Piketty and Saez, 2008). Taking into account the compression of reported wages in the CPS may not fully reconcile the differing conclusions made from the CPS and IRS data - indeed, certain other differences (such as the extensive top-coping of income data for richer Americans on the CPS) make this task challenging. However, the persistent s-shaped pattern, found across all racial and ethnic groups and education levels, even when accounting for rounding, suggests that changes over time in the degree of compression may account for some of the divergence in conclusions made from the CPS versus IRS data. Whether or not such an effect is found remains to be the subject of future research.

Tax Unit Level Analysis

Finally, for completeness, we repeat the regression analysis of the last section at the tax unit level, comparing form 1040 wages to CPS wages. Forms 1040 are filed at by tax units, including married units, so demographic indicators are collapsed to the tax unit level using the average characteristics of the unit. Thus a unit with one spouse with some college but another spouse with a different level of education would be coded as a 0.5 for the “Some College” variable. Interaction terms are excluded because half values confound the interpretation of such terms. Regression results are presented in Table 19.

In general, coefficients take similar signs, magnitudes, and significances as in the individual level analyses. Age is robust, the same pattern emerges amongst the educational dummies, and the coefficients on the racial and ethnic indicators are similar to earlier estimates, though attenuated in the case of White Hispanic

TABLE 19.
Tax Unit Regressions

Dependent Variable	(1) CPS1040wagegap	(2) CPS1040wagegap	(3) CPS1040wagegap
Married	-0.0316*** (0.0036)	-0.0237*** (0.0028)	-0.0067*** (0.0018)
Black	-0.0018 (0.0054)	-0.0044 (0.0044)	-0.0143*** (0.0029)
Asian	-0.0190** (0.0076)	-0.0220*** (0.0055)	-0.0254*** (0.0042)
Native American	0.0080 (0.0175)	-0.0018 (0.0122)	-0.0107 (0.0080)
White Hispanic	-0.0406*** (0.0070)	-0.0361*** (0.0050)	-0.0265*** (0.0032)
Less Than High School	-0.0838*** (0.0076)	-0.0611*** (0.0060)	-0.0356*** (0.0044)
Some College	0.0157*** (0.0043)	0.0217*** (0.0035)	0.0166*** (0.0025)
Bachelor's Degree	0.0301*** (0.0044)	0.0281*** (0.0035)	0.0251*** (0.0027)
Age	-0.0023*** (0.0002)	-0.0022*** (0.0001)	-0.0013*** (0.0001)
Observations	180,000	176,000	162,000
Adjusted R^2	0.009	0.010	0.009
State Fixed Effects	Y	Y	Y
Year Fixed Effects	Y	Y	Y
Percentile Range	0-100	1-99	5-95

Source: CPS ASEC, IRS 1040, 2001-2016

This table presents the results of regressions of the CPS versus 1040 wage gap on key demographic indicators at the tax unit level. This gap is defined as the difference between total log wages reported to the CPS less total log wages reported on form 1040. Where tax units are married, these demographic indicators are based on the average demographics of the couple. The baseline tax unit is White, unmarried, and has a high school diploma but no college. Standard errors are clustered at the sampling unit level. For the reported coefficients, those marked with three stars are significant at the one percent level, those marked with two stars are significant at the five percent level, and those marked with one star are significant at the ten percent level. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

and more significant in the case of Asian. However, a new pattern emerges with the Married indicator, where filing as a married unit is associated with a two-thirds to three percent underreporting to the CPS relative to the IRS. This is in stark contrast to the insignificant, near-zero coefficient estimates in the individual level analyses.

Discussion

We now discuss the implications of using income data with systematic measurement errors for statistical compilation or regression modeling. Statistics compiled from CPS wage data can be adjusted, and regressions can be modified to account for non-zero mean errors. The end of this discussion proposes mechanisms by which measurement errors may be systematically introduced into survey income data.

Suppose we are interested in cases where wages are reported with systematic error in a way that is correlated with a demographic indicator variable. In the case of simple ratios that use CPS wage data, where wages are systematically scaled up or down by a factor, the corrections are simple. For example, if we were measuring a natural logarithmic Black-White wage differential of average incomes,

$$\Delta_{BW} = \log(\bar{y}_{White}^*) - \log(\bar{y}_{Black}^*) \tag{3.5}$$

using measured average incomes

$$\bar{y}_{Black} = \bar{y}_{Black}^*(1 - \delta_{Black}), \tag{3.6}$$

and we had estimated the coefficient on the Black indicator, $\hat{\delta}_{Black}$ in our model.

The proper adjustment would be

$$\begin{aligned}\Delta_{BW} &= \log(\bar{y}_{White}^*) - \log\left(\frac{\bar{y}_{Black}}{1 - \delta_{Black}}\right) \\ &= \log(\bar{y}_{White}^*) - \log(\bar{y}_{Black}) + \log(1 - \delta_{Black}).\end{aligned}\quad (3.7)$$

Assuming the wage differential was positive and the error not bigger than the measured differential, this has the effect of tightening the calculated racial wage differential. This case is analogous to the case where both White and Black incomes are measured each with non-zero mean errors, because we are only investigating the differential.

Implications for regressions can be complicated. We present the simple case in which the dependent variable is survey income, and the incomes for one demographic group are measured with error. Suppose we consider a true model,

$$\tilde{Y}^* = X^*\beta + U, \quad (3.8)$$

where Y^* are incomes, \tilde{Y}^* are the natural logarithms of those incomes, and X^* are explanatory variables. Suppose incomes for one demographic group, say, Blacks, are misreported to surveys with both systematic and random errors, as in

$$y_i = y_i^*(1 - \delta_{Black}Black_i)v_i. \quad (3.9)$$

Here, δ_{Black} is the income reporting differential estimated from our regressions as $\hat{\delta}_{Black}$, and v_i is a median one, log-normally distributed error, where $\log(v_i) = \tilde{v}_i$. After taking logarithms, the $\log(v_i)$ will become an additive error term on the right hand side and decrease model efficiency. We normalize so that the Black indicator

is the K^{th} of K independent variables. By using the systematically mismeasured incomes in our dependent variable, an OLS regression will really run the regression

$$\begin{aligned}
& \begin{bmatrix} \log(y_1^*(1 - \delta_{Black}Black_1)) \\ \vdots \\ \log(y_n^*(1 - \delta_{Black}Black_n)) \end{bmatrix} \\
= & \begin{bmatrix} x_{11} & \dots & x_{1K-1} & \Big| & Black_1 \\ \vdots & \ddots & \vdots & & \vdots \\ x_{n1} & \dots & x_{nK-1} & \Big| & Black_n \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{K-1} \\ \beta_{Black} \end{bmatrix} + \begin{bmatrix} u_1 - \tilde{v}_1 \\ \vdots \\ u_n - \tilde{v}_n \end{bmatrix} \\
& = \begin{bmatrix} \tilde{y}_1^* \\ \vdots \\ \tilde{y}_n^* \end{bmatrix} + \begin{bmatrix} \log(1 - \delta_{Black}Black_1) \\ \vdots \\ \log(1 - \delta_{Black}Black_n) \end{bmatrix} \approx \begin{bmatrix} \tilde{y}_1^* \\ \vdots \\ \tilde{y}_n^* \end{bmatrix} + \begin{bmatrix} -\delta_{Black}Black_1 \\ \vdots \\ -\delta_{Black}Black_n \end{bmatrix},
\end{aligned}$$

and the pathology is clear: using measured underreported incomes is akin to inflating the magnitude of the independent variable *Black* itself (rather than it taking values $\{0, 1\}$, having it take values approximately $\{0, 1 + \delta_{Black}\}$). The prescriptions, imputing incomes for the dependent variable or scaling down the related indicator variable on the right hand side by approximately $\frac{1}{1+\delta_{Black}}$, are equivalent. Computationally, the latter fix requires fewer steps. Incomes can be imputed, conditional on $Black = 1$. The latter fix requires no knowledge of the

status of $Black_i$, because all positive and zero values of $Black$ can be scaled down by the exact same factor.

An even simpler fix can be applied *post mortem* in this simple case. Because the estimated coefficient on $Black$, $\hat{\beta}_{Black}$, was estimated in the faulty regression, we have

$$\hat{\beta}_{Black}Black = \beta_{Black}^*(1 + \delta_{Black})Black \quad (3.10)$$

and the adjusted coefficient can be obtained by dividing the estimate through by $(1 + \delta_{Black})$. As in the simple case of comparing median wages, the systematic reporting differential $\delta_{Black} > 0$ causes the magnitude of the Black-White wage gap to be overstated when the dependent variable is recorded with systematic error. If this type of error in the dependent variable occurs due to multiple reporting differentials for multiple demographic groups, the estimated coefficients on the indicator variables corresponding to those errors can each be adjusted by a similar scaling down (or up) of the estimate.

Rather than having to impute incomes, modify statistics, and employ fixes while regression modeling, we would like to measure incomes more accurately. An analysis of the income reporting process suggests avenues where systematic misreporting may occur and potentially be mitigated.

To contrast the incentives for measurement error on the CPS, we note the incentives to properly report wages to the IRS. IRS wages are typically reported as follows: an employee performs work for an employer; when the employee is paid, income and employment taxes are withheld by the employer or the employer's agent; after year end, the employer or employer's agent send duplicate statements of wage earnings for the tax year to the employee and the IRS; the employee or an

agent of the employee prepare a Form 1040 for the employee, using the duplicate wage earnings statement which has instructions on how to report those wages on Form 1040 (and usually attaching the wage earnings statement to Form 1040).

Employment taxes incentivize employees against reporting non-wage income types as wages. Employees are incentivized to fully report wages earned for compliance purposes. Wage expenses are deductible from the taxable income of the employer, so the employer has incentive to report all wages paid to the employee. While simple recording errors occur, the tax unit is strongly incentivized to record no more and no less than actual wages earned on Form 1040. However, the employee may not be aware of the existence of duplicate records, may not know that incorrectly reported wages may easily trigger an audit, may suffer cognitive costs in properly completing Form 1040, or for other reasons make gross, intermittent errors in wage reporting such as failing to report wages earned.¹⁵

On the other hand, CPS wages are reported as follows: the survey respondent receives a letter preparing the household for the upcoming regular CPS interviews; the interviewer establishes personal contact for the initial interview and may perform subsequent interviews by phone; when it is time for the ASEC, the interviewer asks the respondent additional detailed income questions; the respondent answers the items; the interviewer records the items. In the case of IRS wage reporting, a wage earning tax unit has financial incentives not to overreport or underreport wages; here, the household has no such incentives. Furthermore, a

¹⁵One notable exception bears mentioning here: cash tips paid by customers to employees. Such tips are paid to the employee directly by the customer and thus never enter the employer's books as revenue. Thus, the usual check on employee underreporting, the employer's incentive to minimize its taxable income, fails. Additionally, employees are allowed to estimate tips earned. Here, duplicate statements sent by employers will only estimate the amount of tips earned, and thus employees may underreport tips, undetected. A 1993 IRS study found that more than half of all tips were unreported to the IRS. After implementation of programs, by 2004, nominal tips reported to the IRS had more than doubled (U.S. Treasury Department, 2007).

CPS interviewer has no duplicate wage earnings statements and cannot audit the respondent to validate wage amounts reported by the respondent. In considering pathways for systematic error to occur, we note that the ASEC is longer and more time intensive than the usual CPS interview.

Wage incomes can be recorded with error if they are not remembered properly (if the household either literally forgets, or loses or ignores wage statements). In years where wage statements were sent out later than others, it is possible that a household was interviewed for the ASEC before receiving such statements, and we might expect for memories to be worse and reported wages more variable. If actual prior year wages are forgotten, and household members hold the same jobs, they may heuristically recall current incomes and use them as a proxy for prior year incomes. In the case where real wages rise or are stagnant, and there is positive inflation, we should expect these errors to cause wages to be overstated.

Once wages are recalled, either correctly or incorrectly, households may report them accurately or inaccurately to the interviewer. Due to the length of the ASEC, households may cut corners to save time, for example, by rounding numbers or leaving out income items entirely. Households may be incentivized to misreport wages based on the implications of the survey. The introductory letter to the CPS tells households that their answers in the survey will be used to represent hundreds of households like them (U.S. Census Bureau, 2006). Households may strategically underreport or overreport incomes if they believe that their responses may lead to some future beneficial public policy outcome. Also, households may worry about exposing themselves to theft may be more likely to underreport incomes.

These potential sources of error are due to timing, time investment, and self-interest. When surveys are completed before respondent receive wage statements,

responses should be noisier. Many of the problems that arise in properly recording wages on the CPS occur because greater accuracy likely involves spending more time, but there is no penalty for being inaccurate. The CPS could experiment with providing respondents with directions for year-end recordkeeping that aim to reduce response time (such as preparing respondents to keep wage statements in a set place and leaving a message about the upcoming interview). To address self-interest, the CPS could experiment with different introductory letters that do and do not emphasize the fact that survey responses will be used to represent similar households.

Conclusion

In this paper, we match household IRS wage records to CPS wage records and calculate wage reporting differentials. We utilize the accuracy of IRS wage reporting to attribute any reporting differences to wage mismeasurement on the CPS. We find that households of racial and ethnic minorities and less educated households underreport wages to the CPS, relative to non-Hispanic Whites and more educated households. Wage underreporting on the CPS tends to also increase with age.

Based on these results, wage differentials for racial and ethnic groups that are derived from uncorrected CPS data are overstated. Simple calculations of the returns to education will be overstated as well. Implications in regressions are more complicated and can be analytically unpredictable. These results cast doubt on the assumption that CPS recorded wages can be used as a measure of IRS wages earned, unless adjustments are made. Our results suggest that CPS wage underreporting is not independent of demographic factors.

In this paper, we controlled for related demographic characteristics and included interaction terms between race and ethnicity, gender, and education in many regressions. However, many within-group heterogeneities, such as occupation, remain unexplored. Users of these results should be cautious in that these results only estimate degrees of relative reporting between groups and do not represent all members of these groups.

Timing of wage reporting may matter. When respondents have their duplicate wage statements not long before the ASEC is conducted, they should be more accurate. The due dates of the duplicate statements changed throughout the sample. We can see if the variance of wage reports increases when deadlines are later and possibly after the ASEC. We can tailor our approach more to the intricacies of the CPS, exploiting the panel qualities of the survey.

Finally, we note the potential danger in naively applying our results. Estimates of wage reporting differentials by demographic groups represent the average tendencies of those households relative to baseline tendencies. While we control for many factors, we do not provide regression-based heterogeneity analyses within groups. An average effect we detect could be driven by a specific subset of any one demographic group, which Figure 3 suggests may be the case. Further research is needed to determine whether the relative differentials we find are representative of groups as a whole, or merely particular subsets of demographic groups.

CHAPTER IV

GAMES OF FISCAL MANAGEMENT

Introduction and Related Literature

I write this from the city of Eugene, part of Lane County, in the state of Oregon, which in turn is part of the United States of America. These are all geographical areas as well as government jurisdictions. Just as these areas are nested within one another, their governments are nested as well: as a county lies within a state, a county is beholden to the laws and procedures of the state. Typically in this hierarchy, geographical areas that are subsets of other areas have governments that are subordinate to the bigger government.

Up and down this chain of nested governments, the different levels of government produce public goods, and typically collect tax revenue. Ideal design of public goods production would have any specification of public goods produced at as low cost as possible, by delegating the responsibility of implementing those public goods to the different levels of government in the most efficient manner, including considerations for the overhead costs of the different levels, economies of scale, and externalities created by one level of government that spill onto another. Likewise, ideal tax system design would have a given amount of revenue collected with the lowest collection costs possible and the lowest deadweight loss possible. If everyone were to act upon an agreed upon set of rational social welfare preferences, society would agree upon a set of public policies that maximizes social welfare.

It is not apparent that public goods are provided at as low cost as possible, nor is it apparent that that particular public goods are provided by the lowest cost

provider. One might argue, for example, that firefighting is a fairly homogeneous commodity that could best be provided at a national level, utilizing economies of scale in order to provide the same quality and required amount of public good for as low cost as possible. Or, one could argue that firefighting techniques and expectations differ from area to area, where firefighters near airports need familiarity with certain chemicals and firefighters in agricultural areas need to tap ponds as a source of water, so the importance of local knowledge means that local organizations should provide the good. In reality, we often see one type of good provided by one level of government with support from other levels of government, or by multiple levels of government at once. Why should one level of government provide a public good over another? For some public goods, this is sometimes intuitively clear; common sense says that defense should be provided for at the national level. For other goods, it is not obvious. For taxes, it is even less straightforward. Which level or levels of government should collect sales taxes, or income taxes, or excise taxes, and why?

It is not remarkable that different levels of government produce different baskets of public goods; if localness provides informational gains due to closeness and largeness provides economies of scale and power, then we should not be surprised if public goods that rely on customization tend to be provided at the more local levels and public goods that rely on efficiency and force tend to be provided at the highest ranking levels. From a tax perspective, the assortment of taxes collected at different levels of government is even more puzzling; the hierarchy of taxing organizations no doubt creates inefficiency due to redundancy of tasks. In our current information age, it is a wonder why we do not collect the same amount of revenue under one umbrella and costlessly transfer it to the various

government levels, enjoying a great reduction in overhead expenses. Of course, not all government workers are perfectly benevolent public servants, and transfers are not costless.

These naïve ideals break down immediately if one realizes first that the different levels of government may not be aligned in their preferences: the social welfare functions that represent their preferences elicit different rankings of possible policy choices. Furthermore, one can easily imagine the inefficiencies that can arise due to government agents acting in their own interests. One can extend the notion of differing roles in government having incongruent motives to the similar idea that different levels of government may act in ways that hamper the goals and intentions of the other levels of government. Weird and unexpected results that arise due to strategic interactions of different levels of government are of particular interest.

In this paper, I explore the interactions of different levels of government with regard to public policy decisions, particularly asking the questions of who should provide which public good and who should implement taxes. The models presented in this paper are not meant to closely resemble reality. They are abstract along a number of dimensions. First, I simplify the analysis to only emphasize the interactions between but two government entities, which I call the “Empire” and the “Village.” The Village represents a geographical subset of the geographical area of the Empire, and is subordinate to the Empire in that the Empire can dictate some of the actions of the Village, or associate guaranteed punishments with particular actions of the Village. Second, I limit the analysis to the examination of public goods and taxes that pertain to the Village or the stakeholders of the Village. Third, I ignore any complications of representative democracy; here, the Village has no *direct* say in the decisions of the Empire (such as voting power -

though their strategies may cause the Empire to act differently based on strategic considerations). Fourth, the Empire and the Village are non-distinct other than that the Empire is bigger in some general sense, has some unique efficiencies in public goods production, and can set rules for the Village. Finally, I take the abstraction that public goods fulfill certain distinct and understood “roles,” so that one can delineate between policy decisions that fulfill a specific role such as firefighting or policing (this ignores the possibility for combined goods, such as “public safety,” that may be composites of more commonly known public goods such as firefighting and policing). The Village has the benefit of localness, and is beholden to the rules of the Empire. Despite these abstractions, I hope that an analysis of a variety of models will present some insight into the real world.

The models explored in this paper are called games of fiscal management, because a major consideration of these games is which level - the Empire or the Village - will manage (or implement) a particular policy. These policies are meant to fulfill particular roles of government in general, consistent with the assumption of a known set of roles as delineated in the preceding paragraph. I explore two types of games. In the first set of games, the Empire has ultimate control over which level of government implements a particular role, but rebellious actions of the Village can influence the Empire’s behavior. In the second, bureaucrats of different government levels compete for control of roles (rather than the government decision-makers competing). Consistent throughout is the notion that the Empire benefits from economies of scale while the Village benefits from some sense of localness. These games are also consistent in that choices in policy space are among the players’ actions. The primary novelty of this paper is the application of

strategic decision-making to the delegation of public policy roles to different layers of government.

This paper provides a new approach to the growing body of research on fiscal competition. A general review of this literature can be found in Wilson (1999). Horizontal fiscal competition (that is, competition between locales of relatively equal stature) over public goods provision is introduced in Tiebout (1956), which introduces some important insights. First, he acknowledges that the level of local spending in the United States is such that local expenditures cannot be ignored. Second, he notes that the manner in which federal and local governments provide public goods to the people is fundamentally different, in that national governments have to provide goods to a wide swath of individuals with vastly differing preferences, while local public goods tend to be more customized to the specific types of people living in a locality. Third, the key result of the paper is the idea that individuals will move to the municipalities that provide their favored offered set of public goods, provided that said municipalities are not yet at capacity. Thus, horizontal fiscal competition between localities is welfare-improving. Unfortunately, his assumptions of perfect mobility and full relevant information are not particularly realistic. His analysis focuses on quantities and types of public goods, but taxes implicitly linger in the background. While Tiebout does not explicit focus on the role of taxation in his model, Mieszkowski and Zodrow (1989) provide a review of tax extensions to the Tiebout model.

Subsequent to Tiebout, the literature on fiscal competition focuses on externalities created by horizontal and vertical competition (competition among the hierarchy of government levels). Williams (1966) is one of the first papers to call attention to the spillovers between localities when they provide public goods.

Because local public goods may cause spillovers to nearby areas, those areas receive positive externalities. However, if local governments are only concerned with their citizens, they will ignore these externalities, and, as a whole, provide levels of public goods that are too low. On the tax side, horizontal competition may also create externalities for municipalities. If individuals are relatively mobile, high rates in one municipality may incentivize individuals to switch municipalities, lowering the tax base of their original municipality, as discussed in Oates (1972). This results in the condition that, taken as a whole, local tax rates are set too low from a social optimality perspective. Wilson (1999) examines horizontal public goods externalities, and finds that public goods are produced inefficiently, requiring a sub-optimally high level of inputs, while providing a theoretical definition for tax competition of this type. Keen and Marchand (1997) examine the composition of public goods in a horizontal competition setting, finding that more local public goods are provided at low levels that are sub-optimal.

A less explored subfield in public finance concerns itself with the interactions of different government entities in vertical fiscal competition, where competition occurs between multiple levels of government. In general, Oates' treatise concerns itself with interactions of this type. Hoyt and Jensen (1996) investigate the role of the timing of the announcements and commitments to policies at national and local levels, finding that early commitments can be welfare improving. Wrede (1996) investigates the interaction of different levels of Leviathan (revenue-maximizing) governments that share a common income tax base. He finds that, while a system of pure local competition drives Leviathans to lower tax rates so that the net result is that tax rates are on the upward sloping side of the Laffer curve, a combination of vertical and horizontal externalities leads to mixed results. Vertical externalities

are negative, as one level of government taxing a shared tax base reduces the tax base for other levels of government. These conflicting externalities lead to ambiguous results, meaning that uncoordinated Leviathans do not necessarily end up on the wrong side of the Laffer curve. Though not looking at Leviathans, Jametti and Brülhart (2004) investigate the relationship between horizontal and vertical tax externalities among Swiss municipalities, finding that the effects of vertical externalities dominate. Devereux and Redoano (2007) examine both vertical and horizontal externalities in the context of excise taxes, and is the first paper to allow for both types of externalities in an empirical setting with regard to excise taxes. They find that horizontal externalities matter most when the taxed commodities have low transport costs (like cigarettes) and that vertical externalities matter most when the commodities face high transport costs (like gasoline).

A common thread of vertical fiscal competition literature is the idea that local and national governments share the same government roles, for example, income tax collection. What sets this paper apart is the idea that the management or implementation of such roles matter, and, in many cases, there is only one level of government that is the chief implementer (for example, while schools may have to face federal or state restrictions, their ultimate implementation is at the local level in most cases). This creates an all-or-nothing aspect of vertical fiscal competition and is the subject of this paper.

The paper proceeds as follows. In the next section, I describe a class of games of fiscal management that I call games of sabotage. In and of itself, sabotage, unlike other illicit acts like theft, creates net costs for both the sabotaged *and* the saboteur. The mere existence of sabotage implies that it serves a strategic purpose, even if sabotage would never be consumed in isolation. In the section following,

I describe a class of games where bureaucrats of different levels of government compete by bidding for the opportunity to implement some public policy. In these games, the existence of a losing bidder restores consumer surplus. Finally, I conclude.

Games of Sabotage

In these games of sabotage, the Empire, the larger, more efficient layer of government, and the Village, the smaller, more local layer of government, compete for the chance to implement some “role” of government. “Role” here takes a colloquial meaning, and could refer to fire protection, defense, or property tax collection. The players in this game are, abstractly, the Empire and the Village themselves; the Empire and Village each are modeled as one organism, so we can think of these entities as having some unified, decision making process such as a leader or median voter. Regardless of the role of government in question, the role pertains to the domain of the Village; if the role is defense, this can be thought of as defense as it pertains to the Village (even if ultimately produced at a national level), or if it is fire protection, it may be purely local. There are individuals I , $\{1, 2, 3, \dots, i, \dots, I\}$ who are stakeholders of the public policies enacted upon the Village. These individuals do *not* play the game, and include the residents of the Village but may also include outside stakeholders, such as those who trade with the Village or appointees of the Empire who are assigned to work on the Village. The Empire and the Village have social welfare preferences over the utilities $U_I = \{u_1, u_2, u_3, \dots, u_i, \dots, u_I\}$ of the stakeholders of the Village. Each player j , $j \in \{Empire, Village\}$, makes a policy decision. A policy decision is a choice in public policy space $\{T_I, P\}$, where T_I is a tax schedule $T_I = \{\tau_1, \tau_2, \tau_3, \dots, \tau_i, \dots, \tau_I\}$

and P is a vector of public goods (so each “role” can contain multiple public goods, and a tax schedule can contain multiple types of taxes, though the τ s represent net amounts).¹ Utilities U_I are functions of the individual’s tax, the public goods schedule, and other things θ such as private consumption, so $u_i = f(\tau_i, P; \theta_i)$. Each player chooses a policy decision that maximizes $SW_j = f(U_I) = f(T_I, P; \Theta)$, where SW_j is the social welfare function that represents the preferences of player j and $\Theta = \{\theta_1, \theta_2, \theta_3, \dots, \theta_i, \dots, \theta_I\}$. The players’ social welfare preferences may be identical, but it is generally assumed that they are “competing” in the sense of preferring societal outcomes differently.

A further discussion of the policy decision can be found in the appendix. To sum up, each player brings distinct advantages to their policy outcomes. If put in charge of a particular role, the Empire can implement policies at a lower cost than the Village and can take advantage of its large size. On the other hand, the Village has the benefit of having local knowledge of its citizens and can customize policies to increase its citizens’ utility. These policy decisions essentially “fill in” two of the three payoffs that are explained shortly, and, in the game, the policy decisions are known to both players.

The basic game has but two decision nodes. The Empire, being in control over the Village, decides which level of government should implement their policy choice for the government role in question. Village decides whether or not to sabotage the Empire, if the Empire decides that the Empire should implement policy. A further description of sabotage and its related payoffs can also be found in the appendix. The threat of sabotage spurs the dynamics of these games.

¹This abstracts away from the different distortionary effects of different types of taxes. Distinctly different types of taxes can be incorporated into the analysis by changing $\{T_I, P\}$ to $\{T_I^A, T_I^B, T_I^C, \dots, P\}$, where T_I^A, T_I^B , etc. are the different types of taxes, and individual utilities take into account all types of taxes.

Basic Sequential Game

The basic game unfolds as follows: the Empire decides whether to let the Empire or the Village implement. This strategy set is $\{Empire, Village\}$, abbreviated $\{E, V\}$. If the Empire chooses *Empire*, the Village chooses whether to not interfere with the Empire's implementation or to sabotage their endeavor, notated $\{abide, sabotage\}$, or $\{a, s\}$. If the Empire instead chooses *Village*, the game ends (the Village does not have an opportunity to sabotage itself).

This game is represented in extensive form in Figure 4. The payoffs $\pi_{Actions}^j$ refer to the payoff to player j given the history of actions played by each player. For example, $\pi_{E,a}^E$ is the payoff to the Empire if the actions played are *Empire* and *abide*.² Although I shall look at specific parameterizations of payoffs, there is no requirement about the ordering of payoffs in general, except that the nature of sabotage dictates that $\pi_{E,a}^E > \pi_{E,s}^E$ and $\pi_{E,a}^V > \pi_{E,s}^V$.

To fix ideas, let us take the payoff of player j to be five if they get to implement their plan, free of sabotage. If the other player gets to implement, free of sabotage, then a player only receives a payoff of three. If the Empire's plan is sabotaged, then each player receives negative one. These payoffs are shown in Figure 5. These specific numbers represent a typical set of payoffs, as explained in the appendix.

The single shot game with these payoffs is homeomorphic to the famous chain store paradox of Selten (1978). There, the story goes, an upstart store decides whether to enter the market area of an established store, which is taken to be in the more powerful position. The incumbent may retaliate against market entry by

²Due to the multiplicity of strategies in extensions to this basic model, it is more compact in general to denote the payoffs in terms of actions rather than strategies.

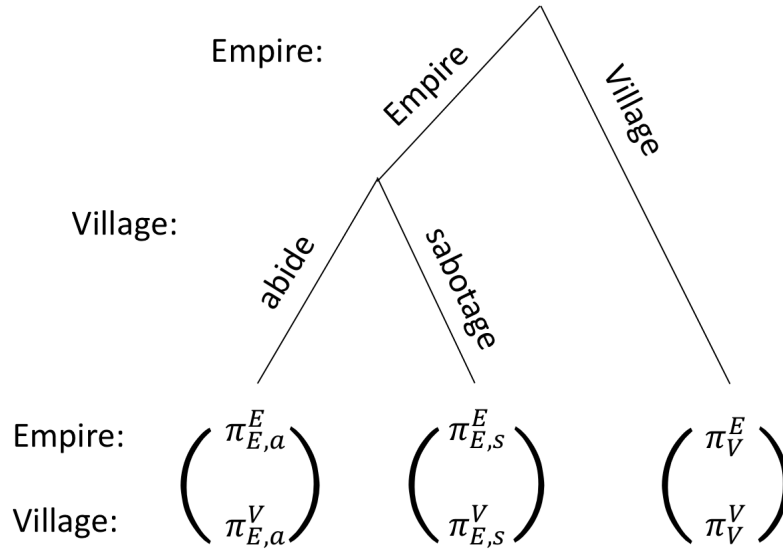


FIGURE 4.
Basic Sequential Game, Generic Payoffs

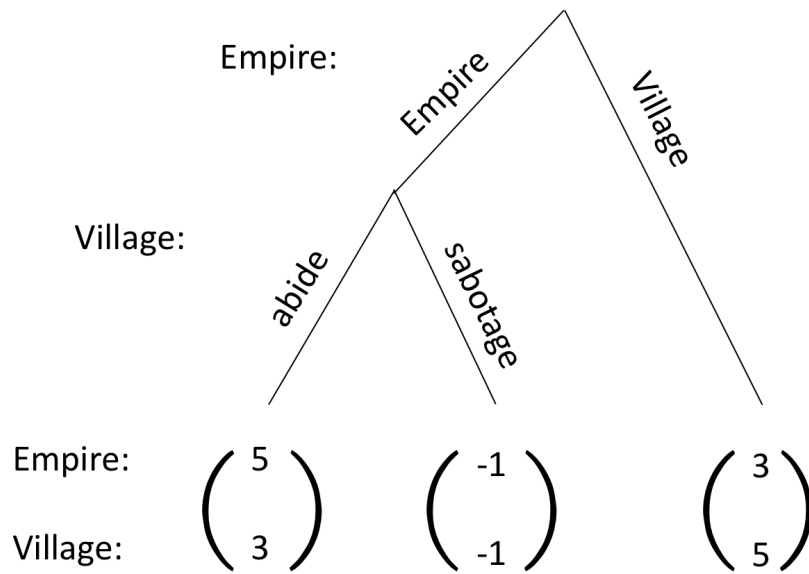


FIGURE 5.
Basic Sequential Game, Specific Payoffs

fighting, which hurts both players. Ironically, here, the weaker Village is the one to respond to the first move of the dominant Empire.

This game has two Nash equilibria, $\{Empire, abide\}$ and $\{Village, sabotage\}$, but the threat of *sabotage* is not credible, making $\{E, a\}$ the only subgame perfect equilibrium. In any single shot game, the Empire will implement all policy roles and the Village will always abide, giving the players payoffs of five and three, respectively.

Reversed Sequence

Suppose the Village can commit to the choice of sabotage or no sabotage before the Empire chooses the policy implementer, and the Village's choice is known to the Empire. Figure 6 shows the resulting game (note that the listed order of the payoffs is reversed, as well as the order of play). Again, there are two Nash equilibria, similar to before (now $\{s, V\}$ and $\{a, E\}$ instead of $\{V, s\}$ and $\{E, a\}$), and again, only one (now $\{s, V\}$) is subgame perfect. By making it clear that the Village will definitely sabotage, fates are reversed and the Village receives its full payoff of five while the Empire receives only three.

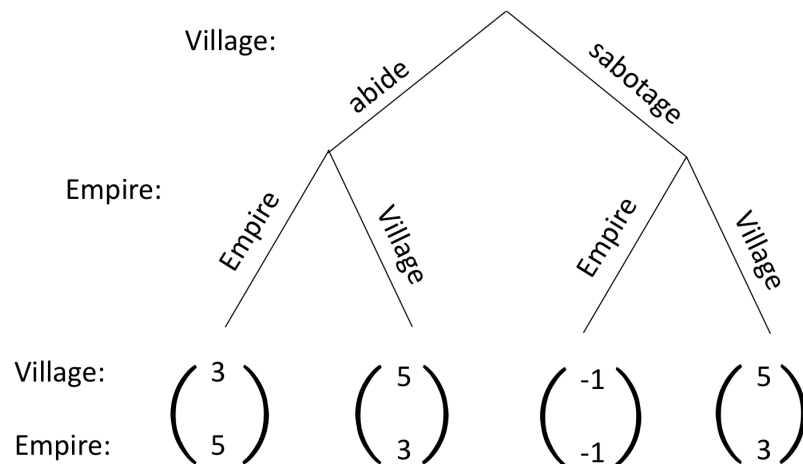


FIGURE 6.
Reversed Order Sequential Game, Specific Payoffs

Simultaneous Play

Suppose that the Empire and the Village choose their actions simultaneously. This can be understood as a situation where sabotage requires advanced planning, before it is known whether the Empire will select *Empire*. The Village's plan to sabotage the Empire is contingent upon the Empire choosing itself as the policy implementer. This game can be represented in normal form, as in the top panel of Figure 7.³ This game has two pure strategy Nash equilibria, where the Empire and the Village play $\{E, a\}$ or $\{V, s\}$. There are no non-degenerate mixed strategy equilibria. The Village's strategy of *sabotage* is weakly dominated and is thus never played if the Empire plays a mixed strategy, leaving $\{E, a\}$ as the only (degenerate) mixed strategy equilibrium.

		Village	
		a	s
Empire	E	5, 3	-1, -1
	V	3, 5	3, 5

		Village	
		a	s
Empire	E	5, 3	-1, -1
	V	3, 5	3, 6

FIGURE 7.
Simultaneous Games, Standard and Modified Payoffs

³For a diagrammatic representation of this simultaneous game in *extensive* form, the reader can view either of the proper subgames in Figure 9.

However, something interesting happens when the payoffs to the Village are changed slightly. Suppose the Village receives some kind of behavioral response to choosing *sabotage*, only to discover that the Empire has chosen *Village* and therefore the Village will not actually have to sabotage the Empire. It is reasonable to believe that, under these circumstances, the Village will experience a feeling of relief, increasing the utility of its citizens and thus increasing the payout to the Village. The modified payoffs are shown in the bottom panel of Figure 7, where the five in the bottom right corner has been replaced by a six. Again, there are the same two pure strategy Nash equilibria as before, but, more importantly, the strategy of *sabotage* is no longer weakly dominated. In fact, the threat of sabotage is so great that the Empire chooses *Village* most of the time. With these particular payoffs, the Village placates with *abide* with two-thirds probability, but the threat of such a low sabotage payoff is so great that Empire only chooses *Empire* with one-fifth probability! This game has an expected payoff of three to the Empire (equal to the Empire's guaranteed minimum payoff), but has an expected payoff to the Village of 4.6, nearly the full five that the Village receives under its most preferred outcome. There are examples of chain store paradox applications where a small chance that the chain store *enjoys fighting* the market entrant is enough to deter most potential entrants (see Kreps and Wilson (1982) and Milgrom and Roberts (1982)), whereas here we have an example where a small premium on *not having to fight* results in the Village nearly always getting its preferred outcome.

Lobbying

I now investigate whether the Village can exhibit a signal, distinct from committing to sabotage, that will induce the Empire to choose the Village's

preferred action of *Village*, giving the Village its highest payoff. In particular, the Village can, at a cost, lobby the Empire. It is assumed that the act of lobbying does nothing to persuade the Empire and it reduces the payoffs to the Village by two. Lobbying does nothing to change the payoffs to the Empire. The idea that sending an expensive signal can possibly benefit the Village is based on Ben-Porath and Dekel (1992), who modify the “battle of the sexes” game to include an initial node where one player may visibly “burn money” in front of the other player. By burning money in their model, their first player reduces all of his payoffs, with the result that only one terminal node leads to a payoff that is greater than his guaranteed minimum payoff if he does not burn money. Their other player, seeing this, realizes that he must only be taking the burning money path if he is expecting his highest payoff down that path. Knowing his intention, their second player chooses her action based on the highest payoff she can get if the first player plays his intention. This way, the money burner gets his preferred payoff, less the cost of lobbying. Their conclusion is arrived at via a *forward induction* argument.

The Village begins the game with a choice of actions, $\{lobby, do\ not\ lobby\}$, or $\{l, d\}$. This is diagrammed in Figure 8. With the pair of actions available to the Empire and the pair of pairs available to the Village, there are now eight strategies available to the Village and four to the Empire, for a total of thirty-two combinations. Due to the multiplicity of strategy profiles, I only focus on subgame perfect equilibria. In this simple sequential game based on the basic game, there is only one subgame perfect equilibrium, where the Village does not lobby, the Empire selects itself as implementer, and the Village abides with the Empire’s plan.

The original burning money example involved non-singleton information sets, so it may be more fruitful to apply the concept of burning money to the

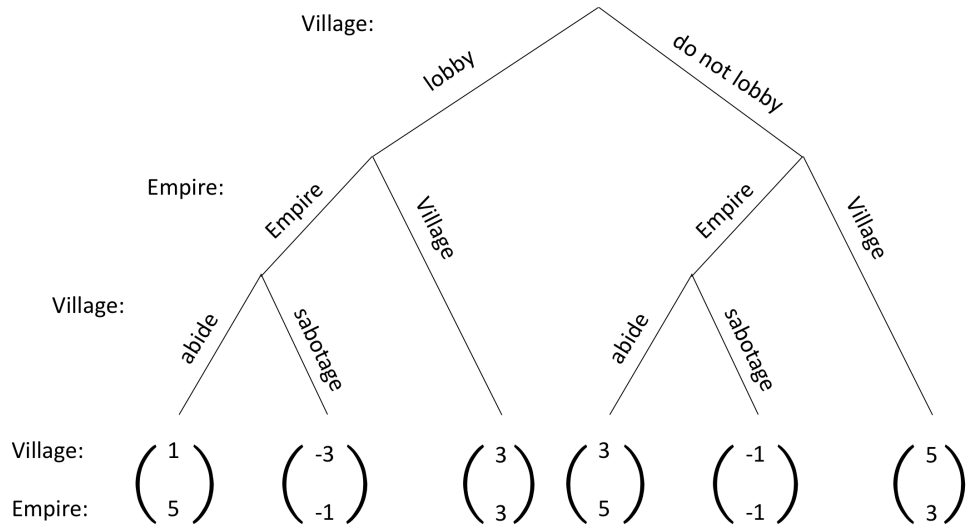


FIGURE 8.
Fully Sequential Game with Lobbying

simultaneous game. This is shown diagrammatically in Figure 9. The logic of the original burning money example is sometimes viewed as contentious (Vega-Redondo, 2003), and it is tough to see on inspection whether it holds up here. One way to show the existence of the burning money equilibrium in the original example is to rewrite the game in matrix form, and successively eliminate *weakly* dominated strategies until the burning money equilibrium (if any) is revealed. The payoff matrix for the simultaneous game of sabotage is shown in Figure 10. The Village has the strategies $\{la, ls, da, ds\}$, which refer to lobby then abide, lobby then sabotage, do not lobby then abide, and do not lobby then sabotage. The Empire has the strategies $\{EE, EV, VE, VV\}$, which refer to select Empire if the Village lobbies and Empire if the Village does not lobby, select Empire if the Village lobbies and the Village if the Village does not lobby, select Village if the Village lobbies and the Empire if the Village does not lobby, and always select the Village. As the modified simultaneous game before, the Village gets a premium for not having to go through with a threat of sabotage.

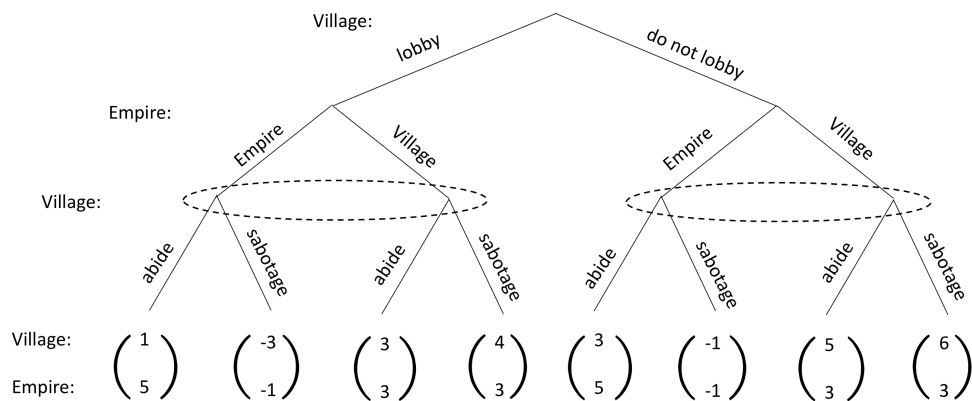


FIGURE 9.
Game with Lobbying Followed by Simultaneous Play, Modified Payoffs

		Empire			
		EE	EV	VE	VV
Village	la	1, 5	1, 5	3, 3	3, 3
	ls	-3, -1	-3, -1	4, 3	4, 3
	da	3, 5	5, 3	3, 5	5, 3
	ds	-1, -1	6, 3	-1, -1	6, 3

FIGURE 10.
Game with Lobbying and Simultaneous Play in Normal Form

The process of removing weakly dominated strategies ultimately proceeds differently than that of the “burning money” version of the battle of the sexes. First, the strategy la is weakly dominated by da , similar to the first elimination in the battle of the sexes. Now, EE and EV are weakly dominated by VE and VV , respectively, meaning that the Empire will always select the Village as implementer as long as the Village lobbies. At this point, no more strategies are eliminated by removing weakly dominated strategies, leaving a three by two game. Solving this game for mixed strategies graphically reveals that there are two mixed or partially mixed strategies of $\{\{la, ls, da, ds\}, \{EE, EV, VE, VV\}\}$: $\{\{0, 0, \frac{2}{3}, \frac{1}{3}\}, \{0, 0, \frac{1}{5}, \frac{4}{5}\}\}$

and $\{\{0, 1, 0, 0\}, \{0, 0, \frac{1}{2}, \frac{1}{2}\}\}$. Interestingly, the Village never gets to implement if the premium for not having to carry out a planned sabotage is taken away (because then all of the Village's strategies are weakly dominated by da). In addition, this game has four pure strategy Nash equilibria of $\{da, EE\}$, $\{ds, EV\}$, $\{ls, VE\}$, and $\{ds, VV\}$, and within the three by two subgame, two pure strategy Nash equilibria of $\{ls, VE\}$, and $\{ds, VV\}$. Just like the battle of the sexes, "burning money" achieves the effect of at least tending to get the initial actor their favored outcome.

Future Work: Repeated Play

By taking the original sequential game as a stage game, it forms the basis for a repeated game. Let there be m rounds of the stage game, with intermediate payoffs to the players paid after each stage. As the number of rounds increased, the number of Nash equilibria gets large, but only one equilibrium is rationalizable. Starting from the final stage m , the Village will either end the game implementing the fiscal policy (as ordained by the Empire), or by choosing whether to abide or sabotage. Maximizing payoffs, the Village will choose *abide*. The Empire, knowing this, knows that the payoff from choosing *Empire* in the final stage will be five, so the Empire chooses this. In the previous round $m - 1$, the Village will understand this, and maximize payoffs by playing *abide* if forced to make this choice. The Empire understands and realizes that *Empire* will result in a higher payoff. This *backwards induction* argument continues back through rounds $m - 2$, $m - 3$ and so on, all the way back to round 1. Thus, in every stage of the subgame perfect equilibrium, the Empire chooses itself to be the implementer and the Village always goes along. Note that this solitary rationalizable equilibrium was arrived at

independently of any understanding of the intertemporal preferences of the players, e.g. their discounting patterns.

If repeated play is infinite, a number of more interesting equilibria emerge. Assuming that players are time consistent in their preferences regarding their intermediate payoffs according to discount factors $\delta_j, \delta_j \in (0, 1)$, each player j maximizes $\sum_{m=1}^{\infty} \delta_j^m \pi_m^j$, where π_m^j is player j 's intermediate payoff from the m th stage of the repeated game. The analysis of these many equilibria is the subject of future work, especially considering situations where players have incomplete knowledge, particularly with respect to the value of the other player's discount factor.

Modern Applications

I have mentioned that these models are abstractions and not meant to emulate the real world. Furthermore, they have some features that are incompatible with a one-to-one comparison with the present condition of the United States, most notably, that I avoid any issues of federalism (the Village has no *direct* say over the decisions of the Empire and can only induce the Empire to act via strategic interaction) and that I only model one Village (the strategic implications of the presence of multiple Villages bears some consideration) and only have one or the other level of government implement policy fulfilling a particular government role. However, this repeated model may be of some assistance in understanding the refusal of some states (who, in this analogy, are like Villages) to refuse Federal funds for the expansion of Medicaid.

Medicaid is a government healthcare program where the Federal government provides funding to the states for the healthcare of poorer individuals and the

states administer these funds. States can opt into or out of the program. Under the Patient Protection and Affordable Care Act of 2010, more commonly known as “Obamacare,” state funds for Medicaid were greatly expanded. From the perspective of the states, these funds could be nearly seen as free money, yet initially, twenty-two states refused these expanded funds, and to this date, fourteen states still refuse these funds (Henry J. Kaiser Family Foundation, 2019). In isolation, these state choices are not rational, implying that either lawmakers in those states are irrational, decision-making processes in those states are irrational (such as legislatures with individual preferences that make voting preferences intransitive), or, more likely, that these choices serve some strategic purpose. In the context of the games of sabotage so far described, refusal of funding is a “sabotage” action that makes states worse off in each stage than if they were to accept the funds. However, if states were to be able to reallocate the taxes used to create funds for Medicaid expansion, and were to control this aspect of the public provision of healthcare, they would pick an altogether different point in policy space. From the dissenting states’ perspective they would be better off in control in the long run. So, while these states may be behaving in a way that seems counterintuitive, irrational, and wasteful, the strategic implications of their acts may be believed on the part of these dissenting states to be leading to an outcome where control over this policy role reverts back to these states, as the Federal government grows tired with their sabotage and realizes that it is better off receiving a string of payoffs where the states are the implementers, rather than continuing to receive a string of sabotage payoffs.

Bidding Bureaucrats

In this section, I describe models where once again, competition is between the Empire and the Village, and the government role in question pertains to the domain of the Village. However, here, the game is static, and the government levels do not compete directly, but rather, bureaucrats - agents of the Empire and the Village - compete for the opportunity to implement a policy role in accordance with their bureaucratic desires. Whereas the Empire had ultimate control before of who is the implementer, here, choice of implementer belongs to some non-player decision maker who chooses the plan that maximizes consumer surplus of the Village. Bureaucrats draft competing bids or proposals, and thus the game bears many similar features to an auction.

The jumping off point for the models in this section is the static budget-maximizing model of Niskanen (1968). I adopt and simplify the notation used by Niskanen. The context of this model is the production of a particular public good, so the choice of taxes shall take a back seat for now (though it is implied that the costs of this public good production are covered, so we may think of these policies as “balanced budget” prescriptions with taxes implicit in the background). The marginal value function, or inverse demand function $MB = a - bQ$ is the vertical sum of the marginal values to the residents of the Village. It will simplify analysis to consider cases where a bureaucrat's of player j 's agency produces public goods at constant marginal cost, or $MC_j = c_j$. Constant marginal cost means that the units of quantity and currency can be redefined so that the linear marginal benefit curve is given simply by $MB = 1 - Q$. In the one-government model, the bureaucrat maximizes his budget $B_j = Q_j - \frac{1}{2}Q_j^2$ by offering a level of production that the public is just indifferent to receiving - he offers a quantity greater than the socially

optimal level in such a way that the consumer surplus is obliterated. This is shown in Figure 11, where area Z shows the area of maximum consumer surplus (and the surplus gained under theoretical perfect market competition for government services). Area Z is perfectly offset by area Z', and the bureaucrat chooses a budget producing Q^* units of the public good (Niskanen imposes the additional constraint that the budget must at least equal the total costs of production $TC = cQ^*$). If marginal cost is less than one-half, then an additional restriction is that the bureaucrat may not produce public goods at a quantity for which there is no positive marginal benefit. In stark contrast to a market valuation situation in which a monopolist produces a quantity that is too low relative to the social optimum, the sole bureaucrat produces a quantity that is too high relative to the social optimum.

With two bureaucrats competing for the opportunity to produce one public good, they shall share a demand curve but have differing cost functions. The Empire, being large and benefiting from economies of scale, can produce the good at lower marginal cost than the Village, or $c_E < c_V$ before local pride is considered. However, the Village's local pride means that the Village can produce the same quality good as the Empire, but it will provide more consumer surplus due to local pride. For each unit of the public good produced, citizens of the Village put a local premium on that good of ℓ . We *could* treat this as the Empire and Village bureaucrats as interacting with different demand curves $MB_E = 1 - Q_E$ and $MB_V = 1 + \ell - Q_V$, but it is far simpler to consider ℓ to be incorporated in the Village's net marginal cost, so that ℓ lowers c_V . After local pride is considered, it is possible that $c_V < c_E$, or that still $c_E < c_V$. The game with two competing bureaucrats is shown in Figure 12.

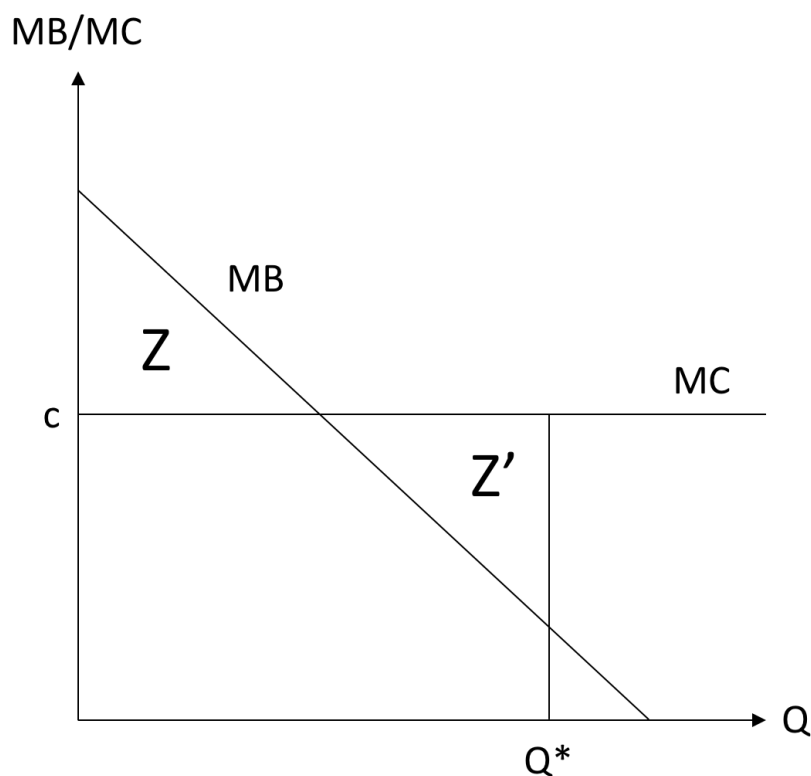


FIGURE 11.
Basic Niskanen Model with One Level of Government

The game works as follows: the Empire's bureaucrat and the Village's bureaucrat simultaneously each choose a quantity Q_j , with no knowledge of the other's bid. Costs and demands are known to all. The decision-maker reviews the quantity bids and chooses the implementer that maximizes consumer surplus for the Village. In this sense, the decision-maker acts in a deterministic manner, so the players of the game are really just the two bureaucrats. If a bureaucrat's bid is chosen, she receives his planned budget B_j . If not, she receives nothing - this is an all or nothing game. The decision maker rejects either bid if a bid creates negative net consumer surplus.⁴

⁴If the decision maker is indifferent between the two proposals, a tiebreaker rule of some sort can be implemented - either the decision maker chooses one plan at random, or has lexicographic preferences over the two bureaucrats in case of a tie.

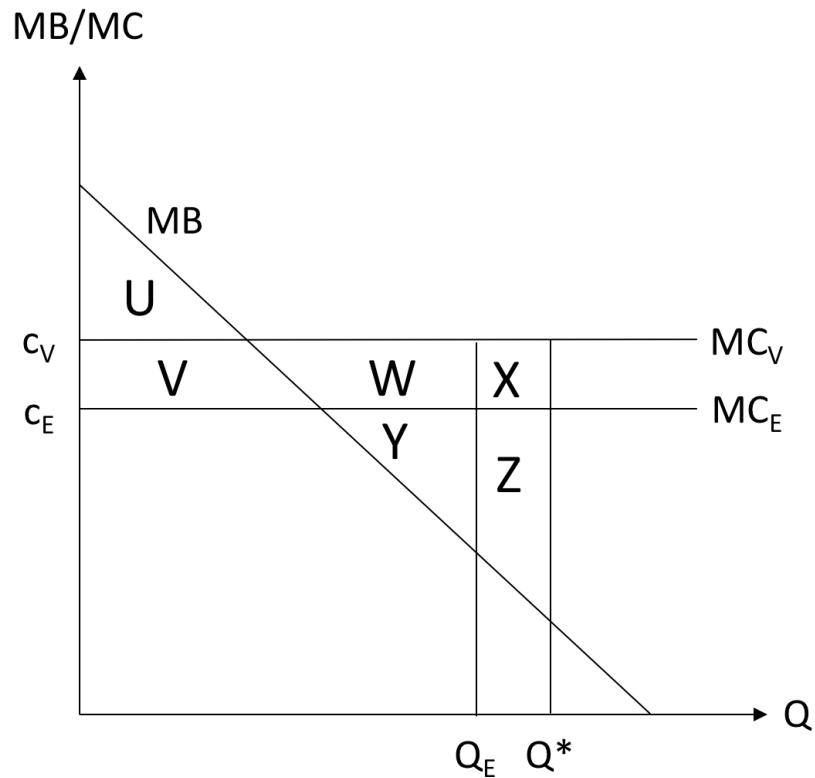


FIGURE 12.
Niskanen Model with Competing Levels of Government

The following analysis refers to the cases where each $c_j > \frac{1}{2}$, but the case of $c_j < \frac{1}{2}$ produces similar results using similar logic, just with slightly different algebra. If $c_E < c_V$, the Empire has the distinct advantage of being big and operating at lower cost, and it is clear that the Empire can force a win. First, consider that no player benefits from a bid of less than $1 - c_j$, the socially optimal (consumer surplus maximizing) quantity given their marginal cost. No player can bid more than $2(1 - c_j)$, the point at which any larger quantity leads to a situation where total costs TC_j exceed budget B_j . Now consider any quantity choice of the Village, $1 - c_V \leq Q_V \leq 2(1 - c_V)$. Because the Empire has the cost advantage, it can always best any resulting consumer surplus CS_V implied by the Village's bid (this is easy to see, because, given $c_E < c_V$ the Empire has the largest maximum

possible consumer surplus CS_j . For any Q_V bid, where $Q_V > 1 - c_V$, the best response of the Empire is to bid the quantity Q_E that produces CS_E such that $CS_E = CS_V$. As the maximum $CS_V = \frac{1}{2}(1 - c_V)^2$, this drives the Empire to produce $Q_E = 1 - c_E + \sqrt{(1 - c_E)^2 + (1 - c_V)^2}$. This can be seen in Figure 12. Without competition, the Empire would produce at level Q^* , as before, wiping out consumer surplus in areas U and V with areas Y and Z. Here, because the Village can offer maximum consumer surplus of area U, the Empire offers Q_E , where consumer surplus of U and V is only offset by area Y (so that $U+V-Y=U$, or $V=Y$), greatly reducing the loss in consumer surplus to the Village, as compared to the theorized perfectly competitive outcome. (If $c_E > c_V$, then the above analysis can be reversed by simply switching the E and V subscripts.)

Supposing $c_E < c_V$, the Empire can always win.⁵ However, as the (net) c_V approaches the value of c_E , the bureaucratic overproduction of the Empire shrinks. This result has echoes of the result of Bertrand competition in a duopoly for market goods (Mas-Colell et al., 1995). There, the entrance of a second supplier drives the duopolists from producing a too-small quantity to producing the perfectly competitive market quantity. Here, quantity competition in an all-or-nothing setting drives the competing bureaucrats from producing too much of a quantity to producing the perfectly competitive market quantity, in the case where $c_E = c_V$. Even more remarkably, because ℓ has the effect of lowering c_V closer to c_E , a mere increase in local pride can be welfare improving for the Village, even if their bureaucrat is never chosen to implement policy. Niskanen (1971) employs

⁵The exact bid may depend on tiebreaker rules - if the decision maker chooses a plan randomly if he is indifferent, the Empire can guarantee a win by bidding a quantity slightly less than $Q_E = 1 - c_E + \sqrt{(1 - c_E)^2 + (1 - c_V)^2}$.

a vertical fiscal competition variant of his model, except that the roles of levels of government in his extension are based on exogenously determined tax shares.

Conclusion

In this paper, I abstractly examine the role of the strategic interactions of levels of government in circumstances where only one level of government is tasked with the implementation of some public policy role. In the first set of models examined, the possibility that a subordinate level of government may sabotage the plans of the dominant level creates a variety of interesting outcomes, where results depend on order of play, simultaneity of play, and repetition of play. In the second set of models, I examine the strategic interactions between two levels of budget maximizing bureaucrats. The mere existence of competition between government levels is welfare-improving, even if the dominant level of government always gets to control policy. As we live in a complex world where a hierarchy of governments are nested within one another, it is important to continue to analyze vertical fiscal competition. Future research shall include further extensions of the models explored in this paper, with particular emphasis on the importance of imperfect information and intertemporal tradeoffs.

CHAPTER V

CONCLUSION

In this dissertation, I have examined facets of public finance relating to the design, use, and implementation of public finance policy outcomes. These aspects of the public finance system impact the daily lives of citizens.

In Chapter II, I examine the movements of stock prices around changes in state corporate income tax rate laws. I find that tax increases of one percent are associated with a decline in stock price of about ten percent. Responses to tax decreases have a similar inverse relationship but are not significant at conventional levels of statistical significance. These results are paramount as state lawmakers must decide what to do with the blend of taxes in their state. My findings suggest that publicly traded companies in states suffer greatly when corporate income taxes are raised, but do not benefit as much as tax decreases may suggest.

Chapter III includes previously unpublished co-authored material. In this chapter, co-authors John Voorheis, Caroline Weber, and I examine wage misreporting by comparing wage income data from the Current Population Survey and administrative records. We find that the gaps between these two sets of data are related to the demographic characteristics of respondents. We find patterns for age, educational attainment, and some racial and ethnic groups. Within all groups examined, survey respondents with lower wages tend to overreport their wages to the CPS, while respondents with higher wages tend to overreport. A great proportion of these gaps are created by rounding numbers, but our results maintain even after accounting for rounding. Our results have important implications

for research into the gender wage gap, racial and ethnic wage gaps, and the measurement of income inequality.

Finally, I augment the empirical chapters with a theoretical chapter. In Chapter IV, I use game theory to model the interactions between multiple levels of government, a less studied aspect of public finance. One set of models deals with the possibility that a subordinate level of government may subvert the policy goals of a larger level of government through sabotage. Depending on the structure of the game, there are times when the subordinate level will sabotage the dominant level, and times when the dominant level will let the subordinate level implement policy to its liking. Finally, a second set of games explores the interactions of bureaucrats. I find that when bureaucrats of different levels compete over the opportunity to implement policy, the results are generally welfare improving.

APPENDIX

Corporate Income Tax Data Collection

In this section, I further detail the data collection process. In addition to using *The Book of the States* for identifying tax rate changes, I also use similar spreadsheets from the Tax Foundation as a double check. Both of these sources list many rates that are different than the actual statutory marginal rates firms would eventually pay, due to the simple fact that a large number of tax law changes are made retroactively, whereas these two sources are forward-looking. Many states have additional types of taxes on corporations, such as franchise taxes, but they are deemed to be too small to worry about. The business tax systems of Michigan, Ohio, and Texas are generally ignored because their alternative tax forms are not directly comparable to taxes on net income.

The SIC codes used in identifying potential bank stock for the sample are codes 6021, 6022, 6029, 6035, and 6036. Some additional criteria are used for dropping firms from my sample. Firms are only kept if they are listed on the New York Stock Exchange (NYSE), the National Association of Securities Dealers Automated Quotation exchange (NASDAQ), the NYSE American exchange (f.k.a. the American Stock Exchange or AMEX), and the NYSE Arca exchange (f.k.a. the Archipelago Exchange or ArcaEx). Firms are also dropped if they are accidentally coded with the wrong SIC code, if their business resembles some type of financial service company other than a community bank (such as a credit card processing company), if they are duplicates, if no stock price data is available or if they have less than one year of data available, and if they have operations in foreign

countries. In general, stocks are identified with their CRSP permanent number or PERMNO, which is the only key unique to a specific financial instrument. When Compustat data are needed, a crosswalk is made between their PERMNOs and their Committee on Uniform Security Identification Procedures (CUSIP) numbers.

Banks' 10-K filings on EDGAR are generally a good source for the number of bank branches in each state. In some cases, the firm could not be found on EDGAR, so filings or annual reports are accessed on Morningstar.com or the company's website. The company profile on Yahoo! Finance is used as a last resort. Firms are dropped if no source of location data is found. When firms list the states of their operations but not the number of branches, values for the proportions of income for those states are set to missing while values for states with no operations are set to zero. If a firm lists a string of known locations "and others," or uses some other vague description, the firm is dropped.

Although income is apportioned to each state using a simple assumption of same payroll, property, and sales at each branch, I do heed state nexus throwback rules. Throwback rules work as follows: if a firm is located in a state with a throwback rule and earns income from a state with no corporate income tax, income from the non-taxing state is apportioned to the taxing state, or "thrown back." Schedules of these state-by-state rules are found via the Tax Foundation.

Estimating the Growth Rate and Rate of Return

Multiple times in this paper, I make use of g , the growth rate of the earnings of bank stocks, and r , the required rate of return on these stocks. Values for these parameters are not readily available, so I estimate them.

To estimate g , I assume that publicly traded regional banks' earnings grow at a constant rate. I identify nineteen states whose top corporate tax rates for financial institutions did not change from 1994-2017 (including zero rate states). I then identify 85 firms in my sample of bank stocks whose income only comes from these nineteen states. This is to ensure that changes in state corporate tax rates do not directly effect these firms' earnings. Using the CompuStat database, accessed through WRDS, I download the daily listings of earnings per share (EPS) for these stocks. These EPS are only updated quarterly, at different dates for different firms. In order to adjust for these timing differences, I collapse the mean of these EPS within each quarter for each stock (essentially smoothing out the EPS). Because the financial crisis beginning in 2007 was an outlier event, I only keep quarterly EPS values up until Q1 2007. I drop the few stocks that have negative values for these smoothed EPS. I then take the logarithm of these smoothed EPS and regress them on an index for each quarter in a time series regression, with fixed effects for each stock. This provides an estimated coefficient on the quarterly index variable of 0.0162, with a p -value of 0.000. This quarterly rate corresponds to an annual rate of 6.7%, which I round down to 6.5% due to the dropped stocks with negative EPS.

To estimate r , I assume prices of bank stocks can be approximated by the aforementioned dividend growth model. For the same set of eighty-five stocks, I download daily stock prices and EPS from CompuStat. I create a variable which is the current EPS divided by the current price for all stock/day combinations. This variable should be equal to the expression $r - g$. I take the mean of this variable and substitute my estimate of g , yielding a quarterly estimate of r of 0.0321, which corresponds to an annual rate of 0.1349. I also estimate r by rearranging equation the dividend growth model so that

$$d_0 = \left(\frac{r - g}{1 + g} \right) P.$$

I regress d_0 on P , forcing the intercept to be zero. This provides an estimate of $\left(\frac{r-g}{1+g}\right)$ of 0.0159 with an R-squared of 0.91. Plugging in my estimate of g , I solve for r , obtaining a quarterly estimate of 0.0325, which corresponds to an annual estimate of 0.1364. I take the two annual estimates of r and round them to 13.5%.

Comparing Tax Law Changes

Comparing different tax laws would be easy if one single new rate went into effect immediately after tax rate change laws were passed. Comparisons are complicated by the fact that most such laws go into effect months after they are passed, and tax rate change laws are often bundles of multiple rates that get phased in over multiple years. To compare these different laws, I use a baseline model that relies on a *ceteris paribus* assumption: that net incomes before state corporate taxes of my sample stocks will keep on their current growth path. This means that the only presumed difference comes from differences between the old top rate(s) and new top rate(s).

I fix the date that a tax law change “happened” as the date the bill was signed into law, rounded to the nearest quarter end. I move the date(s) a law will go into effect to the following January 1st, as the vast majority of stocks in my sample have a December 31st. I create a schedule for each tax law change of the number of quarters left (as of the signing date) under the current top rate, the number of quarters under the first new top rate, the number of quarters under the next new top rate, and so on, until at last there is a “final” rate that continues

indefinitely. I make a similar schedule for the top tax rate(s) that were in effect before the new tax law change was made.

I assume a dividend growth model.

$$P_{it} = \sum_{k=1}^{\infty} (1 - \tau) d_0 \left(\frac{1 + g}{1 + r} \right)^t, \quad (\text{A.1})$$

Normalizing d_0 to unity, I create a column of the next 1,000 quarters of earnings, growing at rate g , assuming no tax. This sum approximates the value of the infinite sum to within a tolerance of 10^{-6} . I create another column that discounts these earnings at rate r . The sum of this second column gives the present value of a prototypical firm's earnings under the assumption of no tax. For each tax law change the sample for which stocks exist with fifty percent or more of operations in the state in question, I make two columns: one for the tax rate τ for each quarter as was scheduled before the law changed, and one with the new rates after the change. I create companion columns for each of these two columns that adjust the original column of earnings by a factor of $1 - \tau$. The sums of these tax-adjusted columns give the present value of a prototypical firm's earnings under the old tax law and the new tax law. The difference of the differences between these sums and the no-tax column's sum gives the percent change, adjusted for the time value of money and future earnings growth. In the case of retroactive tax law changes, a small additional "sweetener" was calculated and added to the after-law-change-column's sum, to take into account the sudden realization of profits from unanticipated increased earnings from quarters past. In future research, it would be interesting to see if such sweeteners are more commonly associated with smaller tax cuts, as this kind of sweetener could be a way to placate people for tax cuts of underwhelming magnitude.

Event Studies

This section documents one of the additional methodologies used to try to determine the impacts of corporate income tax law changes on stock prices, namely financial event studies. The financial event study, developed by Fama, et al. (1969) is a common financial inference tool that employs the use of first differences in stock prices to determine unusual stock return behavior related to an event. Unlike the large number of relatively recent corporate income tax rate changes, tax events often occur in isolation, making causal inference difficult, as the impacts of the tax event may not be distinguishable from other contemporaneous events. For example, Lang and Shackelford (2000) are restricted to this kind of analysis. Event studies of this type are also very useful for visualizing changes in stock prices in time around an event, as time unfolds. Thus, they are useful for pinning down when an impact is felt in the markets.

The intention of an event study exercise in the context of changes in corporate income tax rates is twofold: first, to visualize changes in stock prices, and second, to square this paper with some of the rest of the literature. However, while most of the conclusions derived from these event studies conform to the conclusions derived from the first-differenced regressions, the event studies I created were problematic and in some instances, produced results that are at odds with some of the regression results. They are documented here in the interest of completeness and in the hope that their shortcomings may eventually be overcome. The flaws in these particular event studies are likely caused by a small sample size, a treatment variable that is not restricted to $\{0, 1\}$, and the fact that the timelines for the law changes are all different (e.g. the number of days between introduction of a bill and the passage of that bill is different for all law changes).

The procedure used for this technique is as follows: I regress stocks' returns on market returns (and other factors) well before the occurrence of some event; this establishes a model of how the stock typically moves with the market; then, in an event window surrounding the event in question, I use the model to predict the expected returns for the stocks; deviations from the model's prediction are dubbed abnormal returns, and the sum of these, dubbed cumulative abnormal returns (CAR) explain the portion of returns' behavior due to the event in question. When events for different stocks are associated with different dates, event dates are all normalized to be time zero, effectively stacking multiple events on top of one another and taking the average CAR at each point in time. The time periods involved are known as the "estimation window," the time period over which the model is run, the "gap," a space between the estimation window and the period to be studied, and the "event window," which is the period under study, centered around a particular event.

I create the event studies in this paper using WRDS' built-in EVTSTUDY module, using the Carhart four-factor estimating model (Carhart (1997)),

$$\hat{R}_{it} = \alpha_i + RF_t + \beta_{1,i}(R_{Mkt,t} - RF_t) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t,$$

where \hat{R}_{it} represents the expected return to stock i over some time frame t , based on actual market return $R_{Mkt,t}$ (which usually comes from a broad-based stock index such as the S&P 500), RF_t is the risk-free rate (to adjust for risk), and there are factors for the return premiums of small company stocks over larger stocks SMB_t (based on market capitalization), high book-to-market-value stocks ("value stocks") over low book-to-market-value stocks ("growth stocks") HML_t , and

momentum MOM_t , which adjusts for the markets' tendency to stick with winners over losers.¹

I run event studies separately for tax increases and decreases, for both using each of the five dates for each tax law change's legislative history, for a total of ten studies. For all event studies, I estimate the Carhart model using an estimation window of five hundred days, a gap of three hundred days, and an event window of thirty days on either side of each event. The long estimation window was chosen so as to establish a strong relationship between the stocks in my sample and the market, and the long gap was chosen so that the estimation window avoids other events in the legislative history of the same tax law change.² The event study stacks the abnormal returns of multiple stocks on top of one another, which is problematic for visualization if the impact of a change on one stock is tiny (i.e. if a small proportion of a firm's bank branches are in the state of a law change) - then the combination of returns for these smaller impact stocks should visually mitigate the stock returns for firms with a larger tax impact, meaning that some kind of cutoff should be chosen. For all event studies, I only include firms whose treatment effect (log change from Table 5 multiplied by the proportion of that firm's operations in the affected state) is greater than 0.005 (in other words, if the

¹A number of other models are commonly used. Event study models available in WRDS' EVTSTUDY module include the market model (Fama et al. (1969)),

$$\hat{R}_{it} = \alpha_i + \beta_i R_{Mkt,t},$$

the Fama-French three-factor model (Fama and French, (1992)),

$$\hat{R}_{it} = \alpha_i + RF_t + \beta_{1,i}(R_{Mkt,t} - RF_t) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t,$$

the market-adjusted model, which adapts the market model by forcing β to equal unity, and the above described Carhart model. These other models produce substantially similar images.

²For example, if the gap were shorter, and I run an event study on one of the later dates in a law's legislative history, estimation of the model may be tarnished by stock price changes due to the anticipation of said law's passage.

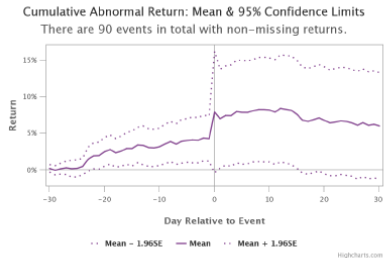
profits of a firm are expected to change by more than half of one percent, *ceteris paribus*). The number of stocks applicable to each tax change varies even within a single tax change's legislative history, as stocks may have been trading during the estimation window or event window relative to one date but not another. The result of this is that each event study utilizes only dozens of stock observations, rather than the thousands utilized in the first-differenced regressions. This leads to weak power and rather noisy event studies, with results driven by companies with operations that tend to be concentrated in states where the largest tax changes occurred.

Figures A1 and A2 show the results of the event studies. Tables 3 and 4 are helpful in figuring out how each of the five panels in each figure fit together with one another across time. Figure A1 shows the CAR for tax decreases for firms with treatment sizes greater than 0.005 in magnitude. Notably, there is an upward spike exactly coinciding with the first news date, suggesting that markets are incorporating news of impending tax law changes in close to real time. However, this spike is largely driven by firms located in the state of New York (an event study, excluding New York tax changes (not shown) looks similar, but without the dramatic spike). Also, there is some upward movement in stock prices beginning about twenty days prior to lawmakers publicly announcing interest in a tax cut. This could be due to noise, misspecification of the first news date for some law changes, or investor pre-knowledge of upcoming, unannounced events. After the initial spike, prices seem to be somewhat mean-reverting. The remaining panels in Figure A1 are mostly flat and noisy.

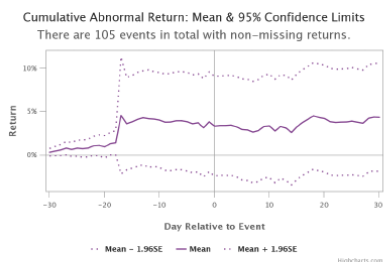
Figure A2 shows the CAR for tax increases. In general, the CAR exhibit a downward trend, as predicted. However, significant negative CAR do not appear

FIGURE A1.
 Event Studies, Tax Decreases, Firms with Log Treatments Greater
 Than 0.005 in Magnitude

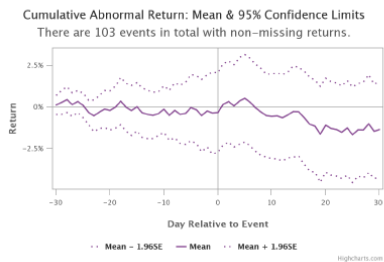
(a) Decreases, First News Date



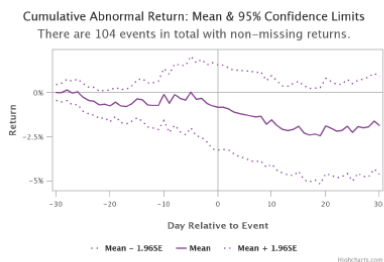
(b) Decreases, Legislation Introduced



(c) Decreases, Legislation Passed



(d) Decreases, Legislation Signed



(e) Decreases, Legislation Effective

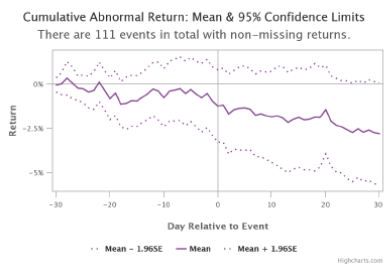
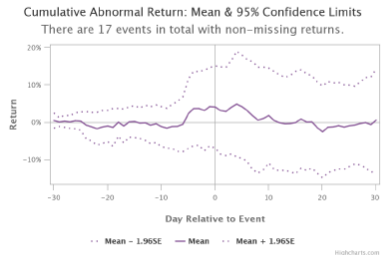
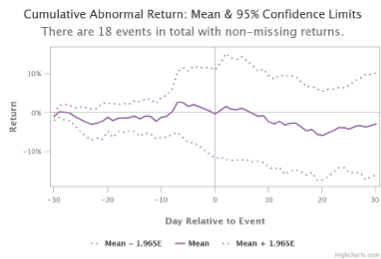


FIGURE A2.
 Event Studies, Tax Increases, Firms with Log Treatments Greater
 Than 0.005 in Magnitude

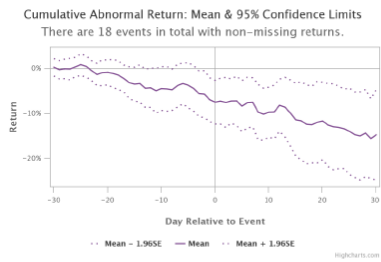
(a) Increases, First News Date



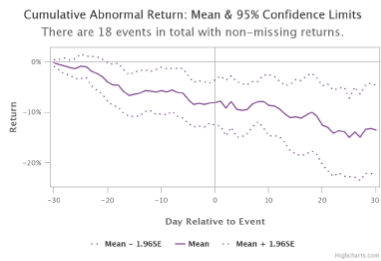
(b) Increases, Legislation Introduced



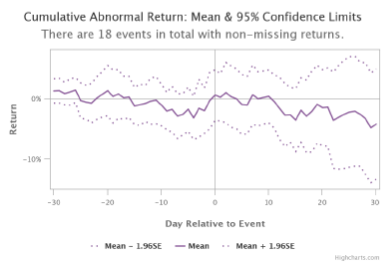
(c) Increases, Legislation Passed



(d) Increases, Legislation Signed



(e) Increases, Legislation Effective



until panels (c) and (d), around the dates tax increase bills are actually passed. Although these images are more descriptive than causal, they suggest that markets believe potential tax decreases will occur as they are announced, but markets wait to incorporate news of tax increases until they are nearly certain (and are possibly slower to do so even after they are signed into law). As is discovered in the main body of this paper (from the within-subwindows regressions), this latter suggestion is likely erroneous. The images in Figure A2 are noisier than those in Figure A1, as the event studies for tax increases only employ about one fifth of the firms (there are far fewer tax increases in the sample).

The timing of the reactions of stock prices, coupled with the EMH, suggests that lawmakers signal potential tax decreases well before they are introduced as bills, and that these signals are heeded. This is consistent with the conclusions drawn from the within-subwindows regressions in the main body of the paper. On the other hand, the timing for increases based on these event studies suggests they do not send signals in advance or that the signals are unheeded early on in the legislative process. While this disagrees with the findings of the within-subwindows regressions, it is possible that the subsample of firms with absolute value treatments greater than 0.005 behave differently than other firms.

Perhaps shockingly, the magnitudes of the CAR seem quite large relative to the size of the tax rate changes, even when compared to regression results. However, these figures suffer from the wide range of treatment sizes, which in these figures range in (log) magnitude from 0.0050 to 0.0370. Therefore direct comparisons about the magnitude of changes rather than the direction of changes are difficult.

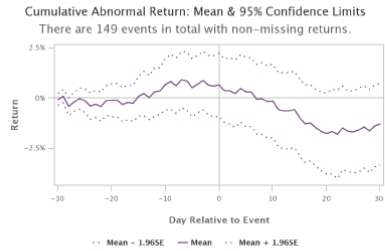
I challenge my event study findings, as tenuous as they may be, with a series of placebo event studies. For the first set of tests, I rerun the original event studies, but substituting twenty random stocks from my sample of bank stocks for each event, conditional on the random stocks having no operations in the states in question. Flat, insignificant results should help to alleviate concerns that my results in the previous section were driven by events where the banking sector as a whole diverged from its usual relationship with the market. Figures A3 and A4 show these results for tax decreases and increases, respectively. In general they are flat and insignificant, however, they do sometimes show trends and occasional spikes. These movements in what should be a null result show that these highly correlated stocks and the banking sector as a whole can diverge from the market sometimes in unpredictable ways.

Next, I rerun the original event studies using the original companies, except that I shift all dates backwards by four years. I choose this date because it is far before the events in question (so that the event window of later dates like the date the law becomes effective do not coincide with earlier actual dates, such as the first news date). Figures A5 and A6 display these results for tax decreases and increases, respectively.

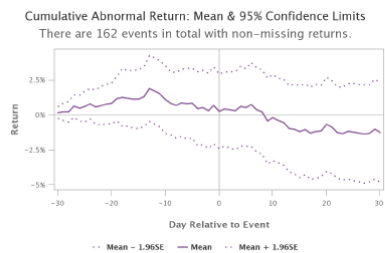
These results are less convincing. While most of these are flat and insignificant, the CAR in panels (c) and (d) of Figure A6 veer sharply downward. The basket of stocks used to make this graph come from only six states, and thus have highly correlated returns within states, so the coincidence of unusual events in just a few of these states could significantly drive results. As these same stocks were used in the original event studies, this means that concerns that random events are driving results are not alleviated. While the event studies are helpful

FIGURE A3.
Event Study Placebo Tests, Tax Decreases, Using Random Firms with
No Operations in State of Tax Law Change

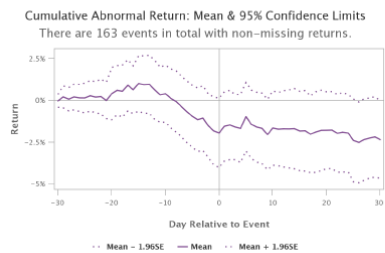
(a) Placebo for Decreases, First News Date (Random Firms)



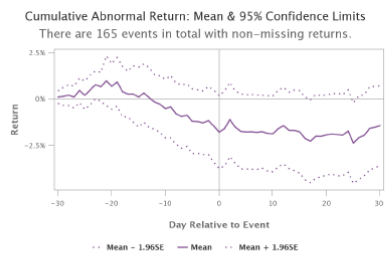
(b) Placebo for Decreases, Legislation Introduced (Random Firms)



(c) Placebo for Decreases, Legislation Passed (Random Firms)



(d) Placebo for Decreases, Legislation Signed (Random Firms)



(e) Placebo for Decreases, Legislation Effective (Random Firms)

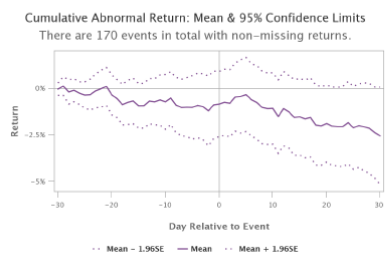
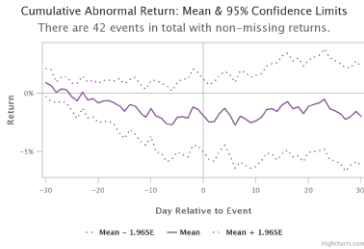
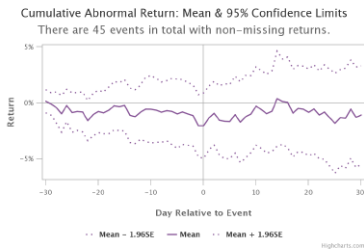


FIGURE A4.
Event Study Placebo Tests, Tax Increases, Using Random Firms with
No Operations in State of Tax Law Change

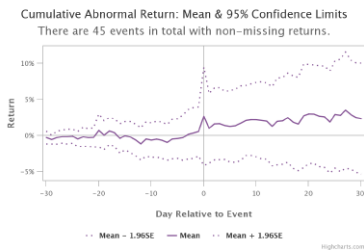
(a) Placebo for Increases, First News Date (Random Firms)



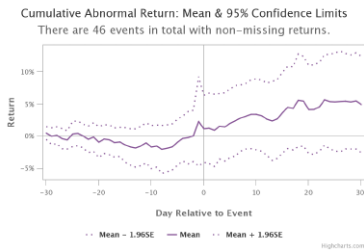
(b) Placebo for Increases, Legislation Introduced (Random Firms)



(c) Placebo for Increases, Legislation Passed (Random Firms)



(d) Placebo for Increases, Legislation Signed (Random Firms)



(e) Placebo for Increases, Legislation Effective (Random Firms)

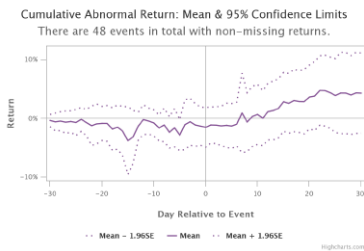
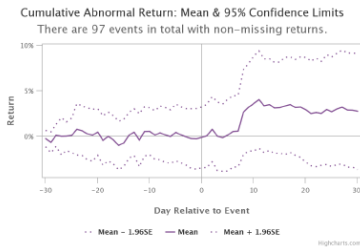
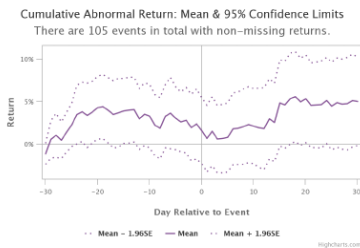


FIGURE A5.
 Event Study Placebo Tests, Tax Decreases, Shifting Event Dates
 Backwards in Time by Four Years

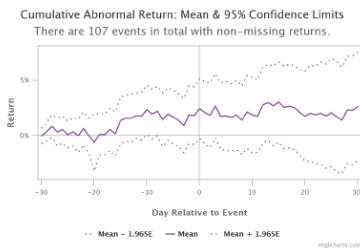
(a) Placebo for Decreases, First News Date (Shifted in Time)



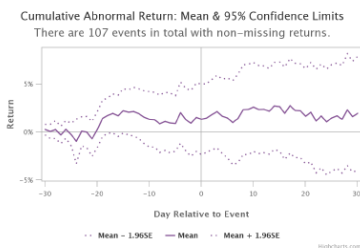
(b) Placebo for Decreases, Legislation Introduced (Shifted in Time)



(c) Placebo for Decreases, Legislation Passed (Shifted in Time)



(d) Placebo for Decreases, Legislation Signed (Shifted in Time)



(e) Placebo for Decreases, Legislation Effective (Shifted in Time)

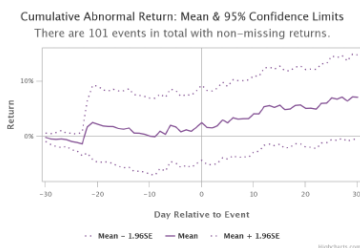
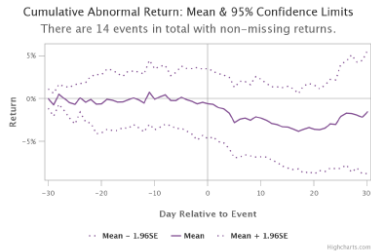
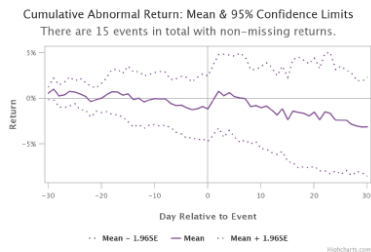


FIGURE A6.
Event Study Placebo Tests, Tax Increases, Shifting Event Dates
Backwards in Time by Four Years

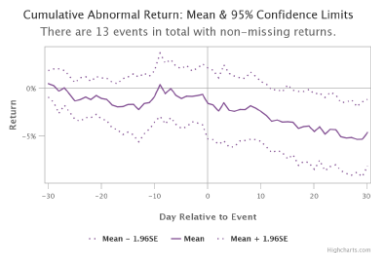
(a) Placebo for Increases, First News Date (Shifted in Time)



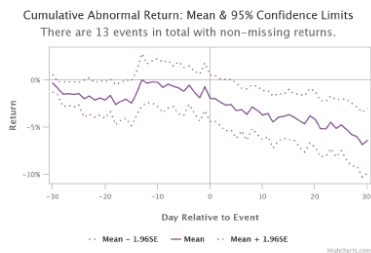
(b) Placebo for Increases, Legislation Introduced (Shifted in Time)



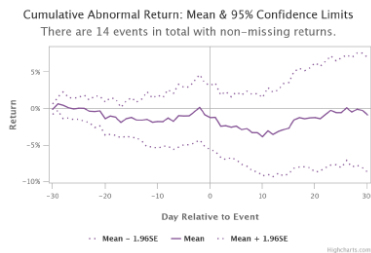
(c) Placebo for Increases, Legislation Passed (Shifted in Time)



(d) Placebo for Increases, Legislation Signed (Shifted in Time)



(e) Placebo for Increases, Legislation Effective (Shifted in Time)



as a diagnostic and descriptive tool to help visualize changes in stock prices brought on by changes in tax laws, they make inference difficult and occasionally produce results that seem to be at odds with the regressions.

Cleaning of Matched CPS/Administrative Data

Table A1 documents the evolution of the sample size as the sample was cleaned from all matchable records to those that are able to be neatly compared using all five income measures (the description of these measures is in section A.6). The sample is restricted along four criteria for comparability, each resulting in the pruning of between two and twenty-eight percent of the existing samples. The age criterion (made to avoid unusual behavior at either end of the age spectrum) is the most restrictive, eliminating thirty-one percent of the existing sample.

**TABLE A1.
Sample Evolution**

Sample	Records Remaining	Percent of Previous Row
CPS/IRS/DER matched unit records	1,034,000	
Units receive no Social Security income of any type	855,000	83
Units represent working age non-dependents	591,000	69
No income amounts are imputed by the CPS	424,000	72
Units appear on only one 1040 per year	417,000	98
Units do not file a Schedule D	349,000	84

Source: CPS ASEC, IRS W-2 and SSA DER, 2001-2016. Approved for release by the Census DRB, authorization numbers CBDRB-FY18-143, CBDRB-FY18-200 and CBDRB-FY18-407.

Variable Construction

The five measures of income that are comparable between CPS and administrative records are: wages, interest plus dividends, total money income, self-employment earnings, and wages plus self-employment earnings. These incomes

are constructed at the tax unit/household level. They are constructed as follows. Wages are simply the sum of IRS wages for the tax unit and the sum of each member of that unit's total wages as reported to the CPS. Interest plus dividends is the sum of IRS taxable interest, non-taxable interest, and dividends.³ CPS interest plus dividends is simply interest plus dividends (the CPS does not make the taxable/non-taxable distinction). IRS total money income, by our construction, adds non-taxable interest back into "total income," and so is really IRS total money income plus non-taxable interest. CPS total money income is the sum of CPS components that most closely matches this IRS line item, given that we dropped data points based on comparability. The CPS total money income therefore is the sum of wages, interest, dividends, sole-proprietor earnings, farm income, alimony, rental income, retirement income, unemployment compensation, and other income. Because IRS total money income includes taxable refunds, credits, and offsets of state and local taxes, but the CPS has no measure of those items, that difference remains a potential wedge in the direct comparability of IRS total money income and CPS total money income. Administrative self-employment earnings are self-employment earnings subject to Medicare taxes. CPS self-employment earnings are the sum of sole-proprietor earnings, farm income, and rental income. Administrative wages plus self-employment is the sum of DER wages and self-employment earnings subject to Medicare taxes, and CPS wages plus self-employment earnings is the sum of CPS reported wages, sole-proprietor earnings, farm income, and rental income.

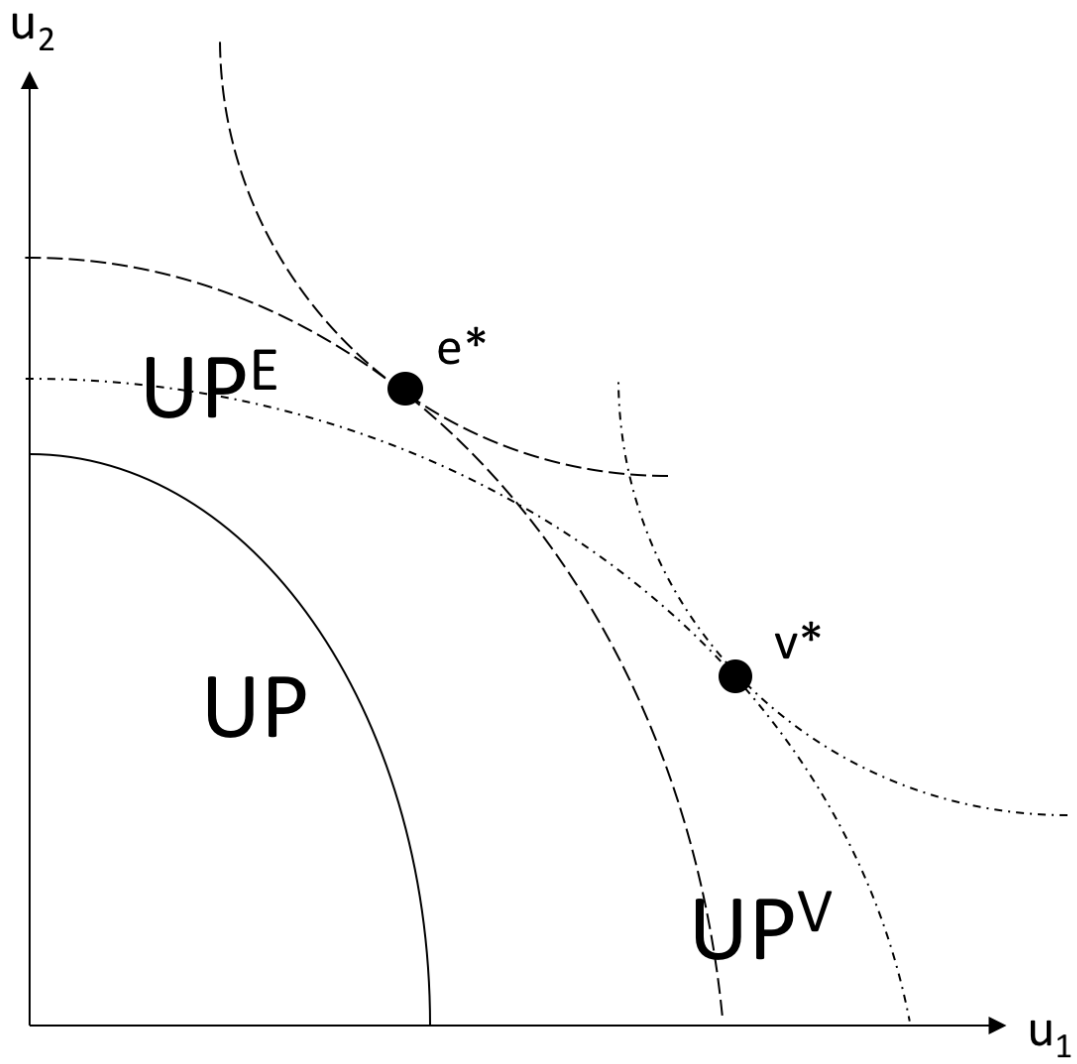
³The IRS does make the distinction between qualified and non-qualified dividends, but both are included in IRS "total income" and the distinction in tax treatment comes later on Form 1040.

Payoff Determination in Sabotage Games

In this appendix I justify the choices of payoffs employed in the games of sabotage. We can think of an arbitrary social welfare considering body of government as having a set of utility possibilities available to it, which can be thought as the union of utility possibilities sets corresponding to any specific public policy choices (so that the utility possibilities present when fulfilling any particular government role can correspond to exactly one policy action or a set of actions). A simple two-individual sample utility possibilities set is labelled “UP” and bounded by the solid curve in Figure A7.

The Empire has a distinct advantage over any arbitrary government body, in that it is large and powerful. This advantage allows it to achieve utility possibilities that an arbitrary government body could not, for example stretching a budget further by producing public goods at a lower cost due to economies of scale. This effectively stretches the utility possibilities set available to the Empire outward, as shown by the set labelled “ UP^E ” and bounded by the regular dashed curve in Figure A7. The Village also has a distinct advantage over an arbitrary government body in that it is close to the people who are the target of government action in these models. The Village knows its people, knows what will please them, knows local obstacles to policy implementation and can avoid potential pitfalls of which an outsider may be unaware. This local knowledge also allows it to achieve utility possibilities beyond that of the arbitrary government body, and is also shown as a stretching of the original utility possibilities set in Figure A7, labelled “ UP^V ” and featuring an alternating dashed curve. In addition, local pride in the efforts of the Village may also stretch UP^V outward.

FIGURE A7.
Maximization Problems of the Empire and the Village



As explained, the Empire and the Village each have preferences over policy outcomes that can be represented by social welfare functions. These are indicated by indifference curves in Figure A7. The Empire's sample indifference curve is shown in the regular dashed pattern and the Village's curve is shown in the alternating dashed pattern. As shown, each entity prefers the set of individual utilities on the frontier of their corresponding utility possibilities set that touches the indifference curve that touches but does not cut into the set. These preferred points are labelled e^* and v^* on Figure A7. These points are the preferred policies of the two players, and correspond to the payoffs $\pi_{E,a}^E$ and π_V^V , respectively. The payoffs $\pi_{E,a}^V$ and π_V^E are the values of the Village's and Empire's social welfare functions at the opposite player's optimum (these indifference curves are not shown).

To ease interpretation, we can think of these payoffs as loosely having three components. First, there is the effect of the player choosing a possible utility level as if it were an arbitrary government body. Second, the Empire confers a premium to any player receiving a payoff when the Empire is the implementer, which is the effect of the Empire's power, size, and ability to implement policies at low cost. Third, the Village confers a premium to any player receiving a payoff when the Village is the implementer, which can be thought of as the increased utility due to local knowledge and pride.

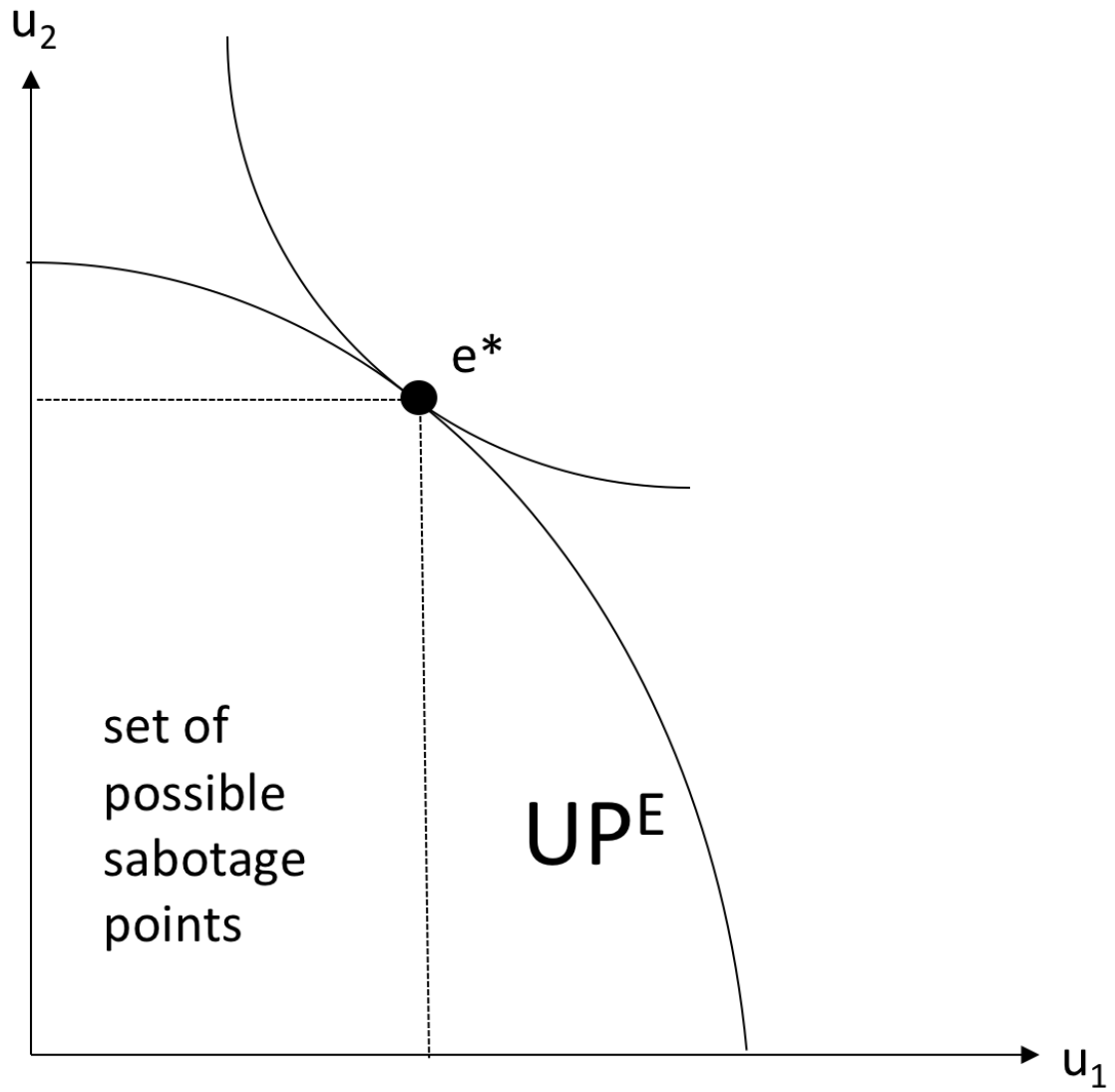
The final payoffs to be determined are the payoffs received when the Empire is the implementer but the Village sabotages the Empire's implementation. As the act of sabotage confers no immediate benefit (unlike other harmful actions like theft, that enrichen the thief) and weakens the saboteur in the short run, the beneficial impacts of sabotage on the saboteur must occur later, due to strategic

interaction. The payoffs under sabotage are thus lower for both the Empire and the Village, relative to the payoffs received at point e^* in Figure A7. The set of the possible utilities under sabotage are shown by the dashed rectangle in Figure A8 (which, though not shown, stretches into negative utilities). While the sabotage, by definition, results in a payoff lower than the utility received at e^* , and thus can be anywhere in that set, it will help interpretation to think of the act of sabotage as particularly damaging to both players, such that the sabotage payoffs are far lower than the non-sabotage payoffs.

Now, the Empire and Village need not have social welfare preferences that are different or even so different that it is always the case that the Village and Empire pick their own preferred policy implementation. For example, in the case of identical social welfare preferences, the premium conferred by the Empire's size may be greater than the premium conferred by the Village's localness, in which case both Empire and Village will prefer that the Empire implements, and vice versa. This simple result explains many historical results that are beyond the current scope of this paper.

In general, however, I assume that, for most government roles, the Empire and the Village have distinctly different views on proper implementation, so that the Empire prefers its own plan despite the Village's local advantage, and the Village prefers its own plan despite the Empire's size advantage. If we take each player's payoffs to be four under their preferred plan but two under the other player's preferred plan before incorporating their local or size premia, and give each player's premium a value of one, then players receive a payoff of five under their plan and three under their opponent's plan. As the sabotage payoffs are meant to be particularly bad, for each player, they shall be assigned a value of negative one.

FIGURE A8.
Effects of Village Sabotage on the Utility Possibilities Set



These payoffs are shown in Figure 5 in the main body of the paper and comprise the “standard payoffs” of this game.

REFERENCES CITED

- Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1(3-4):323–338.
- Arulampalam, W., Devereux, M. P., and Maffini, G. (2012). The direct incidence of corporate income tax on wages. *European Economic Review*, 56(6):1038–1054.
- Auerbach, A. J. (2006). Who bears the corporate tax? a review of what we know. *Tax Policy and the Economy*, 20:1–40.
- Ben-Porath, E. and Dekel, E. (1992). Signaling future actions and the potential for sacrifice. *Journal of Economic Theory*, 57(1):36–51.
- Bloomberg Tax (2014). Iowa governor signs tax reform & conformity bill. <https://www.bna.com/iowa-governor-signs-n57982093083/>. Accessed: 2018-05-18.
- Bollinger, C. R. (1998). Measurement error in the current population survey: A nonparametric look. *Journal of Labor Economics*, 16(3):576–594.
- Borjas, G. J. and Bronars, S. G. (1989). Consumer discrimination and self-employment. *Journal of Political Economy*, 97(3):581–605.
- Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement error in survey data. *Handbook of econometrics*, 5:3705–3843.
- Bound, J. and Krueger, A. B. (1991). The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? *Journal of Labor Economics*, 9(1):1–24.
- Brummet, Q., Flanagan-Doyle, D., Mitchell, J., Voorheis, J., Erhard, L., and McBride, B. (2018). Investigating the use of administrative records in the consumer expenditure survey. CARRA Working Paper 2018-01.
- Buffett, W. E. (1984). The superinvestors of graham-and-doddsville. *Hermes*, pages 4–15.
- Burkhauser, R. V., Feng, S., Jenkins, S. P., and Larrimore, J. (2012). Recent trends in top income shares in the usa: Reconciling estimates from march cps and irs tax return data. *Review of Economics and Statistics*, 94(2):371–388.
- Cameron, S. V. and Heckman, J. J. (2001). The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy*, 109(3):455–499.

- Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling. Technical report, National Bureau of Economic Research.
- Card, D. and Krueger, A. B. (1992). School quality and black-white relative earnings: A direct assessment. *The Quarterly Journal of Economics*, 107(1):151–200.
- Cavaliere, G. and Xu, F. (2014). Testing for unit roots in bounded time series. *Journal of Econometrics*, 178:259–272.
- Chay, K. Y. (1998). The impact of federal civil rights policy on black economic progress: Evidence from the equal employment opportunity act of 1972. *ILR Review*, 51(4):608–632.
- Chetty, R., Looney, A., and Kroft, K. (2009). Saliency and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–77.
- Council of State Governments (1994-2017). *The Book of the States*. Council of State Governments.
- Cragg, J. G., Harberger, A. C., and Mieszkowski, P. (1967). Empirical evidence on the incidence of the corporation income tax. *Journal of Political Economy*, 75(6):811–821.
- De, R. and Mehran, H. (2014). Evolution of s-corporation banks. <https://libertystreeteconomics.newyorkfed.org/2014/11/evolution-of-s-corporation-banks.html>. Accessed: 2019-04-18.
- Deaton, A. (1997). *The Analysis of Household Surveys: a Microeconomic Approach to Development Policy*. World Bank Publications.
- Desai, M. A., Foley, C. F., and Hines, J. R. (2007). Labor and capital shares of the corporate tax burden: International evidence. *Conference on Who Pays the Corporate Tax in an Open Economy*.
- Devereux, Michael P., B. L. and Redoano, M. (2007). Horizontal and vertical indirect tax competition: Theory and some evidence from the usa. *Journal of Public Economics*, 91(3-4):451–479.
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a):427–431.
- Dwenger, N., Rattenhuber, P., and Steiner, V. (2011). Sharing the burden? empirical evidence on corporate tax incidence. *German Economic Review*.

- Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1):34–105.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417.
- Fama, E. F., Fisher, L., Jensen, M., and Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review*, 10(1):1–21.
- Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.
- Feldstein, M. S. and Poterba, J. M. (1980). State and local taxes and the rate of return on nonfinancial corporate capital (revised as w0740).
- Felix, R. A. and Hines, J. R. (2009). Corporate taxes and union wages in the united states. Technical report, National Bureau of Economic Research.
- Giroud, X. and Rauh, J. (2015). State taxation and the reallocation of business activity: Evidence from establishment-level data. Technical report, National Bureau of Economic Research.
- Gordon, R. H. (1985). Taxation of corporate capital income: Tax revenues versus tax distortions. *The Quarterly Journal of Economics*, 100(1):1–27.
- Gravelle, J. (2017). Corporate tax reform: Issues for congress. Technical report, Congressional Research Committee.
- Harberger, A. C. (1962). The incidence of the corporation income tax. *Journal of Political Economy*, 70(3):215–240.
- Harberger, A. C. (2006). Taxation and income distribution: Myths and realities. *The Challenges of Tax Reform in a Global Economy*, pages 13–37.
- Hassett, K. A. and Mathur, A. (2010). Spatial tax competition and domestic wages.
- Heider, F. and Ljungqvist, A. (2015). As certain as debt and taxes: Estimating the tax sensitivity of leverage from state tax changes. *Journal of Financial Economics*, 118(3):684–712.
- Henry J. Kaiser Family Foundation (2019). Status of state action on the medicaid expansion decision.
<https://www.kff.org/health-reform/state-indicator/state-activity-around-expanding-medicaid-under-the-affordable-care-act/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D>. Accessed: 2019-04-12.

- Hoyt, W. H. and Jensen, R. A. (1996). Precommitment in a system of hierarchical governments. *Regional Science and Urban Economics*, 26(5):481–504.
- Internal Revenue Service (2016). Federal tax compliance research: Tax gap estimates for tax years 2008-2010. Technical Report 1415 (Rev. 5-2016), Internal Revenue Service Office of Research, Analysis, and Statistics, Washington, D.C.
- Jametti, M. and Brühlhart, M. (2004). Horizontal versus vertical tax competition: An empirical test. *Econometric Society 2004 North American Winter Meetings*.
- Juhn, C. (2003). Labor market dropouts and trends in the wages of black and white men. *ILR Review*, 56(4):643–662.
- Kane, T. J. and Rouse, C. E. (1995). Labor-market returns to two- and four-year college. *The American Economic Review*, 85(3):600–614.
- Keen, M. and Marchand, M. (1997). Fiscal competition and the pattern of public spending. *Journal of Public Economics*, 66(1):33–53.
- Kreps, D. M. and Wilson, R. (1982). Reputation and imperfect information. *Journal of Economic Theory*, 27(2):253–279.
- Krzyzaniak, M. and Musgrave, R. A. (1963). *The Shifting of the Corporation Income Tax*. Johns Hopkins Press.
- Lenter, D., Slemrod, J., and Shackelford, D. (2003). Public disclosure of corporate tax return information: Accounting, economics, and legal perspectives. *National Tax Journal*, pages 803–830.
- Li, G. (1985). *Exploring Data Tables, Trends, and Shapes*, chapter Robust regression. John Wiley & Sons, New York.
- Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, 20(4):587–615.
- Ljungqvist, A. and Smolyansky, M. (2014). To cut or not to cut? on the impact of corporate taxes on employment and income. Technical report, National Bureau of Economic Research.
- Ljungqvist, A., Zhang, L., and Zuo, L. (2017). Sharing risk with the government: How taxes affect corporate risk taking. *Journal of Accounting Research*.
- Lo, A. W. (2005). Reconciling efficient markets with behavioral finance: the adaptive markets hypothesis. *Journal of Investment Consulting*, 7(2):21–44.

- Malkiel, B. G. (1973). *A Random Walk Down Wall Street*. Norton.
- Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1):59–82.
- Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press.
- Mieszkowski, P. and Zodrow, G. R. (1989). Taxation and the tiebout model: the differential effects of head taxes, taxes on land rents, and property taxes. *Journal of Economic Literature*, 27(3):1098–1146.
- Milgrom, P. and Roberts, J. (1982). Predation, reputation, and entry deterrence. *Journal of Economic Theory*, 27(2):280–312.
- Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, 66(4):281–302.
- Morgenstern, O. (1963). *On the Accuracy of Economic Observations*. Princeton University Press.
- Niskanen, W. A. (1968). The peculiar economics of bureaucracy. *The American Economic Review*, 58(2):293–305.
- Niskanen, W. A. (1971). *Bureaucracy and Representative Government*. Transaction Publishers.
- Oates, W. (1972). *Fiscal Federalism*. Edward Elgar Publishing.
- Peoples, J. and Talley, W. K. (2001). Black-white earnings differentials: Privatization versus deregulation. *The American Economic Review*, 91(2):164–168.
- Piketty, T. and Saez, E. (2008). Income inequality in the united states, 1913-1998 (tables and figures updated to 2006). <http://www.econ.berkeley.edu/saez/TabFig2006.xls>.
- Reimers, C. W. (1983). Labor market discrimination against hispanic and black men. *The Review of Economics and Statistics*, pages 570–579.
- Romer, C. D. and Romer, D. H. (2010). The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks. *American Economic Review*, 100(3):763–801.
- Rosen, H. S. (1976). Taxes in a labor supply model with joint wage-hours determination. *Econometrica*, 44(3):485.

- Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2):41.
- Selten, R. (1978). The chain store paradox. *Theory and Decision*, 9(2):127–159.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442.
- Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *The Journal of Economic Perspectives*, 21(1):25–48.
- Slemrod, J. and Bakija, J. (2008). *Taxing Ourselves: a Citizen's Guide to the Debate over Taxes*. MIT Press.
- Street, J. O., Carroll, R. J., and Ruppert, D. (1988). A note on computing robust regression estimates via iteratively reweighted least squares. *The American Statistician*, 42(2):152–154.
- Suarez-Serrato, J. C. and Zidar, O. (2016). Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. *American Economic Review*, 106(9):2582–2624.
- Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5):416–424.
- Trejo, S. J. (1997). Why do mexican americans earn low wages? *Journal of Political Economy*, 105(6):1235–1268.
- U.S. Census Bureau (2006). Current population survey: Design and methodology. Technical Report 66, United States Bureau of the Census.
- U.S. Census Bureau (2015). Cps non-response. <http://www.census.gov/programs-surveys/cps/technical-documentation/methodology/non-response-rates.html>.
- U.S. Treasury Department (2007). Irs and the tax gap. Technical report, Treasury Inspector General for Tax Administration, Washington, D.C.
- Vega-Redondo, F. (2003). *Economics and the Theory of Games*. Cambridge University Press.
- Williams, A. (1966). The optimal provision of public goods in a system of local government. *Journal of Political Economy*, 74(1):18–33.
- Wilson, J. D. (1999). Theories of tax competition. *National Tax Journal*, pages 269–304.
- Wrede, M. (1996). Vertical and horizontal tax competition: Will uncoordinated leviathans end up on the wrong side of the laffer curve. *FinanzArchiv/Public Finance Analysis*, pages 461–479.