Three Essays in the Public Economics of Perception and Belief

by

Giacomo Brusco

A dissertation presented in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2020

Doctoral Committee:

Professor James R. Hines Jr., Chair
Associate Professor Ying Fan
Associate Professor Kyle Handley
Professor Joel Slemrod

Giacomo Brusco

gbrusco@umich.edu

ORCID iD: 0000-0003-4398-7266

A Giulia, per avermi sopportato in tutti questi anni.

Acknowledgements

When I entered my Ph.D. in economics at University of Michigan seven years ago, I could have never imagined the wealth of personal and professional experiences that staying here provided me. I am immensely grateful for each one of these experiences, and I believe I come out the other end of the tunnel not just richer in my mastery of economic theory and statistical techniques, but also more patient, more adamant to question what I do and do not know, and more driven to pursue further knowledge. None of this would have been possible without the help of some truly extraordinary people.

I was lucky to have access to fantastic scholars in my doctoral committee and beyond. James Hines was motivating and inspiring as a committee chair, guiding me to high aspirations in my upswings, and always there to prevent me from falling in my downswings. Joel Slemrod was selflessly supportive, with both his time and his resources, in shepherding me through the vicissitudes of the world of public finance research – along with countless generations of other public economists. Ying Fan was an invaluable source of insight and wit. Meetings with her were always action-packed; in many years, I have not once left her office without a clear idea of the path forward. Kyle Handley was instrumental in connecting me with the literature on policy uncertainty, which ended up at the core of my academic interests. I am truly thankful to him for his guidance and for the insights he provided. I am especially grateful to E. Han Kim, both for his guidance in finance research and for the financial support he provided through the E. Han Kim fellowship, which sustained me during the last year of my Ph.D.

I was also blessed with stellar peers. The group occupying the "Lorch 109" office will forever be in my thoughts. Benjamin Glass was a faithful friend and a great co-author, forever brimming with ideas and suggestions. William Boning, Ou Nie, Luis Alejos, Luis Baldomero, and Connor Cole were great office mates, with whom I shared many victories and defeats. Chiara Ferrero has been a close friend, and has provided a solid support network for both me and my wife to thrive in Michigan. Tejaswi Velayudhan was an amazingly supportive study-group partner, and has been a great colleague and co-author before and after graduating last year. Yeliz Kaçamak and Ellen Stuart have recently also become close collaborators, and I am truly thankful to them for their ingenuity and hard work.

No list of thanks would be complete without mention of my family. My parents, Tiziana and Sandro, have been incredibly supportive and understanding since I can remember; particularly in recent years, they were a near endless source not just of help and guidance, but also of inspiration. My little sister Bianca has grown to be one of the smartest and most generous people I know, and I am astonished at how much I learned from her in these years.

I would also like to thank a group of friends – some of whom, like Stefano Bonvini, date back all the way back to my early childhood – who have felt like a second family in recent years. The "Gruppo del Pacco", as our WhatsApp chat is called, now spans fifteen time zones, but feels as close-knit as when we were a clumsy gaggle of rebellious teenagers. Being able to rely on such old friends, who feel far removed from work life but close to the heart, has been life-saving more than once.

Last but not least, I would like to thank my wife Giulia. She has been a true, life-long friend, and I count myself unbelievably lucky to have her by my side. Spending the last eleven years together has been an amazing ride, and none of my days – from the busiest night spent writing a paper for an impending deadline, to the laziest Sunday afternoon – would be the same without her. I cannot wait to see where our next adventure will take us.

Preface

This dissertation touches upon several ways in which economic behavior in response to tax policy is modified by what agents believe and perceive about the world around them.

The first chapter analyzes the stock market responses to a change in expectations about future tax rates. Making use of betting markets to measure expectations regarding a major U.S. tax reform, it finds that the most profitable firms, and those in more concentrated markets, gained a large majority of the benefits of the tax reform, which was widely expected to reduce the tax burden of all corporations. This result assumes particular importance in light of recent results in the economics literature, which has found that market power has risen substantially in the U.S. over the past three decades, and that this increase has been driven by a few firms at the top of the mark-up distribution. My results confirm this trend in investor expectations, looking forward. The relationship between rents and corporate tax burden allows me to estimate the distribution of rents across the U.S. sector, showing that indeed rents are concentrated among a few firms, which tend to be intensive in their use of intangible capital.

The second chapter, co-authored with Benjamin Glass and published in the *Journal of Public Economic Theory*, analyzes how agent welfare is affected by the misperception of tax rates, generalizing some existing results in the behavioral public finance literature. We find that in the presence of heterogeneous biases, measuring the welfare effects of a sales tax is impossible even with full knowledge of aggregate demand, as a function of both pre-tax prices and the tax rate. While consumer under-reaction to sales tax tends to harm consumer surplus, as agents are unable to defend themselves against higher after-tax prices, it also tends to mitigate tax burden in excess of revenue raised, as unresponsive agents will generate more revenue for a given tax rate. Nonetheless, heterogeneity introduces a channel that exacerbates the excess burden, as different perceived prices introduce an issue of allocative inefficiency. Our work shows, in a very general setting, that by maximizing or minimizing this allocative inefficiency one can obtain bounds for excess burden, computable with aggregate or average parameters of demand.

The third chapter, also co-authored with Benjamin Glass, shows that even under taxation schemes designed to leave marginal incentives untouched, policy uncertainty can distort the ex-ante decisions of a firm. In particular, we show that if a firm's marginal tax rate is correlated with the success of its ventures, this might discourage or encourage input use, such as investment. More specifically, investment will be discouraged if the firm's marginal tax rate is higher in those states of the world in which the investment's marginal revenue product is high; viceversa, if the marginal revenue product and the marginal tax rate are negatively correlated, investment will be encouraged. We further investigate this channel in an empirical model of capital asset pricing, finding significantly different patters across different industries.

Each chapter is designed to be self-contained, but all inquire the ways in which perception and belief shape the effects of economic policy. The first chapter measures how changes in expected tax rates heterogeneously impacted the stock valuation of firms, and uses this to infer conclusions regarding the distribution of rents throughout the corporate sector. The second chapter studies how the misperception of tax rates influences consumer welfare, and how one might measure excess burden in the presence of these biases. Last but not least, the third chapter is concerned with a "second moment" of agents' belief; it shows how the covariance between two sources of uncertainty, government policy and underlying business condition determining a firm's marginal revenue product, might distort firm decisions.

Table of Contents

List of Figures

List of Tables

List of Appendices

Abstract

My dissertation studies three cases in which the perception of current policy or beliefs about future policy shaped economic decisions. Issues of attention and beliefs about the future can have tremendous importance both because agent actions reveal what they believe about the future, such as in the first and third chapters, or because misperception change how agents are affected by policy, such as in the second chapter.

The first chapter measures the effect of the 2017 Tax Cuts and Jobs Act on share prices of publicly traded firms, finding that the most profitable firms, and those in concentrated industries, benefited the most. The tax bill significantly reduced corporate tax rates, thereby increasing share prices, particularly at the top. Among firms with the highest profit rates, more than 80 percent rose in value on news of the tax reform, whereas among less profitable firms, fewer than 60 percent appreciated. By every measure, stock market gains coincident with the tax bill were concentrated among firms with greater market power. This pattern is consistent with economic rents being important components of the values of large U.S. corporations.

The second chapter of my dissertation looks at how to measure welfare in a world where tax policy might be misperceived by agents. Recent developments in behavioral public economics have shown that heterogeneous biases prevent point identification of deadweight loss. The second chapter replicates this result for an arbitrary (closed) consumption set, whereas previous results on heterogeneous attention focused on binary choice. It finds that one can bound the efficiency costs of taxation based on aggregate features of demand. When individuals have linear

demand functions, the bounds for deadweight loss are easy to calculate from linear regressions.

While the first chapter of my dissertation looks at the effects of expected policy, the third and last chapter looks at higher-order moments. Forecasts of the consequences of tax changes usually assume that economic actors expect these changes are permanent, despite the inevitable political uncertainty that could lead to future reversals or further changes. This reasoning extends to when a firm's tax burden is correlated to the success of its ventures. The third chapter shows how a firm's belief about how government policy is correlated with the input's marginal product distorts its risk profile, leading it to change its input decision. Generally speaking, input use will be discouraged if the firm faces high taxes precisely when the input is more productive. The last chapter shows that in a world of policy uncertainty this holds under an arbitrary tax system, and in particular it holds even if inputs can be perfectly deducted from the firm's taxable income. Whenever the covariance between policy and payoff is zero, the model replicates the classical result that the deductibility of input expenses leaves the decision undistorted. The third chapter uses this theoretical relationship in an empirical model of asset pricing to infer what investors believe about how future government policy correlates with their risky investments in different firms in the stock market.

Each chapter's analysis can be read on its own, but the unifying theme is that in each case issues of perception and belief were central in either identifying current beliefs about the world, or in understanding the real impact of economic policy on agent welfare.

Chapter 1. Tax Reform and the Valuation of Superstar Firms

1.1. Introduction

The last few decades have seen the rise of innovative firms that have managed to quickly gain large market shares in their respective sectors. Some of these companies have joined the ranks of the firms with the largest market capitalizations, sparking public debate regarding their ability to control such large portions of their respective markets. Recent academic literature corroborates some of these concerns. De Loecker and Eeckhout [2017] and De Loecker et al. [2018] have documented a sharp increase in mark-ups over marginal cost since 1980, arguing that this has been driven by a rise in the market power of a few firms at the top of the mark-up distribution. Autor et al. [2017] and Dorn et al. [2017] use a similar methodology to argue that this increase in mark-ups is tied to the secular decline in labor's share of GDP.

this chapter studies how much market power matters for the valuation of the U.S. corporate sector, by investigating how news regarding a corporate tax cut in the U.S. affected the excess returns of U.S. corporations. The effect of a tax cut on the excess returns of a company depends on who bears the burden of corporate taxes. If shareholders bear all of the burden, firm values will simply increase linearly with the cut in tax rates. If, on the other hand, the burden can be entirely shifted away, such as to workers in the form of lower wages or consumers in the form of higher prices, then news of the tax cut will not affect firm valuations. Firms with market power will bear more of the corporate tax burden, because they bear the entirety of the burden of taxes that fall on pure economic profits, or rents.

The empirical analysis consists of two main parts, both of which look at the effects the Tax Cuts and Jobs Act (TCJA). Signed into law on December 22, 2017, the reform reduced corporate income tax rates in the U.S. from a top rate of 35% to a flat rate of 21%. In order to identify news of the tax bill, the paper relies on relevant dates for the passing of the TCJA, identified in other literature through Google trends data and common knowledge of the history of the bill. Further, data from online bets on corporate tax reform in the U.S. allow me to quantify the news on each date and develop clear interpretations for my estimates.

The first part of the analysis exploits variation across firms and, controlling for other characteristics that might have influenced heterogeneity in the effect of the tax bill, such as exposure to provisions on foreign operations, it shows that firms with higher accounting profit margins and bigger market shares tended to earn higher returns upon news of the TCJA. The second part exploits variation within each company across time, and looks at the resulting distribution of individual effects of news of the tax bill in its entirety.

The first exercise shows strong relations between excess returns due to news of the tax cut and measures of profitability and market power. The results show that firms with high shares in more concentrated industries and firms with bigger profit margins gained significantly more from the reduction in corporate tax rates. If market shares and profit margins reflected purely differences in productivity, rather than differences in the potential entry faced by firms, then we would expect to see no relation between the two. Instead, the results suggest that at least some variation derives from real rents that cannot be competed away by potential entry, either in the short run or in the long run.

These results are important for two reasons. First, they are important for policy reasons, as they reveal that differences in market power may be an important source of variation in the burden of corporate taxation faced by the owners of different firms, and that they may be taxed without much distortion. Second, the results put emphasis on the notion put forth by De Loecker et al. [2018] that market power might play a significant role in the market valuation of some companies.

Quantifying the news on each of the dates considered as relevant for news of the TCJA reveals highly heterogeneous effects across different dates, which implies news on the dates of interest represent more than simple changes in the probability of a corporate tax cut. Given this argument the last date considered, in which the bill received final congressional ratification for signature by the president, represents the cleanest identification of how an increase in the probability of a tax cut affected firm values. According to this interpretation, the results suggest a one standard deviation change in a firm's accounting profit margin is associated with an increase in returns of 29.9 percentage points, and that going from perfect competition to a position of monopoly increases the excess return a company obtains due to the tax rate change by 87.7 percentage points.

The second part of the analysis shows that the benefits of the tax bill were highly concentrated among a few firms at the top. This is especially interesting in light of the first set of results, because the firm-level effect of the bill offers an alternative measure of a firm's market power. The distribution of gains from the bill is concentrated around zero but exhibits a long right tail. Quantifying the results with betting data reveals a strong concentration of gains from the TCJA. The top 1% of beneficiaries from the tax bill represented 22% of the original market value of all publicly traded firms but earned 50% of the total gains from the TCJA; the top 5% represented 45% of the original market value but earned 88% of the gains; the top 10% represented 53% of the original value and earned more than the totality of the gains, which were negative for about a quarter of firms in the sample.

While these results should be interpreted with caution, they are important because they provide a forward-looking confirmation of the historical trends detected by Autor et al. [2017] and De Loecker and Eeckhout [2017]. It is also reassuring for this strain of literature, which often relies on structural assumptions for identification, that the ranking of firms in my sample largely matches our prior: Google, Apple, Walmart, Disney, and Microsoft all feature as some of the biggest beneficiaries of the bill. Further, my results provide a picture of the distribution of the heterogeneous burden of corporate taxation across different firms.

The rest of the chapter proceeds as follows. Section 1.2 reviews the literature on corporate tax incidence and gives some background on the present state of the literature, contextualizing the contribution of this chapter. Section 3.3 lays out a simple framework to think about the excess returns earned by a firm upon a cut in the corporate tax rate. Section 1.4 describes the data and gives some institutional background. Section 1.5 describes the methodology and the results of the empirical analysis. Section 3.5 concludes.

1.2. Background

The modern study of corporate tax incidence began with the Harberger [1962] model, in which the incidence of the corporate tax depends on the intersectoral reallocation of capital. Dividing the economy into a corporate and a non-corporate sector, he points out that higher corporate tax rates will increase the cost of corporate capital. If capital is free to move between the two sectors, this will reduce the return to capital economy-wide, and so all owners of capital, not just corporate capital, bear the burden of the tax. Depending on the relative intensity in each sector, other factors may bear some or all of the burden, too, as corporate capital is moved to the non-corporate sector. Subsequent work such as Summers [1981] has noted that if there are frictions that prevent capital from reallocating immediately between sectors, then incidence might be higher for corporate shareholders. On the other hand, Gordon and Hines [2002] noted that if capital is perfectly mobile across borders, corporate tax changes in one country cannot affect the after-tax return to capital, and thus the entirety of the burden falls on fixed factors of production, most notably labor and land.

Issues of imperfect competition also complicate the incidence of corporate taxes, and are of primary importance to this chapter. Any rate of return earned by corporate capital on top of the risk-adjusted rate of return is a rent, and taxes on corporate rents are generally entirely borne by shareholders.[1] The simplest example of this is a tax on a monopolist's profits: because the monopolist maximizes pre-tax profits, regardless of the tax rate, its behavior will not be distorted and shareholders will bear the entire burden of the tax.

One notable exception from the simple monopoly reasoning is that corporate rents can arise as returns on intangible capital held by firms. In this case, the corporate tax can be seen as a tax on entrepreneurial effort, and its incidence will depend on how it affects the process of creation of new ideas. As Auerbach [2006] vividly puts in his review of the literature on corporate tax incidence, "the garages of Silicon Valley might have been used to store cars if the corporate tax rate had been higher." To the degree that the creation of new ideas is not influenced by the corporate tax, a tax on corporate rents will still fall mainly on corporate shareholders. While gathering conclusive evidence on this matter is far from trivial, Akcigit et al. [2018] find evidence suggesting that taxes do matter for innovation, or at least for its distribution across U.S. states. My results suggest that big tech firms did not face a sharp increase in competition due to the reduction of the corporate tax rate – at least not in the short run.

A second issue is that the corporate tax might affect the equilibrium of a dynamic game of imperfect competition. Davidson and Martin [1985] note that a corporate tax increase might affect companies' ability to punish each other in the event of a deviation from collusive behavior, which would also make it distortionary. Furthermore, if the tax base does not coincide with pure rents, imperfect competition might worsen the welfare effects of taxation and can lead to tax overshifting, whereby a tax induces prices to increase more than one-for-one. Fullerton and Metcalf [2002] review the literature on tax incidence and derive the conditions

[1] See, for instance, the discussion in Auerbach [2006].

under which over-shifting would be induced by an excise or an ad-valorem tax.

Empirical work on corporate tax incidence tends to conclude that at least some of the corporate tax burden is shared by factors of production other than corporate capital. Much of this work is focused on assessing the impact of corporate tax on wages. Analyzing union wage premiums, Felix and Hines [2009] find that workers in unionized firms capture slightly over half of the benefits of low tax rates. Arulampalam et al. [2012]'s analysis of wages and corporate taxes in Europe agrees with this finding, suggesting that about half of the burden of corporate taxes falls on labor. Looking at cross-industry differences in corporate tax rates in the U.S., Liu and Altshuler [2013] find that the elasticity of wages with respect to the corporate marginal effective tax rate is higher in more concentrated industries, a fact seemingly at odds with the findings of this chapter, which suggest that greater concentration is tied to larger excess returns due to the tax cut. But as noted by Auerbach [2006], situations of imperfect competition, where market outcomes are already distorted by the desire of companies to artificially limit output, might exacerbate the distortions of the corporate tax. More recent studies of corporate tax incidence have focused on sub-national variation in corporate tax rates. Suárez Serrato and Zidar [2016] use variation in corporate tax rates across U.S. states to estimate a model with imperfect labor and firm mobility. They find that firm owners bear about 40% of corporate taxes, workers bear 30-35%, and landowners 25-30%. Fuest et al. [2018] use variation across German municipalities, finding that workers bear roughly one half of the corporate tax burden.

The most recent U.S. tax reform has spurred a new wave of empirical studies on the effects of corporate tax policy. While arguably the most salient feature of the TCJA for corporations was a reduction in corporate tax rates, it contained a variety of provisions which are reflected in the breadth of the studies on the topic. Wagner et al. [2018b], for instance, document that the stock prices of internationally-oriented firms suffered from news of the TCJA, and that the aggregate market responded positively to lower expected taxes. Blanchard et al.

[2018] find that increases in expected dividends – in which the corporate tax cut played a significant role – were largely responsible for the increase in stock value over the year following the election of President Trump. Gaertner et al. [forthcoming] study the effects on foreign companies, finding that Chinese firms, especially in steel manufacturing, experienced large negative returns, while the rest of the world experienced positive returns. Wagner et al. [2018a] analyze the effects of Trump's election on the U.S. stock market, finding that expected tax rates greatly impact firm value. Hanlon et al. [2018] look at the actions and *statements* about actions of companies in the aftermath of the TCJA, finding, among other things, that share buybacks generally increased in the aftermath of the TCJA, but that this increase was concentrated in a small number of firms.

this chapter stands out from existing literature in its focus on the relation between the excess return earned due to news of the TCJA and traditional measures of profitability and market power. While in much of the paper I will be speaking about the effects of a tax cut, my empirical analysis measures the effects of *news* regarding a possible future tax cut. The late 1980's and the 1990's saw a spur of interest in questions regarding belief about future policy. Auerbach and Kotlikoff [1987], Auerbach and Hines [1988], Cutler [1988], Poterba [1989], Rodrik [1991], and Slemrod and Greimel [1999] are all concerned with how changes in expectations regarding future tax policy affect current decision making, both by firms and individuals, both in financial markets and in the real economy. Cutler [1988]'s analysis of the 1986 Tax Reform Act's (TRA86) effect on stock market prices provides an analysis that parallels this chapter. Among other things, Cutler points out the effect of tax reform on firm value is ambiguous, a reality that in his case was further aggravated by the fact that TRA86 drew a distinction between old and new capital via the Investment Tax Credit. this chapter builds on his framework by using techniques more in line with the modern literature in empirical finance.

A related literature has studied the effects of policy beliefs more generally and in a variety of fields of economics. Friedman [2009], who used news on regulations of drug coverage to estimate the incidence of Medicare Part D, is a methodological predecessor to this chapter: on top of taking an out-right event-study approach, I use a more continuous approach that allows me to do something closer in spirit to a difference-in-difference estimation. Some of this literature has made use of betting data to make sense of magnitudes, such as Meng [2017], who has used prediction market data regarding the passing of an anti-pollution cap-and-trade bill to estimate the marginal abatement cost of proposed policy, or Graziano et al. [2018], who studied the consequences of expectations about the U.K.'s exit from the European Union.

The great potential applications for prediction market data, which this chapter makes use of, have been advocated for some time in the literature, starting with papers like Arrow et al. [2008]. Some early applications include Slemrod and Greimel [1999], who studied how the probability of Steve Forbes winning the presidency influenced the returns for municipal bonds, Wolfers [2006], who studied point-shaving in NCAA basketball, or Snowberg et al. [2011], who advocated the use of betting data in event studies. While some authors, such as Manski [2006], have cast some doubt on the use of these data, arguing that there is no theoretical guarantee that betting odds will reflect average or even median beliefs, Wolfers and Zitzewitz [2006] offer some theoretical justification as well as some empirical evidence in support of interpreting betting prices as average beliefs.

1.3. A Simple Model of Corporate Tax Incidence

This section lays out a simple theoretical model of excess returns, and relates it to the discussion of corporate rents. It then connects the predictions of the theoretical framework to the methodology used in the empirical analysis. As discussed in section 1.2, the effect of a corporate tax cut on the excess returns of a company's equity depends crucially on the degree to which shareholders bear the burden of the tax. The analysis assumes that firms are entirely financed by equity,[2] and considers what happens to the value of this equity starting at a certain corporate tax rate, $\tau \in (0, 1)$.

Firm i's discounted sum of future after-tax profits, $\bar{V}_i = \bar{V}_i(\tau)$, is determined by two distinct components, $V_{i1} = V_{i1}(\tau)$ and $V_{i2} = V_{i2}(\tau)$:

$$\bar{V}_i = V_{i1} + V_{i2}.$$

V_{i1} represents profits from operations where tax burden can be shifted to other entities, such as workers, consumers, or owners of non-corporate capital, while V_{i2} represents profits from operations whose pre-tax value is not influenced by variation in τ, as would be true for pure economic profits. Letting $\Pi_{i1}(\tau)$ and Π_{i2} represent the present discounted value of quasi-rents from each kind of operation, we will have:

$$\begin{aligned} V_{i1}(\tau) &= (1-\tau)\Pi_{i1}(\tau) \\ V_{i2}(\tau) &= (1-\tau)\Pi_{i2}. \end{aligned} \tag{1.1}$$

Appendix A.1 provides an example where firm value can be broken down in two parts with the same properties as in equation 1.1. When the entry of new competitors is impossible, the full burden of the corporate tax falls on firm owners; when entry is costly but possible, firms will bear less than the full burden.

Consider a marginal increase in the tax rate. Since the tax burden imposed on $V_{i2}(\tau)$ falls entirely on equity holders, a marginal tax increase will be entirely reflected in value: $dV_{i2}/d\tau = -\Pi_{i2}$. The same does not hold for V_{i1}, which will change only to the degree that any of the tax burden imposed on this part of

[2] While this assumption significantly simplifies exposition, note that little would be different if firms do not change their ratio of equity to debt financing, which will be true for a marginal tax change if firms were choosing their financing ratios optimally before the change. Although certainly the change in corporate tax rate induced by the TCJA was not marginal, this might alleviate concerns that results are heavily influenced by financing considerations.

profits actually falls on the firm:

$$\frac{dV_{i1}}{d\tau} = -\Pi_{i1}(\tau) + (1-\tau)\frac{d\Pi_{i1}(\tau)}{d\tau}.$$

In the case of a tax cut, $d\Pi_{i1}(\tau)/d\tau$ will be determined by the speed with which new firms can enter the market and compete with incumbents. On one extreme, if entry takes an infinitely long time, then the company's pre-tax profits decrease infinitely slowly, and so

$$\frac{d\Pi_{i1}(\tau)}{d\tau} = 0.$$

On the other extreme, if entry is immediate and the increase in after-tax profits is quickly competed away, then $dV_{i1}/d\tau = 0$, or

$$\frac{d\Pi_{i1}(\tau)}{d\tau} = \frac{\Pi_{i1}(\tau)}{1-\tau}.$$

Assuming that $d\Pi_{i1}(\tau)/d\tau$ can only vary between these two extremes allows us to express it as a convex combination of the two. For some $\alpha_i \in [0,1]$:

$$\begin{aligned}\frac{dV_{i1}(\tau)}{d\tau} &= -\Pi_{i1}(\tau) + (1-\tau)\alpha_i\frac{\Pi_{i1}(\tau)}{1-\tau} \\ &= -(1-\alpha_i)\Pi_{i1}(\tau).\end{aligned}$$

Therefore, one can think of α_i as the fraction of the burden on V_{i1} that is not borne by firm i.

Given these two pieces, the excess returns that the firm earns upon a marginal cut in the corporate tax rate will be:

$$\begin{aligned}r^e \equiv \frac{-d\bar{V}_i/d\tau}{\bar{V}_i} &= \frac{(1-\alpha_i)\Pi_{i1} + \Pi_{i2}}{V_{i1}+V_{i2}} \\ &= \frac{1}{1-\tau}\left(1 - \frac{\alpha_i}{1+\frac{\Pi_{i2}}{\Pi_{i1}}}\right).\end{aligned} \tag{1.2}$$

This expression reveals two important implications. First, the lower α_i the more tax burden falls on shareholders, and so excess returns on the firm's equity upon news of the tax cut are decreasing in α_i. Second, the higher the ratio of rents to other profits Π_{i2}/Π_{i1} the more important rents are in determining firm value, so higher values of Π_{i2}/Π_{i1} will result in a bigger excess returns upon a tax cut.

The empirical work that follows studies how excess returns due to news about the TCJA correlate with several measures of a firm's profitability and market power. More specifically, the analysis focuses on (i) market capitalization, which is a measure of firm size, (ii) the Herfindahl-Hirschman index, computed as the sum of squared shares in an industry and traditionally used as a measure of market power, interacted with own market share and (iii) the Lerner index, a firm's accounting profit margin, computed as the ratio of operating profits minus depreciation over revenues.

Observing a positive relation between one of these measures and excess returns could mean 1) that those firms with higher measures of market power earn high rents, meaning that they have a high V_{i2}/V_{i1}, or 2) that they tend to have a low α_i, which could be due to high adjustment costs of transferring capital in and out of the firm or frictions that prevent the entry of new firms or the exit of incumbents. In either case, the firm's profits are more insulated from new competition upon a reduction in the tax rate. In the limit as the time for competition to enter goes to infinity, the two interpretations coincide.

1.4. Data

Some data for this study come from a merge between the Center for Research in Security Prices (CRSP) and Compustat, including data on the daily holding returns of various stocks in the New York Stock Exchange (NYSE) and the companies to which they are tied. The study focuses on the period starting after the 2016 presidential election, on November 9, 2016, and ending on the day in which the Tax Cuts and Jobs Act was signed into law, on December 22, 2017, using both firm accounting data on items such as foreign profits or revenue and financial data such as holding returns and market value. Summary statistics for the relevant variables can be found in the Table A.2 in the Appendix.

Additional data include information on news regarding legislative prospects of the tax bill. this chapter uses six dates identified by Gaertner et al. [forthcoming] as particularly relevant for the development of tax reform in 2017.[3] They document spikes in Google searches regarding "tax reform", depicted in Figure 1.1, on these dates, and argue that this is due to the fact that these were surprising events at the time, which attracted the public's attention on the tax bill. In their paper, they also use these dates to perform event studies on the stock returns of foreign companies.

In addition to these dates, the paper uses data from a bet undertaken on the web platform PredictIt.org, which organizes bets among users around the world. Betting on PredictIt is entirely user-based; any registered individual can post a "Buy" or a "Sell" contract for each possible outcome, which in the case of our bets of interest is always binary ("Yes" or "No"). Contracts pay out nothing to losing users, and \$1 to winning users, and the price is simply the share of that dollar paid in by each of two users taking part in the bet. No user can wager more than a total of \$850 on a single bet, and PredictIt makes money by charging a 10% fee for all winnings in excess of money invested, plus a 5% withdrawal fee. As a result, users will want to buy "Yes" contracts that have a price (adjusted for fees) lower than their subjective probability of a "Yes", and sell contracts with a price higher.

As Graziano et al. [2018] point out, one need not interpret these implied probabilities as the true, homogeneous belief of all agents in the economy, or even as the true average belief in the economy, as long as one is willing to suppose that the price of these contracts is strongly correlated with individual beliefs, and as

[3]These dates are: September 27, 2017 (United Framework for Tax Reform unveiled - Member retreat), November 2, 2017 (TCJA introduced in the House), November 16, 2017 (House passes TCJA), December 2, 2017 (Senate passes TCJA), December 15, 2017 (Bill reported by the joint conference committee), December 20, 2017 (Final version agreed to by the Senate).

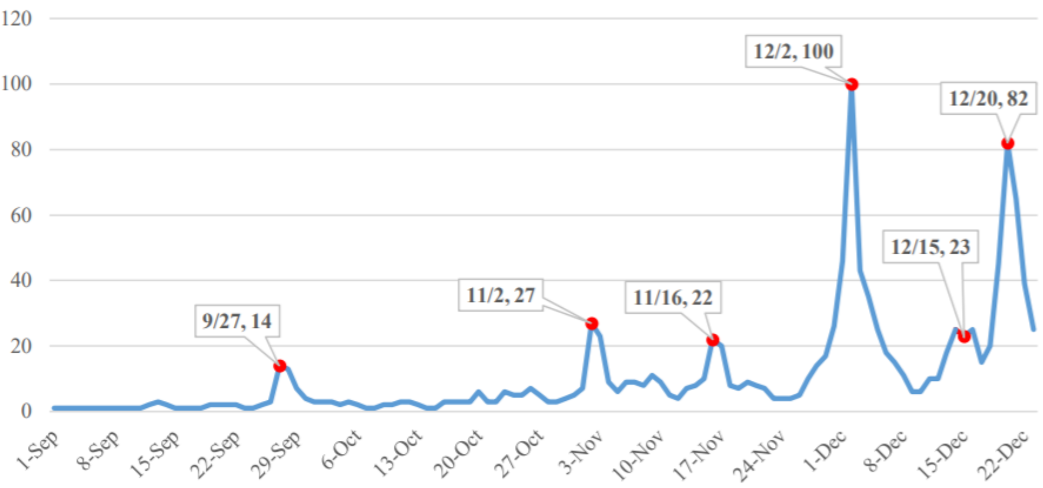


Figure (1.1) Google trends data for searches regarding "tax reform". Source: Gaertner et al. [forthcoming].

long as agents do not systematically change belief in direction opposite to that of the change in odds.[4] this chapter uses changes in betting contract prices to proxy how financial markets react to shocks to policy expectations.

Other literature on the TCJA, such as in Blanchard et al. [2018], has used a PredictIt betting series that started the day after the 2016 presidential election, November 9, 2017, and asked whether there would be a cut in the corporate tax rate by December 31, 2017. Gaertner et al. [forthcoming] point out that the biggest in this series spikes do not seem to match up with either the Google trends data they use nor with conventional wisdom about the developments of the tax bill. This might be driven by the fact that, except for the last few months of the bet, this specific betting market was not very thick, and certainly not as thick as for other bets on the same website,[5] so a few outliers could disproportionately move the price on any given day. Figure 1.2 shows the contract price series overlaid by the salient events described in Figure 1.1. The series declines until late 2017 when legislative framework is unveiled and the bill begins to move through Congress. Not all events match up with the changes as one would expect, and that the largest spike in the price occurs on August 9, 2017, which does not seem to match up with

[4]This could technically happen, for instance, if agents with radically different priors interpreted the same signals in systematically opposite ways.

[5]E.g., betting odds on the election of Donald Trump are more widely used and acknowledged as a good proxy for the public's beliefs

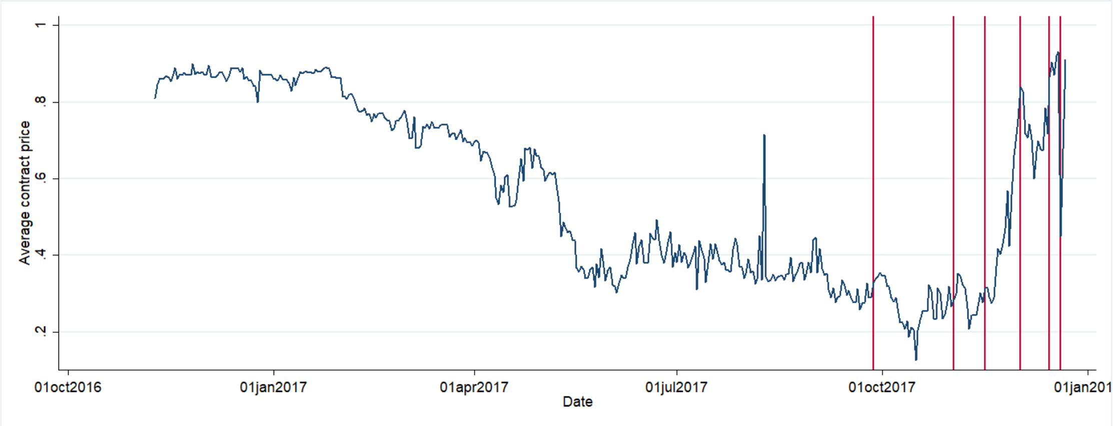


Figure (1.2) Average price for a "Buy Yes" contract in the bet "Will there be a corporate tax cut by December 31, 2017?". Each red line represents one of the events identified by Gaertner et al. [forthcoming] using Google Trends data.

any event of note. A simple explanation for this random spike is that a) limits on the amount that each user can wager on a single bet mean that some mispricing might not get arbitraged away, and b) this is particularly true on days when very few people are betting; it is not surprising to find out that on August 9, 2017, only two users were exchanging contracts on this bet.

Another issue with this particular series of futures prices is that it asked users to bet on whether there would be a cut in the top corporate tax rate *by December 31, 2017*, not necessarily whether one was going to happen imminently or eventually. This can be an issue that is strongly reflected in the betting odds on some days at the very end of the series, as right before passing the law there were some tensions between President Trump and the Republican-controlled Congress regarding funding for immigration control, and for a moment it seemed possible that the signing of the tax bill would have to be postponed to the new year. On December 20, 2017, for example, which is the last date identified by Gaertner et al. [forthcoming] using Google trends data, and only two days before the final signing of the bill into law, the probability that the bill is not signed into law, according to this bet, jumps *down* by almost 48.4 percentage points, which clearly reflects the possibility that there would be a corporate tax cut by year-end 2017, not the possibility that there would be a tax cut at all.

I analyze a new bet started by PredictIt on October 23, 2017, which asked whether there would be a corporate tax cut by March 31, 2018. One drawback of using these data as opposed to the older bet is that it has lower total volume, as the longer-running bet was likely more salient to users: the average number of contracts exchanged in the older bet on any day is 2245.7, but only 893.3 in the newer one; the average number of users exchanging contracts is 52 in the older one, but only 21 in the newer one. Another issue is that because the bet starts on October 23, this leaves out the first of the dates identified by Gaertner et al. [forthcoming], September 27, 2017. However, the new bet is not plagued by timing issues that would be irrelevant to this analysis, and the betting volume is more spread out over the lifetime of the bet (see appendix A.2). This could be due to the fact that this bet was running over a period in which tax legislation was fairly uniformly relevant in the news. Furthermore, the price of the futures contract lines up better with the dates used by Gaertner et al. [forthcoming], as we can see in Figure 1.3. The empirical analysis uses these betting odds as its main point of reference.

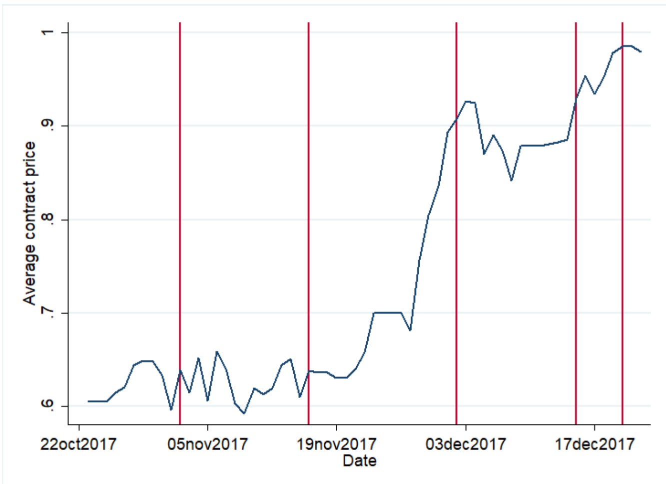


Figure (1.3) Average price for a "Buy Yes" contract in the bet "Will there be a corporate tax cut by March 31, 2018?". Each red line represents one of the events identified by Gaertner et al. [forthcoming] using Google Trends data.

1.4.1. Institutional Background

While tax reform had long been a campaign issue for Donald Trump, the first year of his presidency was spent mostly focusing on health reform. When these efforts stalled in the summer, Congressional Republicans decided to make tax reform their flagship issue for the coming fiscal year,[6] by tying it to the budget reconciliation process under H.Con.Res71, Title II. After that, as Wagner et al. [2018b] point out, the process moved swiftly, making the setting ideal for an event study.

The final version of the bill contained numerous provisions regarding the taxation of both individuals and firms. In particular, it contained a cut in the corporate tax rate, which went from a top rate of 35% to a flat rate of 21%. This was a slightly higher final rate than previous proposals, which were at 15% according to Trump's plan released in April 2017, and at 20% according to the House Republican plan of June 2016. The TCJA also allowed for immediate expensing of capital expenditures.

The TCJA also contained major changes to the taxation of multinational entities, shifting the U.S. from a worldwide to a territorial tax system, by allowing businesses to deduct completely dividends received by 10%-owned foreign corporations. However, it also introduced measures to curtail shifting profits to low-tax jurisdictions with minimum tax schemes such as GILTI (Global Intangible Low-Taxed Income) on intangible assets and BEAT (Base-Erosion and Anti-Abuse Tax) on income from low-tax jurisdictions, as well as measures to encourage U.S. based intangible property, such as FDII (Foreign Derived Intangible Income), a reduced tax rate for income generated abroad attributable to intangible property held domestically. The bill also included a one-time repatriation tax on all foreign-held income, which went from 12% in the first version considered by the House Ways and Means Committee on November 2, to 15.5% in the final version of the bill.

Finally, the TCJA also included various changes to the taxation of non-corporate business income, which might have affected competitors of corporations in the markets in which they operate. In particular, it allowed for a deduction of up to 20% of Qualified Business Income from a firm's tax liability. This deduction is limited

[6] See, for instance, here for Sen. Pat Toomy's call to"focus on taxes right now" after the last Republican attempt to reform health care.

based on whether a business is designated as a personal service firm[7] and on a person's total income.

1.5. Empirical Analysis

1.5.1. Controlling for Market Forces

Analyzing raw stock market returns could conflate the effects of interest with other factors that happen to coincide with them. This issue is particularly relevant if one is trying to infer aspects of the heterogeneity of the effect across firms. Controlling for market forces, while necessary to construct a counterfactual, could distort the overall magnitude of the effect of the tax bill. Suppose for instance each firm's expected excess return is equal to a constant times the market excess return. If the TCJA affected the return on the market portfolio, then we might mistake movements in a firm's return due to changes in the market return with movement due to changes in the corporate tax expectations. This is the reason why Gaertner et al. [forthcoming] use raw returns rather than basing their analysis on an asset pricing model.

this chapter opts instead to use "abnormal returns" as the dependent variable,[8] rather than raw returns in excess of the risk-free rate, in order to control for change in these risk factors. Appendix A.4 shows that the sign of the results is mostly unchanged when using raw excess returns or alternative asset pricing models, although with larger standard errors. [9]

In order to take financial market forces into account, the analysis begins by estimating a model of stock market returns in the spirit of Fama and French

[7] This would be the case, e.g., for a doctor's office or a law firm.

[8] This is akin, for example, to the approach adopted by Wagner et al. [2018b] in a very similar setting regarding the TCJA.

[9] Specifically, the results for market capitalization, which are the weakest among the main results, sometimes change sign. The results for Herfindahl-Hirschman index have the same sign but are not significant under other asset pricing models. The results for the Lerner index have the same sign and magnitude under all specifications, but lose significance under the 5-factor Fama-French model.

[1993] on a period *preceding* the election of President Trump, to get a sense of how each company's stock moves with the rest of the financial market. In their three-factor model the stock market returns of company $i = 1, \ldots, N$ in day $t = 1, \ldots, T$, $R_{i,t}$, depends on the risk-free rate of return, R_t^f, the returns to a market portfolio, R_t^m, the persistent effects of book-to-market equity, a High-Minus-Low portfolio (HML_t), and the persistent effects of firm size as measured by its market capitalization, a Small-Minus-Big portfolio (SMB_t):[10]

$$R_{i,t} - R_t^f = \beta_i^m (R_t^m - R_t^f) + \beta_i^{HML} HML_t + \beta_i^{SMB} SMB_t + u_{i,t}.$$

Assuming that the β's in this model do not change between the pre-period consider and the sample period relevant for the analysis, the estimates $(\hat{\beta}_i^m, \hat{\beta}_i^{HML}, \hat{\beta}_i^{SMB})$ can be used to construct expected returns during our sample period of interest, namely the months leading up to the passing of the TCJA. I will use these expected returns to construct "abnormal returns" for each company in every period, $AR_{i,t}$:

$$AR_{i,t} \equiv (R_{i,t} - R_t^f) - \hat{\beta}_i^m (R_t^m - R_t^f) - \hat{\beta}_i^{HML} HML_t - \hat{\beta}_i^{SMB} SMB_t.$$

Having abnormal returns for every company in the sample, the paper proceeds to study how news shocks regarding the tax bill affected returns relative to what we would have expected. More specifically, in the spirit of the model laid out in section 3.3, the plan is to study whether firms that might be deemed more profitable ex-ante also saw the biggest increase in returns upon news of the tax cut.

[10] These last three values, R_t^m, HML_t, and SMB_t, are taken directly from Kenneth French's website.

1.5.2. Relation to Measures of Profitability and Market Power

As section 3.3 discusses, we should observe excess returns to be higher for firms that do not face potential rent-eroding entry, or that face entry only in longer periods of time. This section uses a continuous differences-in-differences setting to study the relationship between the excess returns upon news of the tax bill and several aspects of firms. The main specification of interest is of the sort:

$$AR_{i,t} = \beta_0 + \beta_1 M_i + \beta_2 M_i \times TCJADates_t + \zeta_t + \xi_{\mathcal{I}(i)} + \gamma' \mathbf{X_{i,t}} + \varepsilon_{i,t}, \qquad (1.3)$$

where $TCJADates_t$ is a dummy equal to one if day t is one of the six dates identified by Gaertner et al. [forthcoming];[11] ζ_t are a set of time fixed-effects; $\xi_{\mathcal{I}}$ are a set of industry fixed effects;[12] and $\mathbf{X_{i,t}}$ is a vector of controls, including what fraction of a company's profits came from abroad in previous years, and the average tax rate that they face abroad (both are assumed to be zero if a firm only operates domestically). This is quite important because the TCJA changed many provisions regarding the treatment of corporations' foreign income, which likely affected future profits differently depending on how much profit a company was earning abroad, and depending on the locations of these operations, including the tax rates faced therein. As a result, the vector of controls includes the interaction between these two variables and $TCJADates_t$. M_i is a measure of firm i's size, profitability, and market power; I explore three different measures: the firm's market capitalization, the firm's market share interacted with the Herfindahl-Hirschman Index

[11] While this procedure weighs each date equally, further analysis in section 1.5.3 will decompose the effect of each date.

[12] Industry here is defined as the first three digits of a company's NAICS code.

(henceforth, HHI) of its industry,[13] and the firm's Lerner Index.[14] All of these, as well as controls regarding foreign income, are measured at the end of fiscal year 2016 to minimize the possibility that firms might have adjusted their behavior in anticipation of the tax reform.

The resulting estimate of β_2 will indicate how firms with a larger measure M_i were differently impacted by the news of the tax bill, by measuring how their stock market returns differed, on average, during our six dates of interest. If the ability to exclude competitors, in the short run or in the long run, mattered for the distribution of excess returns across companies, and if indeed this ability is reflected in M_i, this should result in positive estimates of β_2.

The first version of specification 1.3 uses market capitalization as a measure of size, and looks at different segments of companies quoted on the NYSE. Results are reported in Table 1.1. The coefficient on Market Cap $\times TCJA\ dates$ is negative for firms at the bottom of the distribution of market capitalization, but becomes positive and significant for firms in the top quartile and decile of the distribution. The coefficient is even bigger for firms in the top 1%, though not statistically significant – an unsurprising result, perhaps, given how much sample size was reduced. Further, market capitalization is likely to be a noisy measure of size, especially since it is being measured a year in advance of the periods we are considering to avoid possible anticipatory effects.

While size is generally not very strongly correlated with excess returns upon news of the tax cut, this might be expected. As Appendix A.1 points out, differences in size can be explained by differences in entry costs even in the case where firms face free and immediate entry, and thus bear none of the burden of the corporate tax. Measures of profitability and market power exhibit much stronger

[13] Based on Compustat sales data and 3-digit NAICS sectors. Unfortunately, using Compustat sales data excludes a number of potential competitors who do not file 10-K's and thus do not appear on Compustat. Appendix A.6 shows that results obtained here with Compustat data can be replicated qualitatively using concentration data from the U.S. Census, which are only available for the manufacturing sector.

[14] The Lerner index is also built from Compustat data following established literature, as the ratio of operating income before depreciation minus depreciation, over sales.

Table (1.1) Continuous diff-in-diff regressions of abnormal returns. Standard errors, in parentheses, are clustered by day.

	Dependent Variable: Abnormal Returns						
	All Firms	Quartile I	Quartile II	Quartile III	Quartile IV	Top decile	Top 1%
Market Cap.	-0.000173**	-0.654	-0.0181	-0.0220**	-0.0000458	0.00000641	0.000151
	(0.0000689)	(0.439)	(0.0519)	(0.00878)	(0.0000627)	(0.0000643)	(0.000101)
Market Cap. $\times TCJA\ dates$	0.000155	-8.023	-0.388	-0.0382	0.00130**	0.000754*	0.000560
	(0.000748)	(5.328)	(0.536)	(0.0593)	(0.000484)	(0.000424)	(0.000795)
Proportion of foreign profits	-0.00132	-0.000861	-0.00120	0.00108	-0.00187	0.00301	-0.00145
	(0.000824)	(0.00228)	(0.00139)	(0.00235)	(0.00118)	(0.00223)	(0.0169)
Proportion of foreign profits $\times TCJA\ dates$	-0.00556	0.0115	-0.0224	-0.0112	0.0170**	-0.0171	-0.0843
	(0.00578)	(0.0112)	(0.0155)	(0.0211)	(0.00829)	(0.0299)	(0.184)
Avg. for. tax rate	-0.00196	0.000302	0.0000975	-0.00506	-0.00357	-0.00282	-0.00646
	(0.00162)	(0.00717)	(0.00183)	(0.00313)	(0.00511)	(0.00654)	(0.0490)
Avg. for. tax rate $\times TCJA\ dates$	0.0154	-0.0488	0.0295**	0.0160	0.0692**	0.0392	-0.292
	(0.0136)	(0.0447)	(0.0107)	(0.0250)	(0.0252)	(0.0477)	(0.228)
N	1,126,540	266,018	279,477	284,891	296,437	119,008	12,452
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

	Dependent Variable: Abnormal Returns	
	Herfindahl	Lerner
Market share	-0.173	
	(0.118)	
$HHI\times$ market share	0.226	
	(0.206)	
$HHI\times$ market share $\times TCJA\ dates$	0.898*	
	(0.465)	
Lerner index		5.11e-12
		(1.34e-11)
Lerner index $\times TCJA\ dates$		0.000286**
		(8.05e-11)
Proportion of foreign profits	-0.00143*	-0.00166**
	(0.000818)	(0.000815)
Proportion of foreign profits $\times TCJA\ dates$	-0.00604	-0.00558
	(0.00588)	(0.00575)
Avg. for. tax rate	-0.00119	-0.00129
	(0.00162)	(0.00161)
Avg. for. tax rate $\times TCJA\ dates$	0.0155	0.0175
	(0.0133)	(0.0127)
N	1,274,778	1,166,039
Industry FE	Yes	Yes
Time FE	Yes	Yes

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

Table (1.2) Continuous diff-in-diff regressions of abnormal returns. Standard errors, in parentheses, are clustered by day.

statistical relations. As we can see in Table 1.2, market share and market concentration, as well as the accounting profit margin measured by the Lerner index, increase how a company's return increased in response to news of the tax bill. These results are consistent with the explanation that differences in market power have been a significant factor in explaining the cross-firm heterogeneity in excess returns due to news of the tax bill.

In order for these results to be interpreted as a valid differences in differences design, one should check that there are no significantly different pre-trends between groups with different degrees of treatment. This test is complicated in this case by the fact that differences in M_i can only pick up variation in the intensity of treatment, but there is no group of firms that was not affected by the tax bill, which leaves us without a true control group. Naturally, since we are analyzing stock market returns, any theory which predicts excess returns to follow a random walk would have to satisfy the no pre-trends assumption. Nonetheless, Appendix A.5.1 provides some empirical robustness checks for this assumption, showing that in the months leading up to the passing of the TCJA, the interaction between each date and each measure of M_i tracks changes in betting market odds.

Even though the main results of interest in Table 1.2 are statistically significant, even after clustering standard errors by day, the magnitudes seem surprisingly small. Column (1) implies that compared to being in perfect competition (that is, $s_i \approx 0$ and $HHI_{\mathcal{I}} \approx 0$), being a monopolist in one's sector only raises the gain from the tax bill by less than 0.9 percentage points. Column (2) implies that a one standard deviation increase in a firm's Lerner index is associated with a 0.7 increase in excess returns. These number represent a fraction of the standard deviation in excess returns, which is around 3.5 percentage points in the sample. It should be noted, however, that these results are averaging out across six different dates, throughout a time when the tax bill was undergoing major changes, and which did not necessarily reflect the entire effect of the tax bill. In order to better quantify the effect of news on each date, the next section makes use of betting data.

1.5.3. Cardinal Interpretations

The existence of betting markets specifically on the corporate tax reform quantify the news shocks used in the study, which in turn can yield a better quantitative interpretation of the results. As section 1.2 points out, what the empirical work is trying to measure is not really the reaction of a firm's value to a change in tax rate, but rather the reaction to a change in *expectations* about future tax rates. Suppose that with probability P, the tax rate is decreased by a marginal amount, and that otherwise it remains the same. Taking a weighted average of excess return as defined in equation 1.2, we obtain:

$$\mathbb{E}_P[r^e] = Pr^e + (1 - P)0 = Pr^e.$$

Date	ΔP
September 27	N/A
November 2	0.0428
November 16	0.0286
December 2	0.0316
December 15	0.0442
December 20	0.0065

Table (1.3) Change in average price of the bet "Will there be a corporate tax cut by March 31, 2018?" on each of the dates identified by Gaertner et al. [forthcoming]. Note that betting started on October 23, so there is no data in this series for September 27. The same date on the bet asking "Will there be a corporate tax cut by December 31, 2017?" yields a ΔP of 0.0397. December 2, 2017, the date on which the Senate passed the TCJA, was a Saturday. As a result, I effectively look at returns on the following Monday, December 4, 2017. The change in probability reported here is the sum of the changes between Saturday and Monday.

Note that $r^e = 0$ in the event of no tax change, because when the tax rate does not change, the company's equity is simply earning the normal rate, or in other words is earning zero excess returns.

Observing a change in P, as betting data permits us to do, allows us to back out the full extent of the excess return due to a marginal reduction in the tax rate. Suppose, for instance, that the probability of a tax cut goes from P to P'. Ignoring discounting issues over such short time periods, we have:

$$\Delta\mathbb{E}_P[r^e] \equiv \mathbb{E}_{P'}[r^e] - \mathbb{E}_P[r^e] = r^e(P' - P),$$

or, letting $\Delta P \equiv P' - P$,

$$\frac{\Delta\mathbb{E}_P[r^e]}{\Delta P} = r^e.$$

With this reasoning in mind, I run specification 1.3 allowing for the effect of each of the six dates considered to differ, and then I deflate the result for each date by the change in probability of a corporate tax cut on that date, as measured by the futures prices on PredictIt.org. These changes in probability are documented in Table 1.3.

	Market cap. $\times \ldots$	ΔP-weighted coefficient
Market Cap	November 2	0.027***
		(0.0016)
	November 16	0.0466***
		(0.0024)
	December 2	0.0106***
		(0.0021)
	December 15	0.0482***
		(0.0015)
	December 20	-0.0616***
		(0.0104)
	$HHI \times$ market share $\times \ldots$	
HHI	November 2	5.0879***
		(1.3257)
	November 16	13.2638***
		(1.9838)
	December 2	104.8466***
		(1.7951)
	December 15	22.4177***
		(1.2814)
	December 20	87.7112***
		(8.7725)
	Lerner index $\times \ldots$	
Lerner	November 2	0.0033***
		(0.0006)
	November 16	-0.0014
		(0.0008)
	December 2	-0.0017**
		(0.0008)
	December 15	0.005***
		(0.00055)***
	December 20	0.127***
		(0.0037)

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

Table (1.4) Regression results of specification 1.3 disaggregated by date and re-weighted by ΔP. Raw regression results are reported in Appendix A.3. Note that date 1 is included in the analysis so as to not count it as a "control" day, but omitted in these results because there were no bets on that day in the betting series we are considering. Standard errors clustered by day in parentheses.

Results are reported in Table 1.4. The impact of market capitalization, measured in billions of dollars, remains small even though I am restricting the sample to the top quartile of the market cap distribution, but again is likely to be the victim of attenuation bias. The impact of the Lerner index small, due to the major variation of the Lerner index in the data, as one can see in Table C.1. For December 20, for instance, a one standard deviation change in Lerner index implies an increase in returns of 29.9 percentage points. Results for the HHI look even bigger in magnitude and statistically more significant. Compared to a firm in perfect competition, these results suggest that a monopolist gained between 5 and 105 percentage points more due to the tax bill, depending on which date we consider – a large and somewhat puzzling range. Overall, the betting data help us see that there is a strong relationship between the gains from the tax bill and other traditional measures of profitability and market power, which in turn motivates the next empirical endeavor looking at the effects of the tax bill by individual firm.

Another interesting aspect of these results is how heterogeneous the measured effects are across the dates considered. Even excluding results for September 27, which would rely on a different betting series, there is considerable variation in the magnitudes of these coefficients. If all the information that was released during the dates this study focus on had to do with the probability of the bill passing, and not with different provisions and updated expectations of the bill, then we should observe roughly similar coefficients across each date after adjusting for the "size" of each event date.

One way to explain this variation is that different versions of the tax bill were being considered as time went on. This complicates interpretation, as it is far from trivial to predict what investors were and were not surprised by as new versions of the bill became available. In this sense, perhaps the last date under consideration, December 20, bears the cleanest identification argument – as the bill had undergone major changes which were by then established, and it was merely being sent for signature by the President, which happened two days later. Taking this argument to its extreme conclusion implies very large effects of market power and profitability on excess returns due to the tax cut.

1.5.4. The Distribution of Excess Returns

This section studies how the excess returns due to the news of the tax bill varied across individual companies. The procedure simply models the abnormal returns associated with our dates of interest to differ by each firm individually through the following fixed-effects model:

$$R_{i,t} - R_t^f = \eta_i + \omega_i TCJADates_t + u_{i,t}. \tag{1.4}$$

In this specification, η_i represents the average excess return of firm i on days other than the six dates of interest, and $omega_i$ represents the average difference in firm i's returns on those dates. While this model allows estimating the effects of the news firm by firm, it forgoes controlling for a number of forces that we were instead able to control for in previous work.

The fixed-effects model exhibits two main differences compared to the differences in differences specification of section 1.5.2. First, as section 1.5.1 discusses, the total impact of the bill is bound to include effects on aggregate quantities such as the market return. The identifying assumption while using abnormal returns, then, was that the parameters governing asset pricing did not change upon news of the tax bill. In order to take the same approach while measuring the total effect of the news, we'd need to know what the effect of the tax bill was on the risk factors affecting asset pricing. For this reason, this section instead opts to look at the effect on raw excess returns instead, and simply controlling for the average return η_i.

Second, this model cannot control for time fixed effects. Section 1.5.2 was purely interested in how the returns of companies with different measures M_i reacted differently to news of the tax bill, which meant that all identifying variation for the coefficient of interest was coming from variation across firms. Now that we are interested in the effect of the news *for each firm*, instead, all the identifying variation comes from differences for a given firm across time.

One downside of this approach is that it cannot separately control for how variables concerning firms' foreign operations interact with news about the tax bill, as variables on foreign operations only vary cross-sectionally. This means that ω_i captures the average effect of the tax bill as a whole, not just the reduction in tax rates. One might think that this could mean results are driven by the international provisions of TCJA, but Wagner et al. [2018b] provide evidence that the proportion of income coming from abroad was not significantly correlated with excess returns on some of the dates we are considering, while it was negatively correlated in others. Rather than reflecting the overall provisions of the tax bill, thus, this is likely reflective of an increase in the repatriation tax rate, as the authors themselves point out. This significantly reduces worries that the results of specification 1.4 might be driven by differences in exposure to provisions on foreign income. Further, the empirical results of the previous section should reassure us that at least part of what is being measured has to do with differences in productivity and market power.

Results for specification 1.4 are plotted in Figure 1.4. The average increase in returns is concentrated around zero, but has a long right tail, which is consistent with the explanation regarding the increase in mark-ups and the concentration of market shares offered by De Loecker et al. [2018]. As Appendix A.7 shows, this same distribution is not typical of most dates, and cannot be replicated by repeating the analysis with six dates chosen at random.

Weighing each date by the innovation in betting prices yields a clearer quantitative interpretation of the news on each day, similarly as in section 1.5.3. The dummy variable for each date of interest in $TCJADates_t$ is substituted with the change in odds on that date, and then equation 1.4 is estimated again. The re-

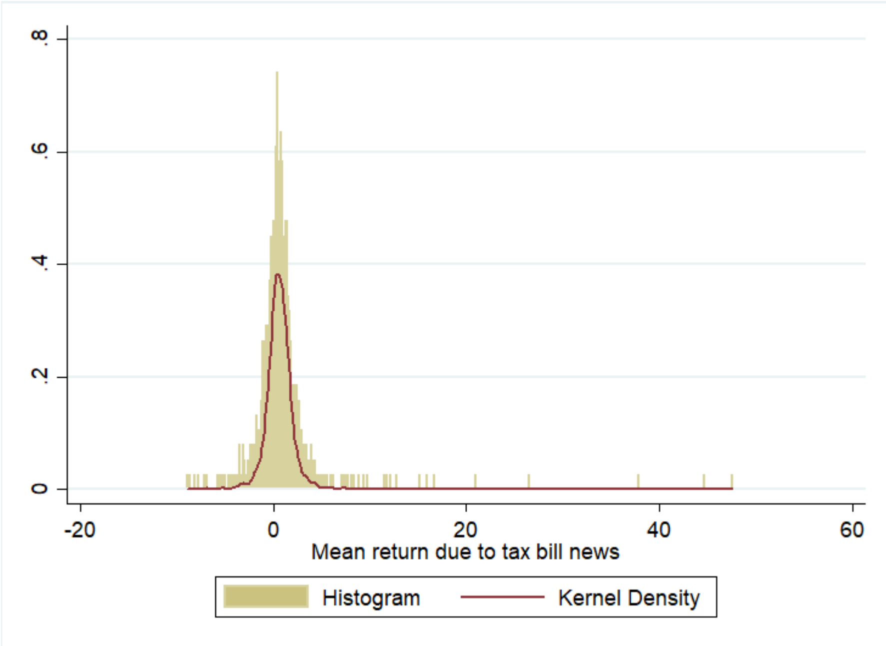


Figure (1.4) Distribution of average change in returns due to the news of the tax bill.

sulting distribution of excess returns is shown in Figure 1.5. Figure 1.6 shows how the distribution of excess returns differed for more and less valuable companies. Companies higher up in the distribution of profitability, as measured by the Lerner index, were substantially more likely to experience a positive gain upon news of the tax bill than other companies. This complements the analysis in section 1.5.2, showing that smaller firms were not just more likely to have a smaller gains upon the the tax cut, but also that they were more likely to have no gain at all – which, in the language of equation 1.2, is what would not happen for a firm with no rents ($\Pi_{i2} = 0$) and facing immediate potential entry ($\alpha_i = 1$).

Multiplying the excess return estimated in this procedure with a firm's market capitalization, one can calculate the implied total effect of the tax bill on a firm's valuation. Since this calculation uses the market capitalization at the end of fiscal year 2016 avoid the possibility that considerations about the tax bill are already reflected in this market capitalization, the result can be thought of as a proxy for the real impact of the tax bill on the valuation of each firm. The resulting distribution of excess returns is reported for all firms in Figure 1.5, while the distribution of projected valuation gains (in billions of U.S. dollars) is reported in Figure 1.7. Again, the distribution exhibits a concentration around zero with a long right tail.

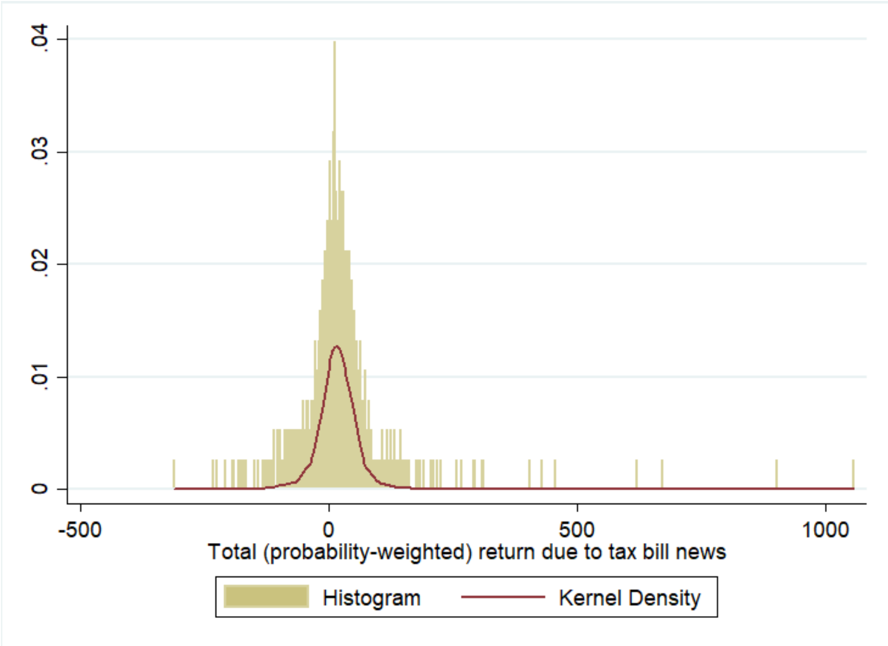


Figure (1.5) Distribution of the total change in excess returns due to the tax bill.

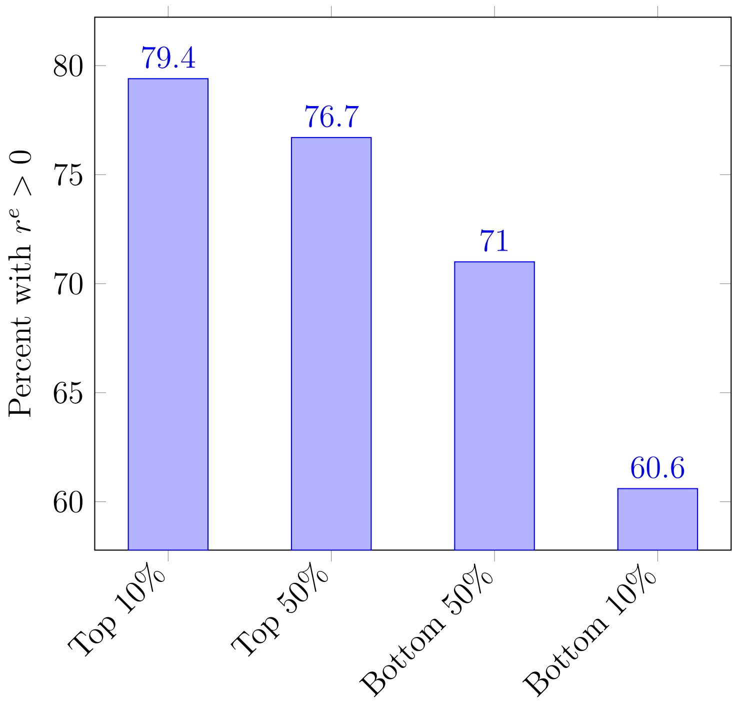


Figure (1.6) Percent of companies with positive excess returns upon news of the tax bill based on where they fall in the distribution of Lerner index.

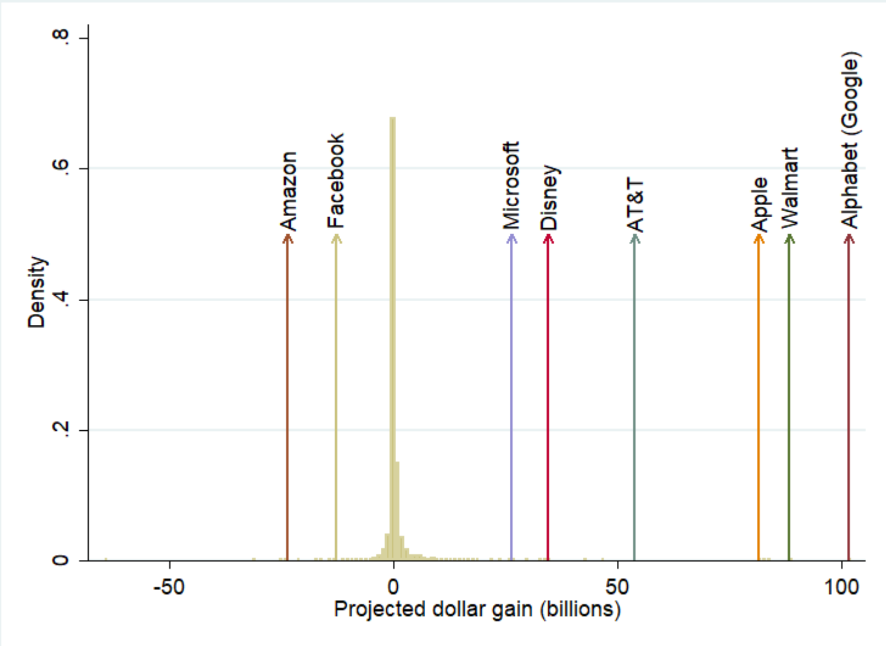


Figure (1.7) Distribution of the total dollar gain due to the tax bill for all firms.

A few well known companies where selected to showcase how much they gained from the bill.

These results should be interpreted with some caveats. For example, only 27% of the estimated excess returns are statistically distinguishable from zero at the 5% significance level. Furthermore, the median magnitude of the t-statistic for each ω_i is 1.17, indicating that many of these estimates have a high standard error relative to their point estimate. Finally, using market capitalization a year in advance means we can only proxy for the original value of each company.

With this in mind, Table 1.5 shows that not only did firms with a higher profit margin gain a bigger share of the gains from tax reform, but that they did so more than proportionally to their initial share of market capitalization. This evidence reinforces the interpretation that the heterogeneity in excess returns due to the tax cut was at least partly driven by differences in market power and the possibility of rent-eroding entry.

Percentile of tax bill gain:	Top 1%	Top 5%	Top 10%	Top 25%	Top 50%	Top 75%
Percent of total market value	22.0	45.1	53.3	63.4	69.6	70.8
Percent of total projected gain	50.6	88.0	103.6	122.5	131.4	132.4
Lerner index percentile:	**Top 1%**	**Top 5%**	**Top 10%**	**Top 25%**	**Top 50%**	**Top 75%**
Percent of total market value	2.3	7.2	10.5	47.3	78.4	93.7
Percent of total projected gain	3.3	13.3	18.6	50.7	80.9	100.7

Table (1.5) Percent of total market value and percent of total gain from the tax reform for several percentiles of the distribution of the Lerner index and of dollar gains. Note that some groups may earn more than 100 percent of the gains because some firms are impacted negatively.

1.6. Conclusion

This chapter studies how the market valuations of different U.S. publicly traded corporations were affected by news regarding an imminent tax cut. The results indicate that there was wide dispersion in the benefits of the tax cut. While the average impact was positive, some firms benefited little or not at all - indeed, roughly a quarter of firms lost value upon news of the tax cut - while others benefited a great deal. As noted in Table 1.5, these benefits were particularly concentrated among the most profitable firms. This heterogeneity is consistent with heterogeneity in the potential competition that a firm faces.

The paper further investigates the link between the excess returns due to news of the TCJA and market power. It finds that among the biggest firms, size is weakly correlated with larger excess returns. It also finds that a firm's excess return is positively related to its accounting profit margin and its market concentration. Using data from political betting markets, it becomes possible to assign cardinal interpretations to these estimates, which suggest sizable effects that are highly heterogeneous across dates. Using the date of last approval of the bill before it was signed into law, these results suggest that on average a monopolist earned an excess return 87.7 percentage points higher than a firm in perfect competition, and that moving up one standard deviation in the distribution of the Lerner index of firms was linked with an average increase of 29.9 percentage points.

This chapter complements a growing literature documenting the rise of market power in the U.S. An analysis of forward-looking financial markets' response to tax cuts suggests that market expectations are consistent with the historical trend of a concentration of economic profit margins at the top of the distribution, in keeping with the evidence found in other parts of the literature. These results also contribute to the literature on corporate tax incidence. Responses to the tax cut are consistent with heterogeneous economic rents, and therefore with the notion that the shareholders of firms with high rents will bear much of the burden of corporate income tax.

Chapter 2. Attending to Inattention: Identification of Deadweight Loss Under Non-Salient Taxes

2.1. Introduction

Taxing a good results in a loss of economic efficiency whenever it distorts equilibrium behavior away from the Pareto optimum. To the extent that agents do not notice a tax, the burden of the tax is exacerbated by the fact that agents cannot adjust the behavior to protect themselves from the tax. However, the burden of the tax in excess of government revenue, or deadweight loss, is mitigated when agents do not pay attention to the tax: if consumers pay the tax without noticing it, they are effectively transferring some of their income to the government in a lump sum. Chetty, Looney, and Kroft (2009, henceforth CLK) were the first to make these points. In addition to their theoretical contributions, they also showed that consumers in the U.S., where sales tax is applied at the register rather than included on the prices displayed on shelves (or, "sticker prices"), tend to under-react to sales taxes.

While CLK (2009) focuses on the case of homogeneous attention, recent work by Taubinsky and Rees-Jones (2018, henceforth TRJ) has noted that introducing the possibility of heterogeneous attention may prevent the computation of deadweight loss from aggregate data. If each person faced a different tax rate when buying a certain good, understanding welfare effects would require us to study not only aggregate demand, but the demand of every individual. Imposing a high tax on low elasticity individuals and a low tax on high elasticity individuals would have a very different effect on welfare than doing the opposite. A similar reasoning applies when all agents face the same tax rate, but perceive different tax rates. TRJ (2018) find that allowing for heterogeneous attention introduces an issue of allocative inefficiency that is normally absent from the study of the welfare effects of taxation. In a world of heterogeneous attention, there is no guarantee that the individuals who end up consuming the good are the ones who value it the most.

TRJ (2018) make these points in a binary choice model. This is well-suited to their experiment, in which people are choosing whether or not to buy a certain object, but their proof does not generalize trivially. Given the predominance of continuous choice settings in much of the literature on tax salience, including CLK's seminal paper, this motivates us to study the issue further.

We begin by developing a model of choice under misperceived prices with an arbitrary closed consumption set, and develop our welfare measure: compensating variation due to the tax, net of tax revenue. Bernheim and Rangel (2009) laid the foundations of welfare analysis with behavioral agents. Our model is similar to the models of CLK (2007, 2009) and TRJ (2018), but we slightly modify the treatment of income effects, along the lines of Gabaix's (2014) model of rational inattention. In the absence of income effects, our model of choice for an individual agent is essentially equivalent to the model in CLK (2009), except that we allow for arbitrary consumption sets. Our model is also similar to the model of Chetty (2009), but for the fact that we specify a particular way in which behavioral agents maximize utility. While this does not impose severe restrictions on behavior, it offers a useful framework when we move on to identification. We confirm that some of the major results in CLK (2009) and TRJ (2018) hold quite broadly: inattention to taxes increases the size of the loss in consumer surplus, but decreases the size of deadweight loss; attention heterogeneity amplifies deadweight loss, and invalidates CLK's sufficient statistic approach.

The main contribution of this chapter is to generalize TRJ's non-identification result to an arbitrary closed choice set. We show that an econometrician who only observes aggregate consumption data can only determine the true value of aggregate deadweight loss to lie on an interval. These bounds were first noted by TRJ (2018) in their proposition A.2. We find these bounds hold generally and propose to use them as a novel empirical tool.

The lower bound for deadweight loss is the calculation one would perform in the case of a representative consumer. Since the loss in efficiency is a convex function of the perceived tax rate, the calculation of deadweight loss from *one* perceived tax-inclusive price consistent with aggregate demand will generically underestimate deadweight loss. Heterogeneity in tax salience creates heterogeneity in perceived net-of-tax prices, which creates an allocative inefficiency across consumers. As the calculation with a representative consumer only accounts for inefficiency from *aggregate* foregone consumption due to the tax, it will under-estimate excess burden. However, in the case in which all agents pay the same amount of attention to the tax, there is no allocative inefficiency between consumers, and so performing the calculation as with a representative consumer yields the correct value for deadweight loss. The formula for this lower bound to deadweight loss is an extension of formulas provided by CLK (2009) and TRJ (2018).

Following TRJ (2018), we obtain an upper bound for deadweight loss by letting the econometrician assume that tax salience has support on a known bounded non-negative interval. The upper bound comes from maximizing perceived price heterogeneity, again exploiting the convexity of deadweight loss with respect to the perceived tax. This is achieved by positing that agents have either zero or maximal salience. Generalizing introduces two additional considerations in calculating the upper bound for deadweight loss. One, a distribution yielding the upper bound for deadweight loss assigns high tax salience precisely where it will "hurt" most: to those agents whose particular preferences yield maximal deadweight loss from that agent relative to the change in consumption of that agent. This distribution allocates high tax salience to those agents who have more convex demand curves, keeping the aggregate change in quantity demanded constant. Two, deadweight loss is maximized for a given aggregate demand if any agent with multiple optimal decisions at the perceived price consumes the highest amount consistent with their preference when they perceive low prices, whereas they consume the lowest amount consistent with their preferences when they perceive high prices.[1] This is because heterogeneity in perceived prices permits different equilibria with the same sticker price, tax rate, and aggregate consumption, yet yielding different values of deadweight loss due to different distributions of consumption among individuals.

Our approach to compute the upper bound forces the econometrician to impose limits on the possible values of attention – something that the empirical literature has not settled yet and might be highly context-dependent. While it might seem natural to assume that attention varies between zero and one, many papers have found evidence of salience above one – see for instance Allcott and Taubinsky (2015). In theory, one might allow for unlimited (positive) tax salience, under regularity conditions that avoid the possibility of unlimited distortion to consumer behavior.[2]

Our general results follow the work of TRJ (2018), and rely on the fact that the distribution of preferences is independent of taxes and prices, but that the distribution of salience might change depending on the value of the tax. Indeed, we show that, even assuming that the distribution of preferences is entirely known to the econometrician and invariant to observables, one cannot identify deadweight loss.

We then turn to the special case in which demand is linear in a relevant range where consumption is positive. This case is of particular interest for three reasons: first, its ease of applicability; second, its special relationship to the second order approximation of deadweight loss; and third, its illustrative value in the kind of

[1] TRJ (2018) do not deal with cases like this because they restrict attention to non-atomic distributions of willingness to pay, so almost every agent has a unique choice that is perceived to maximize utility.

[2] One might assume that consumer surplus is uniformly bounded to ensure that the upper bound for deadweight loss is finite even if the upper bound for tax salience is infinite.

problems we can face in identifying deadweight loss. In the case of binary choice, non-identification ultimately comes from the possibility that taxes and attention to taxes are not independent of each other. Although the experimental evidence in TRJ (2018) suggests that attention varies with how large the tax is, if one assumed away this possibility we could identify a full distribution of responsiveness to both taxes and sticker prices using existing models of discrete choice with random coefficients, as in Masten (2017) and Fox (2017). This estimated distribution could then be used to yield a point-estimate of deadweight loss. In the case of linear demand, instead, even assuming that attention and taxes are independent would not help identify deadweight loss with aggregate data beyond the bounds described above.

Our results complement a growing literature on tax salience. Rosen (1976) does not find evidence of limited tax salience, but besides CLK (2009) and TRJ (2018), Finkelstein (2009), Gallagher and Muehlegger (2011), and Goldin and Homonoff (2013) all find strong evidence of dramatically limited tax salience. Most of this literature looks at sales taxes in the U.S., as they lend themselves very credibly to a story about lack of salience. However, work on salience has also looked at other settings: Finkelstein (2009) studies car tolls; Weber and Schram (2016) study whether income taxes being remitted by the employer or the employee affects differently people's attitude towards public spending and the burden of the tax;[3] Morone, Nemore and Nuzzo (2018) explore a similar question in the context of a double-auction market. Blake, Moshary and Tadelis (2017) study how people react differently to back-end and upfront fees in online purchases; and Bradley and Feldman (2018) study how changes in the disclosure of ticket taxes affect the demand for airlines. As we mentioned above, most of these empirical papers, as well as other theoretically focused papers like Goldin (2015), use models where the choice set is continuous or mixed discrete-continuous.

The paper proceeds as follows. In section 2.2, we develop a general model of

[3]Interestingly, assuming some agents face credit constraints as in Boadway, Garon, and Perrault (2018), also breaks traditional optimal tax theory.

choice under misperceived prices. Once we have replicated some of the major theoretical results in CLK (2009) and TRJ (2018), we shift focus to identification of deadweight loss. Section 2.3 lays out the main results of our paper, establishing the non-identification result and the bounds. Section 2.4 focuses on the special case in which demand is linear, and provides a straightforward way to compute our bounds in the context of linear models. Section 3.5 concludes. Proofs and other minor results are relegated to the online appendix.

2.2. Choice and Deadweight Loss under Non-Salient Taxes

This section describes the theoretical model and results that underlie the rest of this chapter. Many of our results here simply mirror previous literature, but we make slightly different modeling choices. The main modeling challenge in dealing with misperceived prices is to allow for the misperception of prices while keeping agents financially solvent. CLK (2007, 2009) assume that one good "absorbs" all optimization mistakes. In contrast our approach, inspired by parts of the model in Gabaix (2014), has agents conjecture a certain income such that they end up consuming on their true budget constraint when presented with the relative prices they perceive. While this framework preserves all results of interest from CLK (2009) and TRJ (2018), we find that our approach eases exposition while freeing the researcher from having to make ad-hoc assumption about which good (or goods) absorb optimization mistakes. It should be noted that while we do generalize the model to include multiple taxed and non-taxed goods in the online appendix, in the body of the paper agents will face a choice over two goods, only one of which is subject to tax.

The agent has a closed consumption set $X = X^T \times \mathbb{R}_+ \subseteq \mathbb{R}^2_+$. She also has a choice function for the taxed good, $q(\bar{p}, p^{NT}, \tau, W)$, with $(\bar{p}, p^{NT}) \in \mathbb{R}^2_{++}$, where $\bar{p}$ and p^{NT} are respectively the sticker price of the taxed and non-taxed good, $\tau \in \mathbb{R}$

is the sales tax on the taxed good, and W is the income of the agent.[4] We express taxes as if they were specific, so that $\bar{p} + \tau$ is the tax-inclusive price of the taxed good.

The agent has a continuous and strictly monotonic utility function $u(q, q^{NT})$, where q denotes generic consumption of the taxed good. The choice vector function $\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W) = (q(\bar{p}, p^{NT}, \tau, W), q^{NT}(\bar{p}, p^{NT}, \tau, W)) \in X$ meets two requirements. One, the agent spends all available income:[5]

$$(\bar{p} + \tau)q(\bar{p}, p^{NT}, \tau, W) + p^{NT} q^{NT}(\bar{p}, p^{NT}, \tau, W) = W. \tag{2.1}$$

Two, the agent correctly optimizes in the choice of all consumption bundles when there is no tax:

$$\boldsymbol{q}(\bar{p}, p^{NT}, 0, W) \in \underset{\{\bar{p}*q + p^{NT}*q^{NT} \leq W\}}{\arg\max} \; u(q, q^{NT}). \tag{2.2}$$

It turns out that this model is quite general, in the sense that it rules out very few possible behaviors. Indeed, this model is equivalent to one in which agents pick rationally given a perceived price p^s, and conjecture themselves an income W^s so that they satisfy their true budget constraint at their perceived price. Proposition 2 and subsequent work in the online appendix shows that, under weak convexity assumptions on preferences, one can find a pair (p^s, W^s) to satisfy equations 2.1 and 2.2 for any choice function $\boldsymbol{q}(\cdot)$.

We now want some measure of the incidence of the tax on the consumer. For concreteness, we consider the compensating variation due to the tax with complete pass-through, defined as:

$$\Delta CS \equiv \inf\{\Delta W | \; u(\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W + \Delta W)) \geq u(\boldsymbol{q}(\bar{p}, p^{NT}, 0, W))\}.$$

[4]We implicitly restrict consideration to sticker prices, taxes, and income such that $\mathbf{q}(\bar{\boldsymbol{p}}, \tau, W)$ is well-defined at those values.

[5]When we consider multiple non-taxed goods in the online appendix, we also have agents optimally choose q^{NT} given their choice of q.

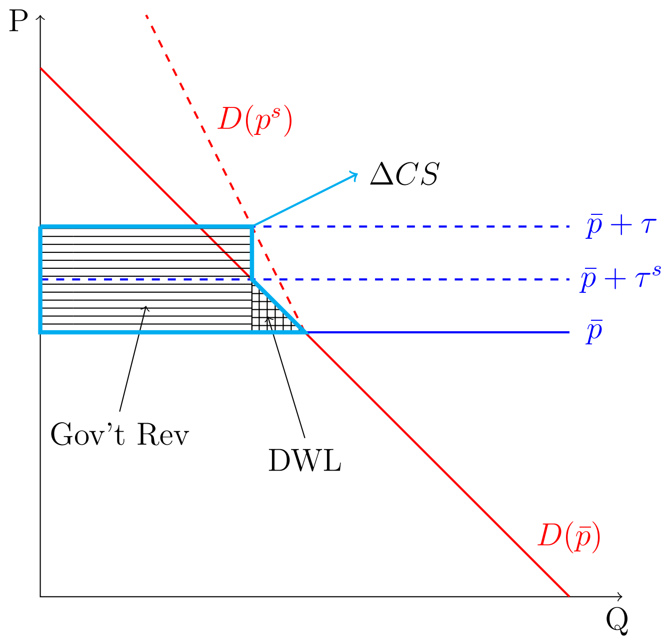


Figure (2.1) Welfare effects from the imposition of a non-salient tax.

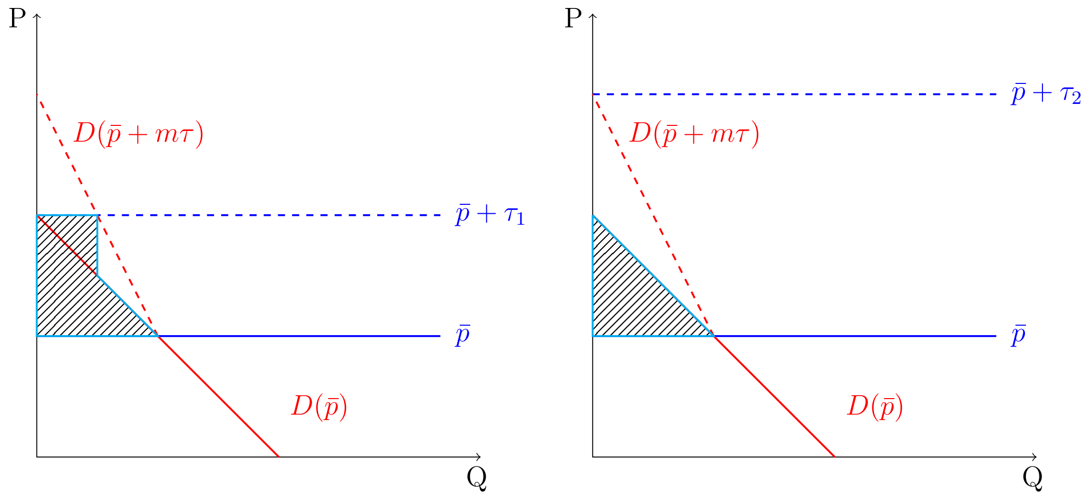
In words, the change in consumer surplus is the greatest lower bound of the amount of money the agent requires to achieve the utility reached before the imposition of the tax. Online appendix proposition 3 shows that the change in consumer surplus can be written as the sum of what consumer would have to be compensated if the tax-inclusive price were *actually* p^s and the income they "lost" due to their inattention:

$$\Delta CS = \underbrace{(\bar{p}+\tau-p^s)h(p^s)}_{\text{Income lost}} + \underbrace{e(p^s)-e(\bar{p})}_{\Delta CS \text{ under } p^s}. \tag{2.3}$$

This representation, which is graphically illustrated in figure 2.1, readily gives us two interesting results. First, as noted by CLK (2007, 2009), if $p^s \in [\bar{p}, \bar{p}+\tau]$, then the consumer will be worse off than if she paid attention to the tax:

$$\begin{aligned}
\Delta CS &= (\bar{p}+\tau-p^s)h(p^s) + \int_{\bar{p}}^{p^s} h(p)dp \\
&\geq \int_{p^s}^{\bar{p}+\tau} h(p)dp + \int_{\bar{p}}^{p^s} h(p)dp \\
&= \int_{\bar{p}}^{\bar{p}+\tau} h(p)dp.
\end{aligned}$$

Second, a consumer can be made better off by an increase in the tax, which we do not believe previous literature to have noted. This is because a tax increase can induce inattentive agents to reduce their consumption of the taxed good to zero, improving welfare for consumers who would have already avoided consuming the taxed good had they been attentive to the tax.[6] We provide a graphical example in figure 2.2.



(a) Under τ_1 the consumer loses more than total CS (b) Under τ_2 the consumer loses exactly total CS

Figure (2.2) The consumer in (a) is worse off than in (b), although she is subject to a lower tax

To obtain deadweight loss, we need to adjust ΔCS for the change in tax revenue:[7]

$$dwl \equiv \Delta CS - \tau q(\bar{p}, p^{NT}, \tau, W + \Delta CS) = e(p^s) - e(\bar{p}) - (p^s - \bar{p})q(\bar{p}, p^{NT}, W + \Delta CS). \tag{2.4}$$

Note that deadweight loss is exactly as if the agent was correct that the tax-inclusive price were p^s.

To introduce heterogeneity, let $i \in \mathcal{I}$ index consumers. Each consumer is characterized by her perception of the price of the taxed good, p_i^s, type, θ_i, standing in for her preferences $\succeq_{\theta_i}$ and income W_{θ_i}, and tie-breaking parameter ζ_i, which we

[6]More formally, consumption does not necessarily have to be reduced to zero for the consumer to be made better off. The loss of consumer surplus decreases in τ whenever the tax is sufficiently high such that consumption is sufficiently small (but positive).

[7]We maintain the convention that deadweight loss is a generically positive value.

need for technical reasons.[8] These consumer-specific parameters are distributed according to $F^*_{p^s,\theta,\zeta}$. Each agent has a choice function for the taxed good, satisfying:

$$q(\bar{p}, p^{NT}, \tau, W_{\theta_i}; \theta_i, \zeta_i) = q(p_i^s, W_i^s; \theta_i, \zeta_i) \in \{q | \exists q^{NT} : (q, q^{NT}) \succeq_{\theta_i} \boldsymbol{q'} \ \forall \boldsymbol{q'} \in X : (p_i^s, p^{NT}) * \boldsymbol{q'} \leq W_i^s\},$$

where p_i^s and W_i^s are determined as in the model of section 2.2, with corresponding expenditure function $e(p_i^s; \theta_i)$. If demand is single-valued, letting us ignore the tie-breaking parameter ζ, total deadweight loss is:

$$DWL = \int_{p_i^s, \theta_i} [e(p_i^s; \theta_i) - e(\bar{p}; \theta_i)] - (p_i^s - \bar{p}) q(p_i^s; \theta_i) dF^*_{p^s,\theta}(p_i^s, \theta_i).$$

this chapter emphasizes continuous choice, as the binary choice case is worked out in TRJ (2018). We momentarily assume that h is continuously differentiable with respect to its own price, and that p^s is continuously differentiable with respect to τ. Then tax salience at a tax rate of zero is $m = \frac{\partial p^s}{\partial \tau}|_{(\bar{p},0)}$.[9] In line with existing deadweight loss analyses, we can consider a second-order approximation, letting us characterize our object of interest in terms of first derivatives:

$$DWL \approx -\frac{\tau^2}{2} \int_{p_i^s, \theta_i} m_i^2 \frac{\partial h(\bar{p}; \theta_i)}{\partial p} dF^*_{p^s,\theta}(p_i^s, \theta_i). \tag{2.5}$$

This process of aggregation makes apparent two important points that confirm the analysis of TRJ (2018) extends naturally to continuous choice. First, allowing for attention heterogeneity introduces an issue of allocative inefficiency, as it is no longer guaranteed that consumers who value some units of the taxed good the

[8] ζ_i acts as a tie-breaker among bundles that could all have been chosen: choices do not necessarily reflect true preferences when agents misperceive prices, and agents might appear indifferent between choices that do not actually yield the same ex-post utility. This is in sharp contrast with the neo-classical model, where the actual choice that one selects among indifferent bundles has no impact on consumer surplus. But in this model of choice, different values of conjectured income might yield different choices, even given the same preferences, perceived prices, and true income.

[9] Formally, the claim is that there exist (p^s, W^s) such that $\frac{\partial p^s}{\partial \tau}$, where the derivative is taken while the consumer is being compensated, is well defined. If $\frac{\partial h}{\partial p}(\bar{p}) \neq 0$, then the Inverse Function Theorem implies that $\frac{\partial p^s}{\partial \tau} = \frac{\frac{\partial q}{\partial \tau} + \frac{\partial q}{\partial W}\frac{\partial \Delta CS}{\partial \tau}}{\frac{\partial h}{\partial p}}$. If $\frac{\partial h}{\partial p} = 0$ in a neighborhood around $\bar{p}$, then set $\frac{\partial p^s}{\partial \tau}|_{\tau=0} = 0$ and $\frac{\partial \Delta CS}{\partial \tau} = -\frac{\frac{\partial q}{\partial \tau}}{\frac{\partial q}{\partial W}}$.

most will be the ones who end up purchasing those units. It should be noted that we are assuming throughout this chapter that supply is perfectly elastic, and so tax increases will be reflected one-for-one in the after-tax price. This is not really an issue of concern: as TRJ (2018) note, all that is needed to generalize this to an arbitrary supply function is to account for the change in sticker price and the change in profits to suppliers. Nonetheless, as in TRJ's work, it is interesting to deviate for a moment from this assumption, to consider what happens to aggregate deadweight loss when supply is perfectly inelastic. In that case, we can use appendix proposition 4 to show that

$$DWL \approx -\frac{\tau^2}{2}\int_{m_i,\theta_i}\left[m_i^2\frac{\partial q_i}{\partial p}dF_{m,\theta}(m_i,\theta_i) - \frac{(\int_{m_i,\theta_i} m_i\frac{\partial q_i}{\partial p})^2}{\int_{m_i,\theta_i}\frac{\partial q_i}{\partial p}dF_{m,\theta}(m_i,\theta_i)}\right] \geq 0,$$

and $DWL = 0$ when attention is homogeneous, i.e. $m_i = m\ \forall i$. Thus, a non-salient tax may yield excess burden even without changing the equilibrium quantity, due to its effects on allocative efficiency.

Second, allowing for heterogeneous attention introduces a serious problem of identification, as neither aggregate price responsiveness $\int_{p_i^s,\theta_i}\frac{\partial h(\bar{p};\theta_i)}{\partial p}dF^*_{p^s,\theta}(p_i^s,\theta_i)$ nor aggregate tax responsiveness $\int_{p_i^s,\theta_i} m_i\frac{\partial h(\bar{p};\theta_i)}{\partial p}dF^*_{p^s,\theta}(p_i^s,\theta_i)$ are sufficient statistics for the deadweight loss in equation 2.5. These points effectively extend some of TRJ's (2018) major results to continuous choice. In the next section, we formalize this non-identification result with an arbitrary choice function, and show how one might bound deadweight loss with mere information on aggregate parameters.

2.3. Non-Identification with Aggregate Data

This section discusses to what degree one can infer deadweight loss from aggregate choice data. Our results generalize TRJ's work on binary choice to an arbitrary choice set. We find this to be of interest for two reasons. First, many papers dealing with tax salience, including the seminal paper by CLK (2009), operate in a continuous choice setting. Second, while point identification is impossible using aggregate parameters, we can still provide tight bounds based solely on aggregate (or average) quantities.

For simplicity, we assume that the econometrician already knows the distribution of preference types, but this should not be considered a limiting assumption. Consumer preferences can be identified with sticker price variation when there are no (non-salient) taxes. Regardless, even when the econometrician can fully observe the true distribution of preferences, and how much aggregate consumption there is at every tax level, she still cannot infer the exact value of deadweight loss. However, we provide a lower bound and an upper bound for deadweight loss. The lower bound is achieved by assuming that all agents perceive the same tax-inclusive price, i.e. assume there is no attention heterogeneity. The upper bound for deadweight loss is achieved by imposing maximal attention heterogeneity. Since the data do not reveal the individual variation in tax salience, one cannot point identify deadweight loss from aggregate data. Deadweight loss can take on any value between the upper and lower bounds.[10]

The results in this section are described as if all agents face the same sales tax. Also, since we are considering the problem of identification with aggregate demand, we assume there are no income effects. This is because even the standard model with fully salient taxes requires strong restrictions on income effects in order to achieve identification with aggregate data, as the same income can yield different consumption bundles whenever there are several conjectured incomes that solve the agent's problem. Suppressing income and price for the non-taxed good, we denote the consumption function for agent i with type θ_i and perceived tax-inclusive price p_i^s for the taxed good by $q(p_i^s; \theta_i, \zeta_i)$. However, all of these results follow if one reinterprets $q(p_i^s; \theta_i, \zeta_i)$ as the *compensated* choice of agent i.

To ensure integrability, we assume the econometrician knows that $F^*_{p^s}$ has support with lower bound greater than zero, and so only considers marginal distri-

[10]This claim follows by taking any weighted average of the distributions of parameters yielding upper and lower bounds of deadweight loss.

butions of subjective prices bounded above zero. The econometrician observes aggregate demand:

$$\int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i).$$

Deadweight loss for an individual i is a function of their expenditure function $e(p)$ and prices via:

$$dwl(p_i^s;\theta_i,\zeta_i) = e(p_i^s;\theta_i) - e(\bar{p};\theta_i) - (p_i^s - \bar{p})q(p_i^s;\theta_i,\zeta_i).$$

We are interested in aggregate deadweight loss:

$$DWL = \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i).$$

The problem of identification is to find conditions for which any joint distribution $F_{p^s,\theta,\zeta}$ of (p^s,θ,ζ), as a function of observable variables $\bar{p}$ & τ, satisfying these conditions and such that aggregate demand is rationalized; that is, such that for any observed values of observable variables, any F satisfying

$$\int_{p_i^s,\theta_i\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta_i}(p_i^s,\theta_i,\zeta_i) = \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i), \quad (2.6)$$

also yields the same value for deadweight loss:

$$\int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) = DWL.$$

The main message of this section will be the failure of such a result obtaining. We show that there are at least two distributions satisfying 2.6 that yield different values of DWL. These two distributions also turn out to yield tight bounds to the possible values of DWL that are consistent with aggregate demand.

Finally, we impose regularity conditions throughout this section to rule out ill-defined integrals. Formally, we insist that the econometrician only consider distributions that satisfy the *integrability conditions*, described below.

Definition 1. *A distribution $F_{p^s,\theta,\zeta}$ satisfies the integrability conditions if:*

1. *q and dwl are integrable on any measurable set.*
2. *$q(p;\theta_i,z)$ is integrable on any subset of the support of θ for any $p>0$ and any z in the range of ζ.*

For instance, all distributions with a finite support of (p^s,θ,ζ) satisfy the above conditions.

2.3.1. Lower Bound on Deadweight Loss

Consider arbitrary $\bar{p}$, p^{NT}, and τ. For arbitrary $F_{p^s,\theta,\zeta}$ consistent with the data, we can choose a price $\hat{p}^s$ that could also rationalize the data if perceived homogeneously.

Proposition 1. *For any $F_{p^s,\theta,\zeta}$ that yields integrable aggregate demand, there exists $\hat{p}^s$ such that for some distribution $F'_{\theta,\zeta}$ such that $F'_\theta = F_\theta$:*

$$\int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) = \int_{\theta_i,\zeta_i} q(\hat{p}^s;\theta_i,\zeta_i)dF'_{\theta,\zeta}(\theta_i,\zeta_i).$$

We can always rationalize the data with a joint distribution of (p^s,θ) in which θ has marginal distribution F^*_θ, whereas $p^s=\hat{p}^s$ with probability one. We now show that such a joint distribution provides a generic underestimate to the possible values of deadweight loss.

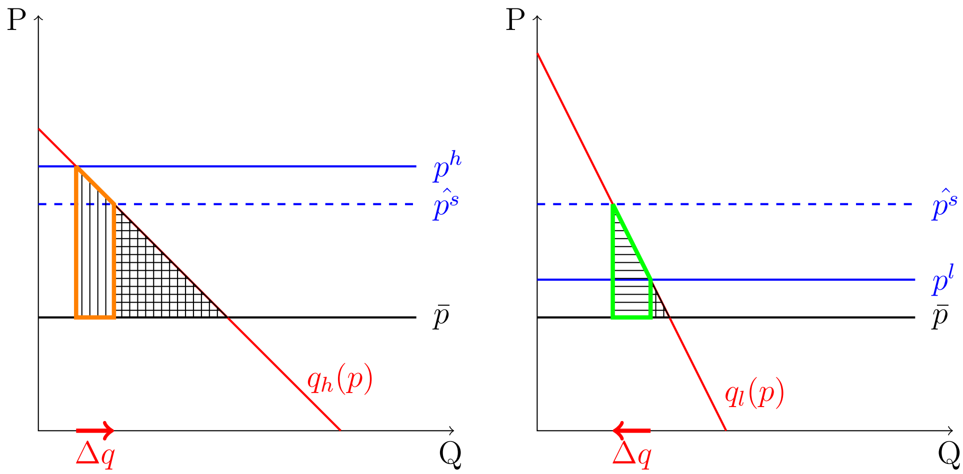
Theorem 1. *Consider any joint distributions $F_{p^s,\theta,\zeta}$ and $F_{\theta,\zeta}$ with corresponding value $\hat{p}^s$ such that:*

$$\int_{\theta,\zeta} q(\hat{p}^s;\theta_i,\zeta_i)dF_{\theta,\zeta}(\theta_i,\zeta_i) = \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i).$$

Then the following inequality obtains:

$$\int_{\theta_i,\zeta_i} dwl(\hat{p}^s;\theta_i,\zeta_i)dF_{\theta,\zeta}(\theta_i,\zeta_i) \leq \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i).$$

Intuitively, introducing heterogeneity in perceived prices can mute gains from trade. If someone with a higher marginal valuation for the good has a higher perceived price than someone with a lower marginal valuation, they could both gain by trading with each other after making their consumption decisions. If they could exchange with each other, the one who perceived the higher price could purchase some of the good from the other agent, making both agents better off. Thus, ruling out perceived price heterogeneity by assuming a homogeneous perceived price $\hat{p}^s$ eliminates the possibility of an allocative inefficiency. Figure 2.3 offers graphical intuition for the lower bound.



(a) High perceived price

(b) Low perceived price

Figure (2.3) A graphical illustration of Theorem 1. When one picks $\hat{p}^s$ as to make the change in demand equal for the consumer in (a) and in (b), the decrease in dwl for the consumer in (a) (orange) must be at least as large as the increase in dwl for the consumer in (b) (green).

Theorem 1 points out that for any distribution that rationalizes the data, i.e. that explains the observed aggregate demand, one can alternatively rationalize the data with a homogeneous perceived price that yields (weakly) less deadweight loss. From this, we can reach two conclusions. One, we generally cannot identify deadweight loss because we could always alternatively rationalize the data with a homogeneous perceived price.[11] This holds even if we already knew the distribution of preference types F_θ^*. Two, if there is a minimum value of deadweight loss that is consistent with the data, that value of deadweight loss comes from a distribution with no heterogeneity in tax salience.

2.3.2. Upper bound on deadweight loss

The upper bound comes from an assumption on the limits to tax salience:

Assumption 1. *There is some value $\bar{m} \geq 0$ such that p^s has support known to be contained in $\mathcal{P} \equiv [\bar{p}, \bar{p} + \bar{m}\tau]$.*

This assumption says that agents must perceive a non-negative tax τ^s no greater than fraction $\bar{m}$ of the true tax.[12] The econometrician is allowed to assume an arbitrarily large $\bar{m}$, but the gain in robustness will likely come at the expense of precision. For instance, setting $\bar{m} = 1$ would be to assume that agents never over-react to a tax rate. Imposing that $\tau^s \geq 0$ with probability one already ensures that deadweight loss is no greater than the original consumer surplus.[13] But the interval restriction implies any distribution yields no more deadweight loss than a distribution with "binary" perceived prices, i.e. where p^s can only take on values in $\{\bar{p}, \bar{p} + \bar{m}\tau\} \equiv \partial\mathcal{P}$.

The gist of the upper bound of deadweight loss is that, for any data-generating process that rationalizes observed aggregate demand, there is another data-generating process that also rationalizes observed demand, but which would yield at least as much deadweight loss. This alternative explanation of the observed demand insists that all agents pay either zero or maximal attention.

Before formally stating our main result, we demonstrate how one can always

[11] This claim holds generically, but would not hold, for instance, if there was no heterogeneity in tax salience.

[12] This description implicitly assumes that $\tau > 0$.

[13] Recall that deadweight loss equals its calculation as if taxes actually satisfied $\tau_i = p_i^s - \bar{p}$, so excess burden cannot exceed the original consumer surplus for any agent. One can show that if τ^s has support on negative values, then it's possible to have total deadweight loss substantially greater than the original total consumer surplus.

pick such a distribution of attention to rationalize observed demand, for any underlying (known) distribution of preferences. Then, we state the main result, theorem 2, providing some intuition for why such a distribution would yield a weakly higher deadweight loss. Because our model of choice may result in several choices given the same sticker prices and taxes, sometimes there can be multiple equilibria with the same level of aggregate demand, but different consequences for welfare. We briefly discuss appendix theorem 3, which deals deals with such cases.

Consider any $F_{p^s,\theta,\zeta}$ that rationalizes the data, and such that:

$$\lim_{m \to \bar{m}^-} F_{p^s}(\bar{p} + m\tau) - F_{p^s}(\bar{p}) > 0.$$

In words, the distribution assumes some positive mass of agents pay neither zero nor maximal attention, i.e. $m \in (\bar{p}, \bar{p} + \bar{m}\tau) \equiv int(\mathcal{P})$. Pick $\tilde{p}^s \in int(\mathcal{P})$ and a corresponding $p^b(p_i^s) \equiv \bar{p} + \mathbb{I}(p_i^s > \tilde{p}^s)\bar{m}\tau$ such that:

$$\begin{aligned}\int_{p_i^s \in int(\mathcal{P}),\theta_i} q(p^b(p_i^s);\theta_i,l)dF_{p^s,\theta}(p_i^s,\theta_i) &\leq \int_{p_i^s \in int(\mathcal{P}),\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \\ &\leq \int_{p_i^s \in int(\mathcal{P}),\theta_i} q(p^b(p_i^s);\theta_i,h)dF_{p^s,\theta}(p_i^s,\theta_i).\end{aligned}$$

In words, for any distribution that puts mass on $int(\mathcal{P})$, we pick a value $\tilde{p}^s$ that acts as a divide: those below it get assigned to a group that does not perceive the tax at all, while those above it get assigned to a group that perceives it "maximally". Since demand is monotonic in p, and given our definitions of l and h, one can always pick $\tilde{p}^s$ such that the above inequalities hold weakly. Thus, one can always find $\lambda \in [0,1]$ such that:

$$\begin{aligned}\lambda \int_{p_i^s \in int\mathcal{P},\theta_i} & q(p^b(p_i^s);\theta_i,h)dF_{p^s,\theta}(p_i^s,\theta_i) \\ &+ (1-\lambda)\int_{p_i^s \in int\mathcal{P},\theta_i} q(p^b(p_i^s);\theta_i,l)dF_{p^s,\theta}(p_i^s,\theta_i) \\ &= \int_{p_i^s \in int\mathcal{P},\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i).\end{aligned} \tag{2.7}$$

Equation 2.7 implies that we can always pick a threshold $\tilde{p}^s$ that rationalizes demand, so long as we randomly assign a fraction λ of consumers with $m \in int(\mathcal{P})$ to the tie-breaking parameter $\zeta = h$, and the remaining $1-\lambda$ to $\zeta = l$. But in turn, this implies that we have found binary distribution of p^s and ζ that rationalizes aggregate demand.

Let us call such distribution $F''_{p^s,\theta,\zeta}$. This new distribution has the same marginal distribution of preferences, $F_\theta = F''_\theta$. If a consumer perceived a price in the boundary region in the original distribution $F_{p^s,\theta,\zeta}$, then so will she in the new distribution: $F''_{p^s,\theta,\zeta|p^s\in\partial\mathcal{P}} = F_{p^s,\theta,\zeta|p^s\in\partial\mathcal{P}}$. As for those consumers who perceived a price in the interior, we propose to split them up in a manner akin to what we just did in equation 2.7: $F''_{p^b(p^s),\theta} = F_{p^s,\theta}$, and ζ is assigned randomly as we described above. One can quickly confirm that this distribution rationalizes the same demand as the original distribution $F_{p^s,\theta,\zeta}$:

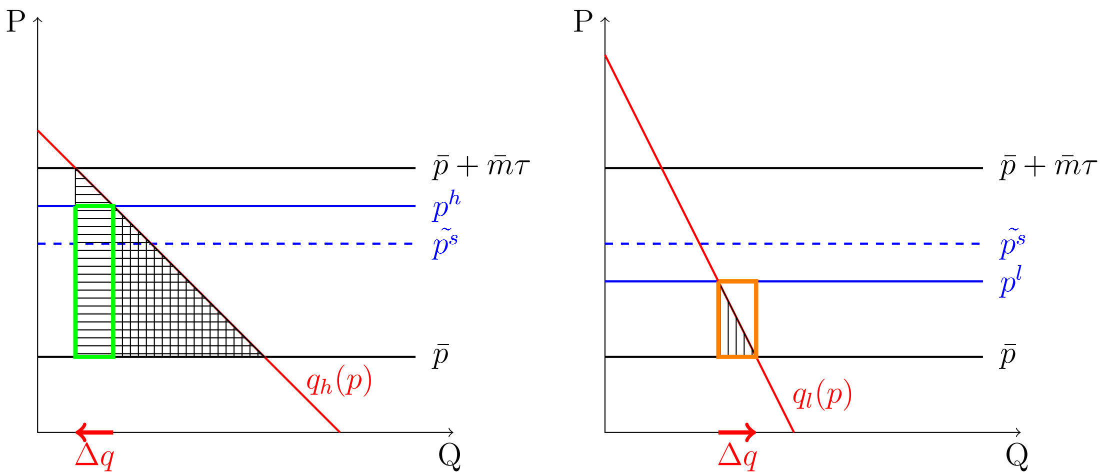
$$\begin{aligned}
\int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) = &\int_{p_i^s\in int(\mathcal{P}),\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\\
&+ \int_{p_i^s\in\partial\mathcal{P},\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\\
&= \int_{p_i^s\in int(\mathcal{P}),\theta_i} [\lambda q(p^b(p_i^s);\theta_i,h)\\
&+ (1-\lambda)q(p^b(p_i^s),\theta_i,l)]dF_{p^s,\theta}(p_i^s,\theta_i)\\
&+ \int_{p_i^s\in\partial\mathcal{P},\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\\
&= \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i).
\end{aligned}$$

Furthermore, such a distribution provides a generically larger value of deadweight loss than does $F_{p^s,\theta,\zeta}$.

Theorem 2. *Under assumption 1, for any $F_{p^s,\theta,\zeta}$ and any corresponding $F''_{p^s,\theta,\zeta}$ as described above:*

$$\int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \geq \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i).$$

We can obtain intuition in two ways. One is to note that the method of forcing binary perceived prices increases heterogeneity of perceived prices compared to $F_{p^s,\theta,\zeta}$. Another is by considering the case where $\bar{m} = 1$ and F^*_θ is known to be degenerate, so that all agents have the same preferences. For a given aggregate demand, deadweight loss is maximized under these preferences when some perceive price $p^s_i = \bar{p}$, while others correctly perceived the true tax rate $p^s_i = \bar{p} + \tau$. This is because for each individual agent, deadweight loss is convex in the perceived price. Hence, for a given aggregate demand, *aggregate* deadweight loss will be highest when it is as high as possible for some – namely, those who fully perceive the tax – while it is null for everybody else – as those who don't perceive the tax at all are effectively subject to a lump-sum tax. We provide a graphical illustration of this argument in figure 2.4.



(a) High perceived price (b) Low perceived price

Figure (2.4) A graphical illustration of Theorem 2. The watershed price $\tilde{p}^s$ is chosen to make the change in demand equal for the consumer in (a) and in (b). As long as we are dealing with weakly decreasing demand functions, the increase in deadweight loss for (a) is at least as big as the green box, while the decrease in deadweight loss for (b) is at most as big as the orange box. By assigning a perceived price of $\bar{p} + \bar{m}\tau$ to the consumer in (a) and $\bar{p}$ to the consumer in (b), we have increased aggregate deadweight loss holding aggregate demand constant.

Theorem 2 illustrates that for any distribution of (p^s, θ) that rationalizes the data, we can alternatively rationalize the data with a distribution with support for p^s_i on $\{\bar{p}, \bar{p} + \bar{m}\tau\}$ that yields (weakly) greater deadweight loss. Again, we see that identification of deadweight loss is not generally possible even if we knew the distribution of F^*_θ, as different marginal distributions of p^s and ζ could have different implications for deadweight loss. Also, any upper bound to the possible values of deadweight loss must be generated from a distribution with support of perceived prices on $\{\bar{p}, \bar{p} + \bar{m}\tau\}$.

However, not all distributions that have $p^s \in \partial\mathcal{P}$ with probability one yield the same value of deadweight loss, even when rationalizing the same data with the same distribution of preference types. The allocation of the good that yields the highest possible deadweight loss will also assign more consumption to agents with more convex demand curves. Theorem 3 in the online appendix spells out how to assign consumption of the good in the way it will "do the least good", and deals with cases where the tie-breaking parameter ζ is relevant, to get a general expression for the upper bound for deadweight loss consistent with aggregate demand and the distribution of preferences.

2.4. Linear Special Case

In this section, we discuss the special case in which q is known to be linear in $\bar{p}$ and τ (for fixed p^{NT}). We focus on this example both because of how frequently economists estimate linear models and because of its relationship to the second order approximation of deadweight loss.

We can also use the linear special case to better illustrate the general identification problem. In this subsection, we will no longer assume that the distribution of preference parameters is known; the econometrician will, as is usually the case, be able to identify preferences with exogenous price variation. As in TRJ (2018), we assume the distribution of preferences does not depend on taxes (or prices). Further, we permit the econometrician to assume that the distribution of tax salience does not change as sticker prices and taxes vary. This entirely rules out any sort of endogeneity between attention and taxes – which was driving non-identification in the binary case – and yet we will get the same non-identification result.

Naturally, no demand curve can be entirely linear, for the simple reason that agents cannot consume negative quantities. But in practice, one rarely gets the privilege of such rich variation in prices when doing empirical work. What we are implicitly assuming in this section is that the values of $(\bar{p}, \tau)$ that are considered are all such that $\bar{p} > 0$, and $q_i > 0$ for all consumers in the market, so that aggregate demand is also linear at those values. In other words, one can think of our work here as modeling linear demand conditional on buying at all prices under consideration.[14]

We might recall from section 2.2 that one can express a second order approximation to deadweight loss as a function of derivatives. If the choice function is linear in regressors $\bar{p}$ and τ, the second order approximation is an exact calculation of deadweight loss, and our results from the previous sections apply.

Formally, each preference type θ_i takes the form $\theta_i = (\beta_i, \epsilon_i) \in \mathbb{R}^2$.[15] To maintain linearity in regressors, we also assume that tax salience m is constant with respect to τ. The choice function q then takes the form:

$$q_i = \alpha + \beta_i p_i^s + \epsilon_i = \alpha + \beta_i[\bar{p} + m_i \tau] + \epsilon_i$$

We are suppressing the tie-breaking parameter ζ because in this linear example $\theta_i, m_i, \bar{p}$, and τ always uniquely determine consumption. We have the parameter α so that we can assume without loss of generality that $\mathbb{E}[\epsilon] = 0$.

Defining $\tilde{\beta}_i \equiv m_i \beta_i$ yields:

$$q_i = \alpha + \beta_i \bar{p} + \tilde{\beta}_i \tau + \epsilon_i, \tag{2.8}$$

[14]We thank an anonymous referee at the Journal of Public Economic Theory for helping us clarify our thinking on this matter.

[15]Agents have quasi-linear utility $u_i = \frac{q_i^2/2 - (\alpha+\epsilon_i)q_i}{\beth_i} + q_i^{NT}$. For a given p^{NT}, we define $\beta_i \equiv \frac{\beth_i}{p^{NT}}$, yielding utility representation $U_i = \frac{q_i^2/2 - (\alpha+\epsilon_i)q_i}{\beta_i} + p^{NT} q_i^{NT}$.

with corresponding deadweight loss per agent, from equation 2.4:

$$\begin{aligned} dwl_i &= \int_{\bar{p}}^{p^s} [\alpha + \beta_i p + \epsilon_i] dp - (p^s - \bar{p})[\alpha + \beta_i p^s + \epsilon_i] = \int_{\bar{p}}^{p^s} (p - p^s)\beta_i dp \\ &= \frac{1}{2}\Big[\frac{p^{s2} - \bar{p}^2}{2} - (p^s - \bar{p})p^s\Big]\beta_i = \frac{1}{2}(p^s - \bar{p})[(p^s + \bar{p}) - 2p^s]\beta_i = -\frac{1}{2}\tau^{s2}\beta_i \\ &= -\frac{1}{2}m_i^2\beta_i\tau^2. \end{aligned}$$

We assume that not just the distribution of preference parameters, but also the joint distribution of preference and salience parameters remains unaffected by the specific values of $\bar{p}$ and τ. The econometrician observes for various values of regressors:

$$\mathbb{E}[q|\bar{p},\tau] \equiv \int\limits_{\beta_i,\tilde{\beta}_i,\epsilon_i} [\alpha + \beta_i\bar{p} + \tilde{\beta}_i\tau + \epsilon_i] dF^*_{\beta,\tilde{\beta},\epsilon}(\beta_i,\tilde{\beta}_i,\epsilon_i) = \alpha + \mathbb{E}[\beta]\bar{p} + \mathbb{E}[\tilde{\beta}]\tau. \quad (2.9)$$

where $F^*_{\beta,\tilde{\beta},\epsilon}$ is the true distribution of $(\beta, m\beta, \epsilon)$. The challenge is to use the observed values of triplets $(\bar{p}, \tau,\ \mathbb{E}[q|\bar{p},\tau])$ to infer aggregate deadweight loss, which in this case is equivalent to its second order approximation around $\tau = 0$:

$$DWL = -\frac{1}{2}\int_{\beta_i,m_i} m_i^2\beta_i dF_{\beta,m}(\beta_i,m_i)\tau^2 = -\frac{1}{2}\mathbb{E}[m^2\beta]\tau^2 = -\frac{1}{2}\mathbb{E}[m\tilde{\beta}]\tau^2.$$

The only restriction that the econometrician imposes on the distribution of tax salience m is that the support of tax salience is contained within the interval $[0, \bar{m}]$. The econometrician can also use the Compensated Law of Demand as defined in appendix lemma 2, which shows that compensated demand is always weakly decreasing, so that $\mathbb{P}[\beta \leq 0] = 1$. In fact, we can permit the econometrician to know the entire distribution of $\theta = (\beta, \epsilon)$. It will not affect our results.} First, we can find a homogeneous perceived price that rationalizes the data for any τ. In particular, a linear regression of aggregate demand on sticker prices and taxes may permit identification of $\hat{\beta} \equiv \mathbb{E}[\beta]$ and $\hat{\tilde{\beta}} \equiv \mathbb{E}[\tilde{\beta}]$, respectively.[16] We define a

[16]Such identification requires exogenous and non-collinear variation in sticker prices and taxes. If the econometrician cannot identify these terms, so much the worse for identifying aggregate

measure of central tendency of tax salience:[17]

$$\hat{m} \equiv \frac{\hat{\tilde{\beta}}}{\hat{\beta}}.$$

Then the homogeneous perceived price that rationalizes the data is $\hat{p}^s = \bar{p} + \hat{m}\tau$. To see this, note that assuming all agents have tax salience $m_i = \hat{m}$ yields aggregate demand as in equation 2.9:

$$\begin{aligned}
\int_{\beta_i,\epsilon_i} [\alpha + \beta_i \hat{p}^s + \epsilon_i] dF^*_{\beta,\epsilon}(\beta_i, \epsilon_i) &= \alpha + \bar{p} \int_{\beta_i,\epsilon_i} \beta_i dF^*_\beta(\beta_i) + \hat{m}\tau \int_{\beta_i} \beta_i dF^*_\beta(\beta_i) \\
&= \alpha + \hat{\beta}\bar{p} + \hat{m}\hat{\beta}\tau \\
&= \alpha + \hat{\beta}\ \bar{p} + \hat{\tilde{\beta}}\ \tau.
\end{aligned}$$

Thus, the econometrician cannot rule out all agents perceiving the same price $\hat{p}^s$, and so cannot rule out $m_i = \hat{m}\ \forall i$. For tax τ, this would yield deadweight loss:

$$DWL_{low} = -\frac{1}{2}\hat{m}\hat{\tilde{\beta}}\tau^2.$$

By theorem 1, this is a lower bound for deadweight loss.

Alternatively, the econometrician cannot rule out the perceived tax τ^s having support in $\{0, \bar{m}\tau\}$. To see this, consider $\mathbb{P}(p^s = \bar{p} + \bar{m}\tau) = \frac{\hat{m}}{\bar{m}}$ and $\mathbb{P}(p^s = \bar{p}) = 1 - \frac{\hat{m}}{\bar{m}}$ independently of other parameters and regressors.[18] This will rationalize aggregate demand:

$$\begin{aligned}
\int_{\beta_i,\epsilon_i} \left[\alpha + \beta_i \bar{p}_i + \frac{\hat{m}}{\bar{m}} \beta_i \bar{m}\tau_i + \epsilon_i\right] dF_{\beta,\epsilon}(\beta_i, \tilde{\beta}_i, \epsilon_i) &= \alpha + \mathbb{E}[\beta]\bar{p} + \hat{m}\mathbb{E}[\beta]\tau \\
&= \alpha + \mathbb{E}[\beta]\bar{p} + \mathbb{E}[\tilde{\beta}]\tau.
\end{aligned}$$

deadweight loss.

[17]If $\hat{\beta} = 0$, then let $\hat{m} = 0$.

[18]In the true distribution, it must be that $\hat{m} \in [0, \bar{m}]$. Alternatively, one could consider checking whether $\hat{m} \in [0, \bar{m}]$ as a weak test of the null hypothesis that tax salience is bounded within that interval.

This yields deadweight loss for tax τ:

$$DWL_{high} = -\frac{1}{2}\frac{\hat{m}}{\bar{m}}\mathbb{E}[\beta]\bar{m}^2\tau^2 = -\frac{1}{2}\hat{m}\hat{\beta}\bar{m}\tau^2 = -\frac{1}{2}\bar{m}\hat{\tilde{\beta}}\tau^2.$$

For instance, if $\bar{m} = 1$, then the value of deadweight loss under a homogeneous perceived price is fraction $\hat{m}$ of the above calculation of deadweight loss.

Proceeding from theorem 2, we noted that there is a specific distribution of perceived prices on $\{\bar{p}, \bar{p} + \bar{m}\tau\}$ that maximizes deadweight loss. We describe that distribution in theorem 3, noting that it involves assigning high or low perceived prices based on the ratio of per-person deadweight loss to the change in consumption for that individual. But in this context:

$$\frac{dwl_i}{q_i(\bar{p}) - q_i(p^s)} = \frac{\tau^s}{2}.$$

Thus, the distribution of tax salience independent of all other parameters and regressors in which $\mathbb{P}(m = \bar{m}) = \frac{\hat{m}}{\bar{m}}$ and $\mathbb{P}(m = 0) = 1 - \frac{\hat{m}}{\bar{m}}$ maximizes deadweight loss. More generally, the econometrician cannot rule out this maximal value of deadweight loss so long as they cannot rule out the possibility of some distribution F with $F_\beta = F_\beta^*$ such that $supp(m) \in \{0, \bar{m}\}$ with:

$$\mathbb{P}_F(m = \bar{m})\mathbb{E}_F[\tilde{\beta}|m = \bar{m}] = \hat{m}\hat{\beta} = \hat{\tilde{\beta}}.$$

One can check that this distribution rationalizes the data,

$$\mathbb{E}[q|\bar{p}, \tau] = \alpha + \mathbb{E}[\beta]\bar{p} + \mathbb{E}_F[\tilde{\beta}]\tau = \alpha + \hat{\beta}\bar{p} + \mathbb{P}_F[m = \bar{m}]\mathbb{E}_F[\tilde{\beta}|m = \bar{m}]\tau = \alpha + \hat{\beta}\bar{p} + \hat{\tilde{\beta}}\tau,$$

and yields the maximal value of deadweight loss,

$$-\frac{1}{2}\mathbb{E}_F[m^2\beta]\tau^2 = -\frac{1}{2}\mathbb{P}_F[m = \bar{m}]\bar{m}\mathbb{E}_F[\tilde{\beta}|m = \bar{m}]\tau^2 = -\frac{1}{2}\bar{m}\hat{\tilde{\beta}}\tau^2 = DWL_{high}.$$

More intuitively, once one knows $\hat{\beta}$ and $\hat{\hat{\beta}}$, one can rationalize the aggregate data. Since the ratio of deadweight loss to the change in quantity is constant, the relationship between tax salience and preferences doesn't matter upon attaining the observed aggregate demand.

Finally, consider a distribution with $m \perp (\beta, \epsilon)$ with $supp(m) \subseteq \{0, \hat{m}, \bar{m}\}$, $\mathbb{P}(m = \hat{m}) = \lambda$ and $\mathbb{P}(m = \bar{m} | m \neq \hat{m}) = \frac{\hat{m}}{\bar{m}}$. Varying λ from zero to one yields:

$$DWL \in \Big[-\frac{1}{2}\hat{m}\hat{\hat{\beta}}\tau^2, -\frac{1}{2}\bar{m}\hat{\hat{\beta}}\tau^2 \Big].$$

We can conclude from this result that one cannot even identify a second order approximation of deadweight loss with aggregate data alone.[19] Imposing structure on preferences to facilitate identification of F^*_θ still only permits interval identification. Nonetheless, we can use aggregate data to obtain bounds, or at least $\hat{m}$, which gives us a sense of the uncertainty over the possible values of deadweight loss.

Besides its illustrative value, this linear framework also gives researchers a quick and easy way to compute bounds for the deadweight loss of non-salient taxes or fees in a variety of empirical contexts. In the online appendix, we apply these findings to the framework of CLK's study of aggregate beer consumption and Goldin and Homonoff's (2013) study of cigarette consumption. Details on the estimation procedures are provided in the online appendix. In the baseline specification of the CLK (2009) data, we estimate that $\hat{m} \approx 0.31$. This estimate suggests that even assuming that salience cannot exceed one, $\bar{m} = 1$, the upper bound of deadweight loss is about three times the lower bound. These estimates, however, all seem fairly imprecise. Across the two data-sets, there is no specification in which we can reject the null hypothesis that $\hat{m} = 0$, permitting the upper bound to be arbitrarily large in proportion to the lower bound.[20] Similarly, in most specifications we cannot reject that $\hat{m} = 1$, which would imply that upper bound and lower

[19] One can identify a first order approximation trivially: it is zero.

[20] Note that $\frac{DWL_{high}}{DWL_{low}} = \frac{\bar{m}}{\hat{m}}$, so that as $\hat{m}$ approaches zero from above, this ratio of upper to lower bounds blows up to infinity.

bound are identical. This underlying uncertainty is mirrored in previous work on tax salience – e.g., TRJ (2018) find that individual differences increase excess burden by at least 200% relative to the case of homogeneous attention – but our wide confidence intervals may be the result of the specific data sets we are using here – for example, Goldin and Homonoff (2013) often cannot reject that consumers do not react at all to sales taxes. Our procedure, however, seems so straightforward to carry out that it might turn out useful in future research on tax salience.

Finally, in both the setting of CLK (2009) and Goldin and Homonoff (2013), functional form assumptions seem to matter. The limitations of the linear setting would prompt us to undergo more sophisticated and less parametric exercises, but we are dissuaded by the fact that our statistical power is already very low.

2.5. Conclusion

In this chapter, we studied deadweight loss in a model where agents misperceive prices. We started by generalizing the theoretical results of CLK (2007, 2009) with an arbitrary closed choice set. Inattentiveness to taxes makes agents worse off while reducing deadweight loss by preventing agents from avoiding the tax.

As in the binary choice model of TRJ (2018), heterogeneous attention adds another layer of complexity to deadweight loss. In our general setting, we show that aggregate consumption can be consistent with a wide variety of co-distributions of attention and preferences, each with a different implication for deadweight loss. This is because it matters who gets the good and who doesn't: when prices are misperceived, there is no guarantee that agents who end up with some units of the good are the ones who value those units most. By minimizing and maximizing this allocative inefficiency, we show that deadweight loss can only vary between two extremes for any given aggregate demand. The lower bound holds generally, while the upper bound relies on the assumption that tax salience has support contained in a known non-negative interval.

Finally, we explore the special case in which demand is linear, which is of special interest due to its relationship to both the empirical literature and the second order approximation of deadweight loss. Our analysis shows that, while identification of deadweight loss under binary choice may be restated as an endogeneity problem, the same cannot be said regardless of the choice set. Indeed, when individual demand is linear, assuming independence of tax salience from taxes and prices does not change the interval of possible values of deadweight loss.

The linear model yields bounds for deadweight loss that one can easily compute from linear regression estimates. While this doesn't necessarily doom any future application in empirical work, our own applications of this method on the existing work of CLK (2009) and Goldin and Homonoff (2013) leave us without many answers about how tight these bounds might be. While some point estimates seem reasonable, they also can be imprecise and dependent on functional form specification.

Chapter 3. Risky Business: Policy Uncertainty, Valuation, and Investment

3.1. Introduction

Analyses of the effects of tax changes often assume that these changes are permanent, or that they will unfold in a deterministic way. While this does offer insight into policy recommendations, it poses a serious problem from a positive stand-point. Since firms can observe the political process, they form expectations about future policy, which will influence their present behavior. Indeed, tax uncertainty distorts firms' incentives by affecting both the timing of their investments and the risk profile they face. While previous work has been focused on the former, this chapter is concerned with the latter.

While previous work about policy uncertainty was concerned with beliefs about the distribution of future taxes, this chapter has a broader concern with the joint distribution of a firm's pre-tax returns and its tax obligation. We demonstrate that agents particularly care about the covariance of future taxes with future productivity. For instance, suppose a risk-neutral investor is considering an investment that costs \$3. The investment will either pay \$10 or \$0, each with equal probability. The government taxes the revenue from this investment at a rate of either 50% or 0%, each with equal probability. The investor's expected payoff clearly depends on when she will get taxed. If the investor believes that the tax kicks in only when the investment is successful, then her expected after-tax payoff will be $\frac{1}{2}10(1-0.5)+\frac{1}{2}0(1-0) = 2.5 < 3$, and she will not undertake the investment. If, on the other hand, the investor believes that the tax applies only when the investment is unsuccessful, then her expected after-tax payoff will be $\frac{1}{2}10(1-0)+\frac{1}{2}0(1-0.5)=5>3$, and she will undertake the investment. Note that in either case the marginal distributions of pre-tax payoffs and tax rates are exactly the same; what pushes the investor to invest or not invest is her belief about their correlation.

As we note in our theoretical model, what ultimately matters to a firm is the covariance between its marginal tax rate and the marginal product of its inputs. Interestingly, the same mechanism can arise when there is no uncertainty about tax policy, but the tax schedule is not linear. This point goes far back in the literature on the effects of income taxation: Domar and Musgrave (1944) pointed out that when the government does not rebate taxes on corporate losses, this creates a piecewise linear tax schedule with a kink at zero: up to zero the marginal tax rate is zero, after zero there is a positive marginal tax rate. Since the government punishes profits more than it compensates losses, this disincentivizes investment[1]. Our model generalizes this point. Whenever the tax schedule is convex, this will automatically generate a positive covariance between the marginal tax rate and productivity, which has the effect of depressing investment. Further, the same effect can arise even if the tax schedule is linear, but firms are uncertain about what the marginal tax rate will be.

This chapter investigates how policy uncertainty affects the risk profile faced by firms and how this affects their choices, both theoretically and empirically. In section 3.2, we enter into the details of the literature on tax changes, policy uncertainty, and investment, with an eye both on how people have thought about the effects of policy uncertainty in the past, and the empirical evidence they have gathered. Section 3.3 lays out our full theoretical model. As we mentioned, this can be done using a simple model where decisions are taken statically, though fac-

[1] While the main point of Domar and Musgrave's paper was how proportional income taxation affects risk-taking by individuals, they also had a number of other points relating to how taxation interacts with uncertain payoffs.

ing uncertainty. Our model can accommodate any tax schedule, and allows for any aspect of said tax schedule to change stochastically. Then, we take our theoretical observations to data in section 3.4. We conclude in section 3.5.

3.2. Background

Public economists have long been thinking about how the prospect of tax changes affects economic choices in general, and investment in particular. We could separate these works in two branches: one concerned with "mean effects", and another concerned with "variance effects". The question in the first branch is: how does the future *expected* level of taxation affect a firm's investment decision? In this, one should include all models that assume perfect foresight. Auerbach and Hines (1988) is an early example of papers in this branch. In a perfect foresight model, they find that anticipated tax changes carry a lot of weight in explaining firm behavior, relative to a model of myopic expectations. Other examples include Poterba (1989) and Slemrod and Greimel (1999). The first paper shows how the yield spread between taxable and tax-exempt bonds responds to changes in expected individual income tax rates by performing event studies around the time of unexpected tax changes. The second goes a step further, and proxies changes in expectations with the odds of Steve Forbes winning the presidency in 1996[2], finding some evidence for a possible causal relationship. Rodrik (1991) also discusses how expectations of future policy interact with current decisions. He makes the point that in order for economic reform to have its desired impact, investors must believe it to be persistent. He models uncertainty as a probability of policy reversal, and shows that if previous policy and the reform are distant enough, even a small probability of reversal can act as a hefty tax on investment.

The second branch of this literature is concerned not only with the effects of changes in expectations, but also (and particularly) with the effects of uncertainty

[2] Steve Forbes was an outspoken advocate of a "flat-tax" that would have erased the differential tax treatment.

itself. While the common view on the topic is that policy uncertainty deters investment, early models of investment under uncertainty, such as Hartman (1972) and Abel (1983), show that, in fact, the optimal capital stock should actually increase in the face of uncertainty about output price. Pindyck (1988) makes the point that these results rely heavily on the assumption of *reversibility* of the investment. Hassett and Metcalf (1999) make a similar point in the context of uncertainty regarding investment tax credits: uncertainty affects the *timing* of investment. We may also include Skinner (1988), who notes that net-of-tax income may become less uncertain if the income tax rate is positively correlated with pre-tax income. In his model, he is thinking about the welfare effects of income uncertainty for risk-averse agents, while we focus on risk-neutral firms.

More recent literature has found similar theoretical insights. When investment is (partially or fully) irreversible, uncertainty has the effect of delaying investment. Bloom (2009) makes this point in a model where firms face costly adjustments both for capital and labor decisions. He shows that if one assumes an (s, S) solution to the firm's optimal investment and hiring policy, a mean-preserving spread in the distribution of taxation results in an increased area of inactivity for the firm.

Whereas these two branches of the literature focus on the mean and variance of future policy, this chapter focuses on the covariance between tax rates and productivity. We are unaware of any previous literature that considers the impact of this covariance. The paper that comes closest to discussing the mechanism we illustrate here may be Domar and Musgrave (1944). They discuss tax regimes that are non-linear functions of profit. We generalize their results in a way that allows us to consider cases where the tax schedule is itself uncertain, which they explicitly exclude.

To understand why it is so important to think about how policy uncertainty affects economic decisions, one should first realize that this represents a very concrete problem to all agents in an economy. To discipline thinking, let us try to think about *why* agents face uncertainty about policy. One reason might be that governments decide to condition their policies on uncertain states of the world. For instance, governments might decide to enact different policies during booms and recessions; if agents are uncertain about when booms and recessions will begin and end, then they will also be uncertain about economic policy. Further, this would automatically generate a non-zero covariance between uncertain fundamentals that affect firm profitability and policy. Another reason why policy uncertainty might arise is that the political processes can be inherently uncertain. In the absence of median-voter theorem type results, which fail, for example, whenever policy is more than one-dimensional, different candidates will run on different platforms. If all candidates are running to win, the outcome of the election will be inherently uncertain. Indeed, one could easily imagine a world where the ultimate outcome of the election depends on some underlying state of the world, unobservable to politicians, that also affects the productivity of firms; if people are more optimistic, say, they will consume more and increase firm profits, and vote for candidate A; if people are more pessimistic, they will consume less and decrease firm profits, and vote for candidate B.

At the root of this second source of policy uncertainty lies the inability of governments to commit. If a government could at any point tie the hands of all future governments to enact a certain policy, then the only possible source of uncertainty would arise from the government conditioning their policy on underlying fundamentals. But, as a matter of fact, governments have at best an imperfect ability to commit. Persson and Tabellini (1999) review many examples of political economy games where equilibria can change radically depending on whether the government has access to a commitment technology. Early analyses of optimal capital income taxation, such as Chamley (1986) and Judd (1985), all mention, almost offhand, that if lump-sum taxation were available, then the optimal policy would be to have a lump-sum levy on all wealth existing in period 0, leaving untouched the product of future investment. But such a policy would only be optimal in an environment where the government can credibly commit to never raising a capital levy again.

Kehoe (1989) takes this a step further, and shows that in fact the ability of governments to commit to cooperating with other governments is not necessarily optimal for society, even in a world where all governments are benevolent. While in this chapter we focus on the problem from the point of view of the firm, it would be complementary to our analysis to think about a government engaged in a game with firms might optimally decide to promise one thing and then do another when a certain state of the world is realized, or might decide to play a mixed strategy in equilibrium, leaving firms guessing as to the action it will actually take.

Our empirical work looks at how the stock market return on different firms, whose profitability is correlated in different ways with tax policy, varies upon a policy uncertainty shock. We decide to take this route rather than directly assessing the impact of policy uncertainty on investment for three reasons. First, as we explain in section 3.3, the mechanism linking the covariance of productivity and policy to the investment decision is exactly the same mechanism linking that covariance to firm profits. Third, our mechanism relates to policy and uncertainty shocks that are ex-ante uncertain to firms. In this sense we differ a lot from previous literature regarding productivity, which tends to treat it as an observable parameter, or at least a parameter known to firms or private individuals. As a result, we do not find it sufficient to look at how firms react to *realizations* of these uncertain parameters, because our theory has to do with their ex-ante *belief*, not with what will ultimately realize. Short of ripping people's heads open to observe their beliefs, the next best thing we can do is to look at scenarios, like stock markets, where expectations about future profitability are instantaneously incorporated into observable prices.

Our empirical strategy is indebted to papers that use stock market evaluations to infer the effects of policy, such as Cutler (1988), or Friedman (2009), besides the aforementioned Poterba (1989) and Slemrod and Greimel (1999). Baker, Bloom, and Davis (2016) develop an index of policy uncertainty based on newspaper coverage, the number of expiring tax provisions, and disagreement amongst forecasters, which will be prominent in the empirical investigations of this chapter. They show that policy uncertainty tends to be correlated, among other things, to decreases in stock market returns. While Baker, Bloom and Davis (2016) take a more "reduced-form" approach, Handley and Limão (2017) have a more structural bend. In the context of trade policy, they develop a structural model, and estimate it around China's 2001 WTO accession, which drastically reduced uncertainty regarding trade policy between China and the U.S.

3.3. Framework

Expectations about future policy will influence decisions of firms in the present by distorting the distribution of their after-tax returns. As we have seen in section 3.2, this can happen in several ways: through changes in the first moment of policy, through changes in the second moment of policy, or through changes in the mixed moment between policy and other stochastic determinants of profitability. In this section we focus on this last mechanism. The model presented here, however, can also be interpreted to study first-moment effects, and we are actively working on extensions that would allow us to incorporate second-moment effects, by adding adjustment costs. Having a fuller picture might help us to see how all these different ways in which policy uncertainty affects firm profits act, both in isolation and in concert with each other.

We start with a model of uncertainty in a simple cash-flow tax. A firm has to decide on its scale of operations, facing a certain rate of deduction today, but an uncertain tax rate on tomorrow's revenue, as well as uncertain revenue. This example is particularly informative because it shows how our mechanism can arise from policy uncertainty in an environment that would otherwise leave the firm's decision undistorted. However, we want to stress that this point holds very broadly, and under several definitions of policy uncertainty. Similarly to the set-up in Rodrik (1991), one can effectively think of "tax collected" as any policy provision that reduces (or enhances) the profitability of investment, as long as government policy depends in any way on an (uncertain) final output. As we show later in this section, our point holds under a largely arbitrary tax system, and the flexibility in our definition of what constitutes a "tax" is reflected in the flexibility of our theoretical framework.

Suppose that an agent had to pick how much to invest, x, in a risky project, which will produce an uncertain amount $\epsilon f(x)$, where $f(x)$ is a known, strictly increasing, and strictly concave function, while ϵ is stochastic. The agents fortunes unfold in two periods: in the first period, the agent decides how much to invest, and gets to deduct expenses, x, against the current tax rate, τ_0; in the second period, the agent's revenue is realized, $\epsilon f(x)$, and it is taxed at rate τ_1. Assuming for simplicity that all prices are equal to unity, and that the agent values equally profits in both periods, we can write down her problem as:

$$max_x \mathbb{E}\Big[(1-\tau_1)\epsilon f(x) - (1-\tau_0)x\Big]. \tag{3.1}$$

Let us start with the case where τ_0 and τ_1 are both known. Letting x^* be the argmax to the problem in equation 3.1, the firm's expected profits are:

$$\begin{aligned}\bar{\Pi} &= \mathbb{E}\Big[(1-\tau_1)\epsilon f(x^*) - (1-\tau_0)x^*\Big] \\ &= (1-\tau_1)\bar{\epsilon} f(x^*) - (1-\tau_0)x^* \\ &\equiv \Pi(\tau_0, \tau_1, \bar{\epsilon}, x^*) \end{aligned} \tag{3.2}$$

where $\bar{\epsilon} = \mathbb{E}[\epsilon]$. Thus, in a world were tax policy is known ex-ante, the two tax rates τ_0 and τ_1 simply reduce (or increase) the firm's costs and revenue, respectively. If $\tau_0 = \tau_1 = \tau$, then an increase in the tax rate simply scales down expected profits proportionally. Note that in this last case the firm's choice of inputs remains undistorted relative to the case of no taxes. In general, the firm's optimal choice of investment will be

$$x^* = \frac{1-\tau_0}{(1-\tau_1)\bar{\epsilon}}.$$

If $\tau_0 = \tau_1$, then the firm will pick $x^* = 1/\bar{\epsilon}$, or the efficient choice it would have made had all tax rates been identically zero.

Now let us suppose that while τ_0, the current tax rate, is known, τ_1, the future tax rate, is not. The key insight here is that now the firm will care not about the marginal distributions of ϵ and τ_1 respectively, but about their *joint* distributions. Expected profits are now:

$$\begin{aligned}\bar{\Pi} &= \mathbb{E}\Big[(1-\tau_1)\epsilon f(x^*) - (1-\tau_0)x^*\Big] \\ &= \mathbb{E}[(1-\tau_1)\epsilon]f(x^*)(1-\tau_0)x^* \\ &= \Pi(\tau_0, \bar{\tau}_1, \bar{\epsilon}, x^*) - f(x^*)\rho\sigma_{\tau_1}\sigma_\epsilon, \end{aligned} \tag{3.3}$$

where $\bar{\tau}_1 = \mathbb{E}[\tau_1]$, $\rho = corr(\tau_1, \epsilon)$, $\sigma_{\tau_1} = \sqrt{Var(\tau_1)}$, and $\sigma_\epsilon = \sqrt{Var(\epsilon)}$. As we can see, the tax still has the same proportional effect on expected revenues and costs as in 3.2, but with the addition of a new term involving the covariance between tomorrow's tax rate and productivity. For given expected tax rate $\bar{\tau}_1$, an increase in the variance $\sigma^2_{\tau_1}$ will reduce profits when ρ is positive, and will increase them when ρ is negative.

More formally, by the Envelope Theorem, since the derivative of expected profit with respect to investment is zero, we have that if we marginally change σ_{τ_1} holding

$(\tau_0, \bar{\tau}_1, \bar{\epsilon}, \rho, \sigma_\epsilon)$ constant,[3] we obtain:

$$\frac{\partial \bar{\Pi}}{\partial \sigma_\tau} = -f(x^*)\rho\sigma_\epsilon \tag{3.4}$$

Thus, when ρ is positive (negative), we should see that higher policy uncertainty, in the form of a higher σ_{τ_1}, decreases (increases) expected profits. This reasoning lies at the foundation of our empirical work. If the firm's stock market evaluation is given by its expected future stream of profits, it follows that unexpected increases in policy uncertainty should be associated with an increased evaluation if $\rho < 0$, and a decreased evaluation if $\rho > 0$.

This same mechanism is reflected in the choice of investment. Under a stochastic second-period tax rate, the optimal choice of investment is:

$$x^* = (f')^{-1}\left(\frac{1-\tau_0}{\bar{\epsilon}(1-\bar{\tau}_1) - \rho\sigma_{\tau_1}\sigma_\epsilon}\right). \tag{3.5}$$

If the tax rate and productivity were independent, so that $\rho = 0$, then one could induce the same efficient choice that the firm would make when all taxes are identically zero by setting $\tau_0 = \bar{\tau}_1$, similarly as in the case where all taxes are known ex-ante. However, if $\rho \neq 0$, then one should adjust the rate of first-period input deductibility to reflect this fact: $\tau_0 = \bar{\tau}_1 + \rho\sigma_{\tau_1}\sigma_\epsilon/\bar{\epsilon}$. Thus, if the tax rate is positively correlated with productivity, the firm will need to be compensated for input expenses *more* than the expected tax rate it will face on revenues. If instead the tax rate is negatively correlated with productivity the opposite will be true, since the firm knows that the tax rate will tend to be lower than average in states of the world where the investment reveals to be more productive. When the covariance is positive (negative), the firm will face a lower (higher) expected profit, the

[3] It might be unclear what we mean by changing a standard deviation while holding other parameters of the joint distribution constant. More formally, one can write $\tau_1 = \bar{\tau}_1 + z$, where z is a random variable with mean zero and variance σ_τ^2. Then we could perform a transformation by scaling z by some constant $c > 0$, $\tau = \bar{\tau} + cz$, which would increase the variance of τ and the covariance of τ and ϵ while leaving the mean tax and the correlation coefficient the same. In that case, equation 3.4 states the partial derivative of expected profit with respect to c when $\sigma_\tau = 1$.

covariance acts as a tax (subsidy), and so the firm is less (more) willing to invest.

As noted in the introduction, this reasoning extends beyond political uncertainty about future policy. Whenever a tax system is progressive,[4] marginal tax rates will automatically increase whenever a business venture is more successful, thereby inducing positive correlation between the tax rate and the productivity of inputs. In this case, an equal increase of tax rates at all corporate income levels would discourage input use. Conversely, if the tax system was regressive, an equal increase of tax rates at all corporate income levels would encourage input use.[5] To clarify ideas, let us consider a more general model.

A firm has to decide how much to invest, x, on a risky project. Pre-tax profits are:

$$\pi(x, \epsilon) = \epsilon f(x) - x,$$

where ϵ is an ex-ante uncertain Hicks-neutral productivity shock. The firm faces a tax $T(\epsilon f(x), x, \epsilon)$, and can vary depending on realized revenue $\epsilon f(x)$, as well as input costs x, and the underlying productivity realization ϵ. The firm takes the decision before ϵ is realized, to maximize expected after-tax profits:

$$max_x \mathbb{E}[\pi(x, \epsilon) - T(\epsilon f(x), x, \epsilon)]. \tag{3.6}$$

Again letting x^* denote the optimal choice of investment, and then taking a first-order Taylor expansion of $T(\cdot)$ in its first argument, we obtain that expected profits are now:

$$\bar{\Pi} \approx \pi(x^*, \bar{\epsilon}) - \bar{T}(x^*) - f(x^*)C(x^*), \tag{3.7}$$

[4]As, e.g., the federal corporate tax system used to be in the U.S. before the Tax Cuts and Jobs Act of December 2017, and as it currently is in many U.S. states.

[5]While this point is a natural implication of our model, the observation originates in Domar and Musgrave (1944). However, they do not consider uncertain tax rates.

where $\bar{T}(x) = \mathbb{E}[T(\bar{\epsilon}f(x), x, \epsilon)]$, and $C(x) = cov(T_1(\epsilon f(x), x, \epsilon), \epsilon)$. For the sake of clarity, let us specify that we use subscripts on $T(\cdot)$ to indicate the argument with respect to which we are taking partial derivatives; thus T_1 indicates the partial derivative of $T(\cdot)$ with respect to its first argument. Again, we see a similar reasoning as in equation 3.3.

Once again, we have similar implications for investment. Taking the FOC of the problem in equation 3.6 gives us:

$$f'(x)\Big[\bar{\epsilon} - \bar{T}_1(x)\bar{\epsilon} - C(x)\Big] = 1 + \bar{T}_2(x)$$

Where $\bar{T}_1(x) = \mathbb{E}[T_1(\epsilon f(x), x, \epsilon)]$ is the expected marginal tax (or subsidy) on revenue, and $\bar{T}_2 = \mathbb{E}[T_2(\epsilon f(x), x, \epsilon)]$ is the expected marginal tax discount (or penalty) given for input use. All expectations are, of course, taken conditional on investment, as that is the control variable in the firm's problem. The firm's optimal choice of investment is now:

$$x^* = (f')^{-1}\left(\frac{1 + \bar{T}_2(x^*)}{\bar{\epsilon}(1 - \bar{T}_1(x^*)) - C(x^*))}\right). \tag{3.8}$$

Note that when there is no productivity uncertainty, or if productivity uncertainty is uncorrelated with the marginal tax rate, then $C(x) = 0$ for any choice of x, so that equation 3.8 yields $x^* = f'^{-1}(1 + \bar{T}_2(x)/\bar{\epsilon}(1 - \bar{T}_1(x)))$. Again, we have that firms are allowed to deduct expenses at the same marginal rate as they are taxed on revenue (that is, $T_1(\epsilon f(x), x, \epsilon) = -T_2(\epsilon f(x), x, \epsilon)$ for all (x, ϵ)), the firm will make the efficient choice that it would make in the absence of taxation. This echoes the long standing result that taxing profits does not distort firm decisions. However, in our case, as the covariance between the marginal tax rate and the productivity of investment becomes larger, since $f'^{-1}(\cdot)$ is a decreasing function, optimal investment declines.

There are two noteworthy cases in which the covariance of marginal tax rates and productivity distorts investment. The first, more well known case is where the tax revenue function $T(\cdot)$ is known ex-ante, so that it does not depend on ϵ conditional on realized profit, $T(\epsilon f(x), x, \epsilon) = T(\pi(x, \epsilon))$. In this case, the covariance may be non-zero because the tax revenue function T is non-linear, and so the marginal tax rate is not constant. This relates to a classic point originally made by Domar and Musgrave in 1944. When the government has a constant tax rate on positive firm profits, but does not remit payment to firms that realize net negative profit, then it is discouraging investment whenever firms fear they could make negative profit upon a sufficiently poor productivity realization. More generally, if $T(\cdot)$ is a convex function of profits, and if there is any uncertainty in ϵ, then $cov(T'(\pi(x, \epsilon))\epsilon)$ must be positive for any positive value of x.[6]

The less studied case where the covariance of the firm's marginal tax rate and its productivity comes into play is one where the tax rate is both uncertain and correlated with productivity. This is precisely the sort of case we consider at the beginning of the section. That example, although it presents an unrealistic tax system, gives us two important insights. First, policy uncertainty might affect firms in a way that we haven't really contemplated in the past, by distorting the risk profile they face with their investments. Even though there is no "timing" of the decision to speak of, as all decisions in the model are taken *before* the uncertainty is realized, firms change their decisions in the face of policy uncertainty whenever they believe that policy correlates with the pre-tax productivity of their investment. It should be noted, however, that by adopting a static model we are implicitly assuming complete irreversibility of investment. "Partial" irreversibility can be obtained by adding adjustment costs, say $\frac{1}{2}\xi(x - x(\tau, \epsilon))^2$ for some $\xi > 0$, which would preserve our main results[7]. Second, how firms react to uncertainty shocks in the tax depends on how their productivity is correlated with said tax. By observing how firms (or their expected profits) react to policy uncertainty shocks, then, we should be able to infer how their productivity is correlated to policy. This observation is at the foundation of our empirical application, which we dive into

[6] To be more precise, $C(x) \geq 0$, with strict equality assured if $T(\pi)$ is strictly convex.

[7] More formal results on this on their way.

next.

3.4. Empirics: a CAPM Approach to Policy Uncertainty

As we mention in section 3.3 our empirical methodology looks at stock market prices, which are tied to the expected profitability of a firm. As we have seen in our theoretical model, the expected profits of a firm will react differently to policy uncertainty shocks depending on how they are correlated with firm productivity. While all of our work so far has presumed a neutral attitude to risk, the distortion of the risk profile of investment will affect risk-loving or risk-averse agents, too. As the foundation of our empirical strategy is a capital asset pricing model that presumes agents are averse to risk, we formally develop a finance model in section 3.4.1. Section 3.4.2 takes a more practical look at the data and methodology we use. Finally, we present preliminary results and discuss future work in section 3.4.3.

3.4.1. An ICAPM Model of Policy Uncertainty

Consider the problem of $h = 1, \ldots, H$ infinitely lived agents who, every period, have to decide how much of their wealth W_{ht} to consume, c_{ht}, which gives them utility $u_h(c_{ht})$, and what share of their saved wealth to invest, $\pi_{ht} = [\pi_{h1t}, \ldots, \pi_{hNt}]$, on N assets. Investment opportunities depend on market risk, as well as policy risk, denoted by state variable P_t. Here we are going to treat P_t as a scalar representing the variance of a policy variable, but one could easily obtain the same results thinking of P_t as a vector of moments for the same variable. P_t is assumed to move according to a first-order Markov process. The Bellman equation for an individual agent is:

$$V_{ht}(P_t, W_{ht}) = max_{c_{ht},\pi_{ht}} \Big[u_h(c_{ht}) + \delta_h \mathbb{E}_t[V_{h,t+1}(P_{t+1}, (W_{ht} - c_{ht})\pi'_{ht} R_{t+1})] \Big| \mathbf{1}'\pi_{ht} = 1 \Big], \tag{3.9}$$

where subscript t indicates expectations taken conditional on information available at time t, for some discount factor $\delta_h \in (0, 1)$. As in Back (2010) (p. 194-195), the envelope condition and the fact that the ratio of marginal utilities is a one-period Stochastic Discount Factor, we have that

$$Z_{h,t+1} = \delta_h \frac{\partial V_{h,t+1}(P_{t+1}, W_{h,t+1})/\partial W_{h,t+1}}{\partial V_{h,t}(P_t, W_{h,t})/\partial W_{h,t}} \tag{3.10}$$

$$\equiv \frac{\delta_h V_{hw}(P_{t+1}, W_{h,t+1})}{V_{hw}(P_t, W_{h,t})} \tag{3.11}$$

Since $Z_{h,t+1}$ is a one-period SDF, we have that, by definition, for any pair (i, t):

$$\mathbb{E}_t[Z_{h,t+1} R_{i,t+1}] = 1. \tag{3.12}$$

When appropriately manipulated, this produces the relation

$$\mathbb{E}_t[R_{i,t+1}] = \frac{1}{\mathbb{E}_t[Z_{h,t+1}]} - \frac{1}{\mathbb{E}_t[Z_{h,t+1}]} cov_t(R_{i,t+1}, Z_{h,t+1}). \tag{3.13}$$

Taking a first order Taylor approximation of $V_{hw}(P_{t+1}, W_{h,t+1})$ around $(P_t, W_{h,t})$, and adding over all households, we obtain our (approximate) CAPM equation:

$$\mathbb{E}_t[R_{i,t+1}] \approx \zeta_t + \alpha_t cov_t(\Delta W_{t+1}, R_{i,t+1}) + \eta_t cov_t(\Delta P_{t+1}, R_{i,t+1}), \tag{3.14}$$

where,

$$\zeta_t = \alpha_t \sum_h \frac{-V_{hw}}{\delta_h V_{hww}}, \tag{3.15}$$

$$\eta_t = \alpha_t \sum_h \frac{V_{hwp}}{V_{hww}}, \tag{3.16}$$

$$\begin{aligned} \alpha_t &= \mathbb{E}_t\left[\sum_h \frac{-V_{hw}(P_{t+1}, W_{h,t+1})}{V_{hww}}\right]^{-1} \\ &\approx \sum_h \frac{V_{hw}}{V_{hww}}, \end{aligned} \tag{3.17}$$

where the argument of $V(\cdot)$ is $(P_t, W_{h,t})$ unless otherwise noted, and the last approximate equality holds as long as $\mathbb{E}_t[\Delta W_{t+1}] \approx 0$ and $\mathbb{E}_t[\Delta P_{t+1}] \approx 0$.

3.4.2. Data and Empirical Strategy

Our data comes from three different sources. First, we use Baker, Bloom and Davis's (2016) Economic Policy Uncertainty (EPU) index, which forms an index of policy uncertainty based on:

1. Newspaper coverage from 10 major newspapers containing policy-related words;
2. How many federal tax code provisions are set to expire in the next 10 years;
3. Dispersion in the forecasts of the CPI, federal expenditures, and state and local expenditures by individual forecasters in the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters.

The EPU index varies monthly and is available for over 30 years, from 1985 onwards. We might also be interested in data on government expenditure in proportion to GDP, which Baker, Bloom and Davis (2016) sometimes use in their regressions as a control for the first moment of government policy. Second, we use monthly returns from CRSP on the stock of publicly traded firms between 1985 and 2014. Third, we match these financial data with Compustat quarterly data about firm fundamentals, namely market capitalization, book value, sector, and foreign income. We restrict attention to firms for which data is available throughout our sample period, which runs from 1985 to 2014.

Conceptually, the data set consists of a set of monthly returns for $i = 1, \ldots, N$ assets over $t = 1, \ldots, T$ periods. Our method is heavily inspired by Fama and French (1993) and their (and others') subsequent work on empirical models of asset pricing. Our empirical model makes use of the Fama-French framework and adds another state variable, policy uncertainty. The idea here is that policy uncertainty poses a separate source of systemic risk which is not necessarily priced into the returns on the market portfolio. The idea here is that if we were looking at *all* savings opportunities of agents we would find a "naked" CAPM to hold, but since we are looking only at a subset of assets there will be other state variables of interest. We think of policy uncertainty as a state variable which in part governs what investment opportunities are available. We give a more formal derivation of an intertemporal-CAPM model in which policy uncertainty pops up as a state variable in a CAPM-style equation in section 3.4.1.

Our empirical model is:

$$\mathbb{E}[R_{it} - R_{ft}] = \beta_{im}\mathbb{E}[R_{mt} - R_{ft}] + \beta_{is}\mathbb{E}[SMB_t] + \beta_{ih}\mathbb{E}[HML_t] + \beta_{ip}\mathbb{E}[EPU_t], \tag{3.18}$$

where R_{it} is the return on portfolio i in period t; R_{ft} is the risk-free rate of return; SMB_t (Small Minus Big) and HML_t (High Minus Low) are the difference in returns between diversified portofolios of small and big firms (in terms of their market capitalization) and between diversified portfolios of high- and low-BE/ME ratio firms (Book Evaluation/Market Evaluation)[8]; EPU_t is the EPU index; and the β's are coefficients on the OLS regressions of $(R_{it} - R_{ft})$ on $(R_{mt} - R_{ft})$, SMB_t, HML_t, EPU_t. Because firm fundamentals like total assets (used to compute

[8]These factors are the ones used in Fama and French (1993) and many subsequent pieces that have been found to work particularly well.

market value and book value) are only available quarterly, quarters are our basic time unit here. This prompts us to estimate the following equation:

$$R_{it} - R_{ft} = \alpha_i + \beta_{im}(R_{mt} - R_{ft}) + \beta_{is}SMB_t + \beta_{ih}HML_t + \beta_{ip}EPU_t + u_{it}, \tag{3.19}$$

where u_{it} is a mean-zero i.i.d. error term.

While our methodology is heavily inspired by the financial literature on empirical tests of CAPM models, our intent is not to re-invent the CAPM. As a result, we find it of interest to run a "naked" version of the CAPM without SMB or HML, as well. This allows us to get a slightly bigger sample size[9], as well as to run the regression monthly. Our second regression estimates:

$$R_{it} - R_{ft} = \alpha_i + \beta_{im}(R_{mt} - R_{ft}) + \beta_{ip}EPU_t + u_{it}. \tag{3.20}$$

What we are ultimately interested in are the coefficients on EPU_t, for two reasons. First, we are interested in documenting which kinds of firm returns tend to do better or worse when policy uncertainty increases. As showcased in section 3.3, this should tell us about how their own productivity is correlated with policy. To study this, we cut the sample according to three different classification: by sector, by multinational status[10], and size (as measured by market capitalization). Second, finding significant coefficients on EPU_t would be a weak test that our mechanism is at work. Of course, we cannot rule out other confounding factors that might be correlated with policy uncertainty. However, finding that all stocks do not vary at all with changes in EPU_t would tell us that our mechanism is *not* at work, which, perhaps, would be interesting in its own right.

[9]Specifically, we use 690 firms rather than the 556 used in the Fama-French CAPM.
[10]More specifically, a dummy for whether the firm has positive foreign income.

Identifying Assumptions

It should be noted that we are making several assumptions here. First, we are assuming that there is one (correct) distribution of all economic and policy-related parameters, a belief shared by all agents in the economy. As a result of this, the changes in the volatility of policy shocks are assumed to be identical for savers operating in financial markets as they are for firm owners deciding on input use.

Second, we are effectively assuming that the correlation between policy and productivity does not change for any specific firm during the entire sample period. We see how this might be a problematic assumption that, as beliefs about the variance of policy are changing, beliefs about the correlation between policy and productivity are held constant. However, this might not be as crazy as it sounds if one buys that said correlation is determined by fundamental relations between the government and a given sector, whereas policy uncertainty is driven by more ephemeral swings in public opinion or current events.

Another way to address this assumption is that we are measuring the average parameter over a number of period-specific parameters in our entire sample period. Of course, we are also holding other moments of policy constant over the sample period. If, say, swings in policy uncertainty were to be correlated with swings in one particular direction of"mean" government policy, this would obfuscate our empirical results.

3.4.3. Results

As we mention in section 3.4.2, one of our main interests is in looking at how policy uncertainty β's vary by three classifications: by sector, by multinational status, and by size. In order to interpret these results, it might be useful to know that the standard deviation of EPU over our sample period was 32.6. Policy uncertainty jumps after big, unexpected events. For instance, after the 9/11 attacks the EPU

Test	Naked CAPM	Fama-French CAPM
All sectors identical	< 0.001	< 0.001
Local = Multinational	0.1415	0.1565
Small = Big	0.9625	< 0.001
All sectors identical (oil)	0.0406	< 0.001
All sectors identical (health)	< 0.001	< 0.001
All sectors identical (reduced)	< 0.001	< 0.001

Table (3.1) Tests of the null hypothesis that all β's are equal to each other in several specifications.

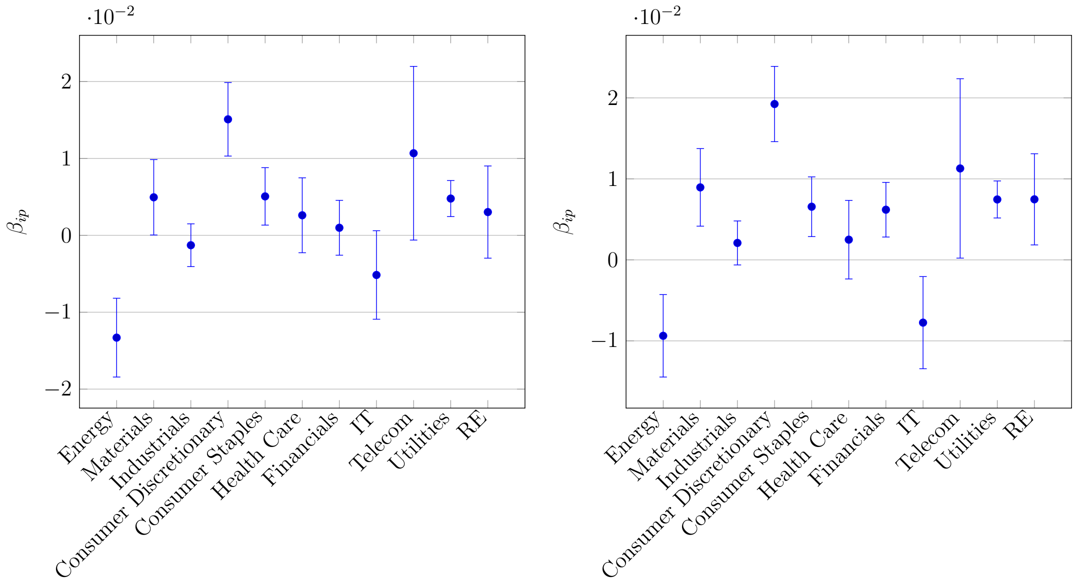
index increased by 103.77[11]; after the 2016 election, it increased by 76.86[12]. Figure C.1 in the data appendix graphs the EPU index during our sample period.

We can reject the null hypothesis that all individual firm β's are zero with more than 99.9% confidence. In all the categorizations below, we can also always reject the null hypothesis that all category-level β's are zero, again with confidence higher than 99.9%. We can also usually reject the null hypothesis that all β's are equal to each other. More details about this are provided in Table 3.1.

Figure 3.1 reports sector-level policy beta's for both models. Interestingly, it appears that three sectors in particular tend to have higher returns in periods of high policy uncertainty, indicating their productivity is negatively correlated with policies harming their profits: Consumer Discretionary, which includes products like cars, other durables, and services; Consumer Staples, which includes food, beverages and tobacco, and household products; and Utilities, which includes electric and gas utilities, as well as renewable electricity producers. On the other hand, the energy sector, composed mostly of oil and gas manufacturers, tends to have lower returns in periods of high policy uncertainty, indicating that policies that decrease their profits tend to be implemented during periods that would otherwise be very advantageous for them.

Figure 3.2 illustrates results for the classification by multinational status. For both the naked and the Fama-French CAPM we cannot reject the null that the effect is the same for multinational and domestic companies. For the naked

[11]From 84.29 in August 2001, to 188.06 in September 2001.
[12]From 92.52 in October 2016 to 169.39 in November 2016.

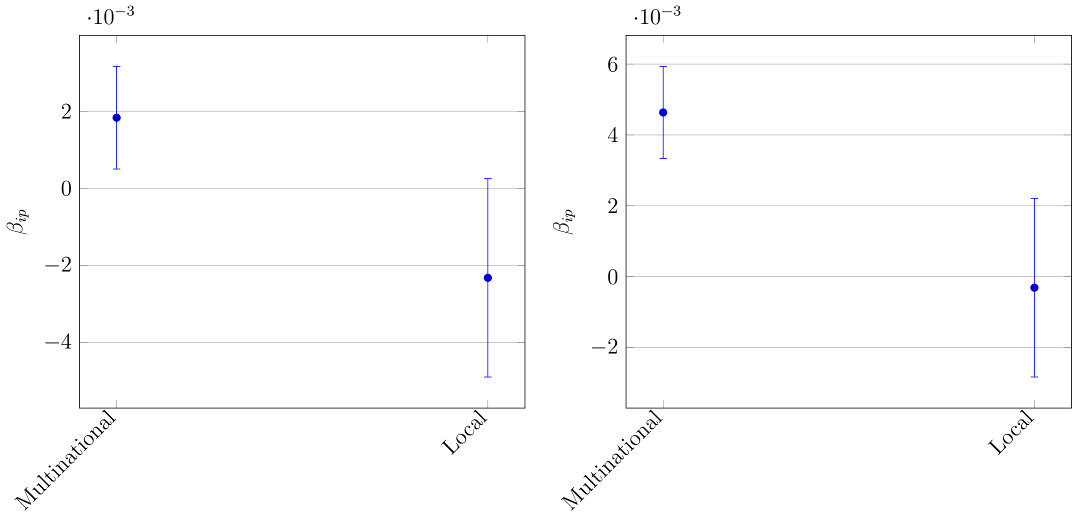


(a) Naked CAPM (b) Fama-French CAPM

Figure (3.1) Mean β_{ip}'s by sector (95% CI)

CAPM, the p-value for the test that the effect is the same is 0.1415, while for the Fama-French CAPM it is 0.1565. It should be noted that firms classified as "multinationals" here are the ones that had any foreign income in a given year. Given the heavy selection of "successfull" firms into our sample, the vast majority of our observations – over 90% – is constituted by multinational companies. While our estimates here are mostly statistically insignificant, they do go in the direction that one would expect, with policy uncertainty in the U.S. being generally worse for companies that earn all of their income domestically.

Finally, figure 3.3 represents mean beta's for small and big firms. "Small" here means a firm was in the smallest quintile of market capitalizations for a given year, while "big" means they were in the highest quintile. While we can reject the null hypothesis that EPU is uncorrelated with the return of firms of all sizes with very high confidence ($> 99.9\%$), we cannot always reject the hypothesis that the result is identical for small and big firms. While in the naked CAPM the p-value for the test that small and big firms have the same average β's is 0.9625, we can reject the null with very high confidence in the Fama-French model. This makes sense given that in the Fama-French CAPM equation we are already accounting for other common risk factors that arise from size differences, which might be con-



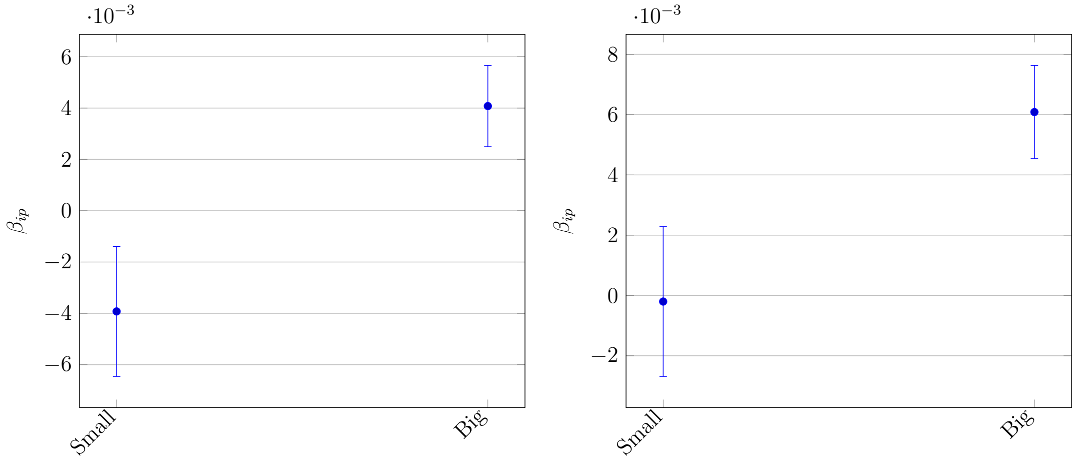
(a) Naked CAPM

(b) Fama-French CAPM

Figure (3.2) Mean β_{ip}'s by multinational status (95% CI). A company is classified as a "multinational" whenever it has positive foreign income.

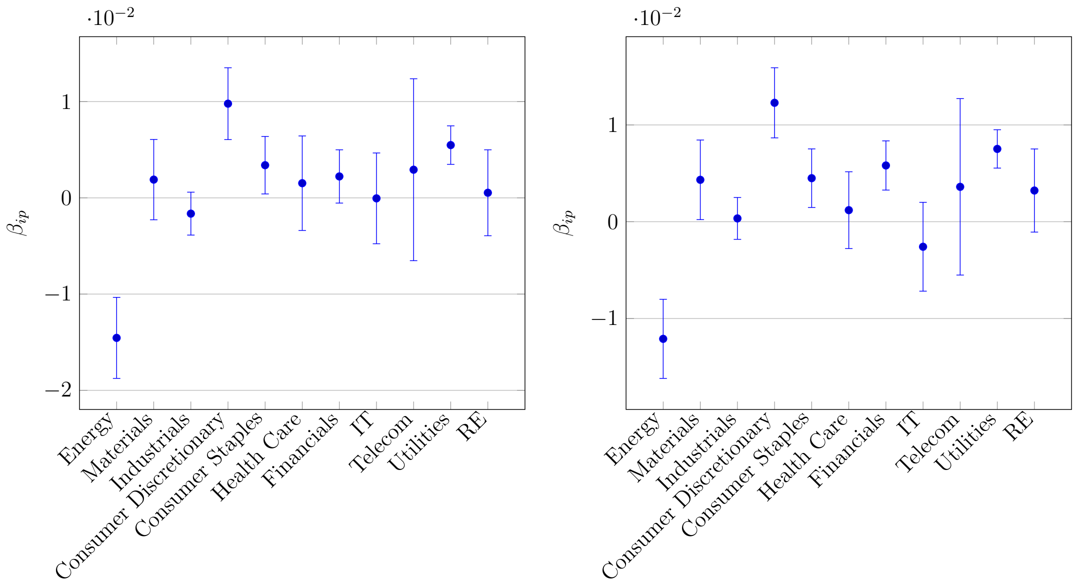
founding results in the naked CAPM. In the Fama-French specification, we can see that the returns of big firms tend to benefit from policy uncertainty, whereas the returns of small firms tend to suffer. This is consistent with our prior that big firms can pull more weight politically to obtain a policy environment that is more favourable to them.

Figure 3.2 illustrates results for the classification by multinational status. For both the naked and the Fama-French CAPM we cannot reject the null that the effect is the same for multinational and domestic companies. For the naked CAPM, the p-value for the test that the effect is the same is 0.1415, while for the Fama-French CAPM it is 0.1565. It should be noted that firms classified as "multinationals" here are the ones that had any foreign income in a given year. Given the heavy selection of "successfull" firms into our sample, the vast majority of our observations – over 90% – is constituted by multinational companies. While our estimates here are mostly statistically insignificant, they do go in the direction that one would expect, with policy uncertainty in the U.S. being generally worse for companies that earn all of their income domestically.



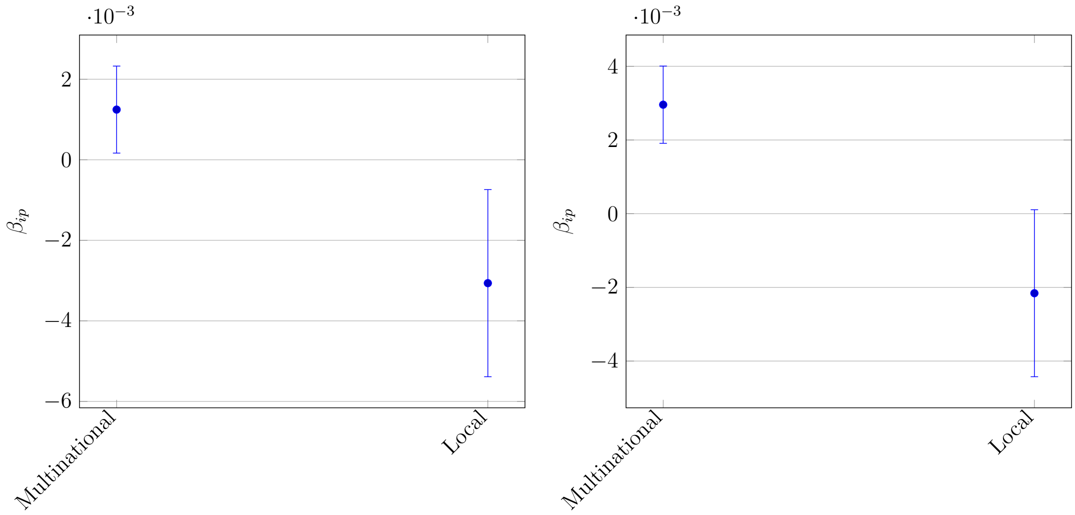
(a) Naked CAPM (b) Fama-French CAPM

Figure (3.3) Mean β_{ip}'s by size (95% CI). "Small" means in the lowest quintile of market cap, "big" means in the highest quintile.



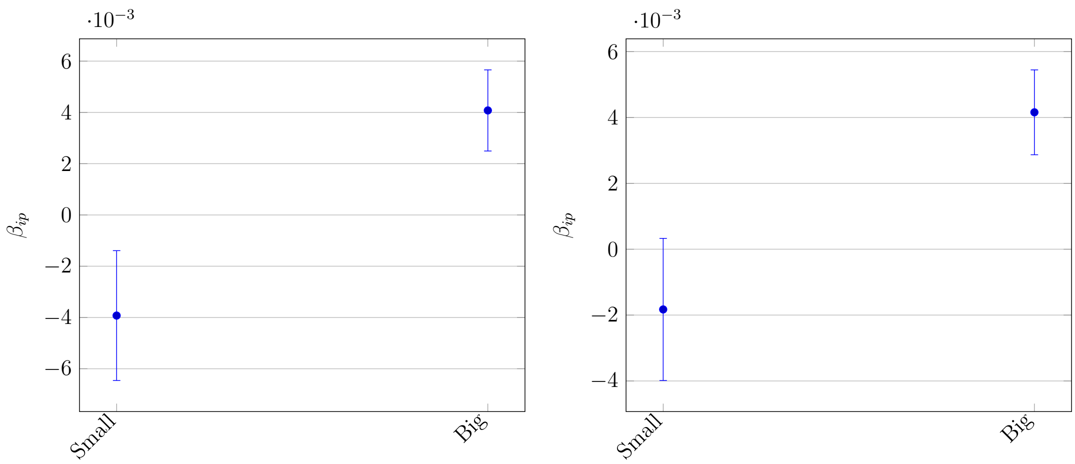
(a) Naked CAPM (b) Fama-French CAPM

Figure (3.4) News mean β_{ip}'s by sector (95% CI)



(a) Naked CAPM (b) Fama-French CAPM

Figure (3.5) News mean β_{ip}'s by multinational status (95% CI). A company is classified as a "multinational" whenever it has positive foreign income.



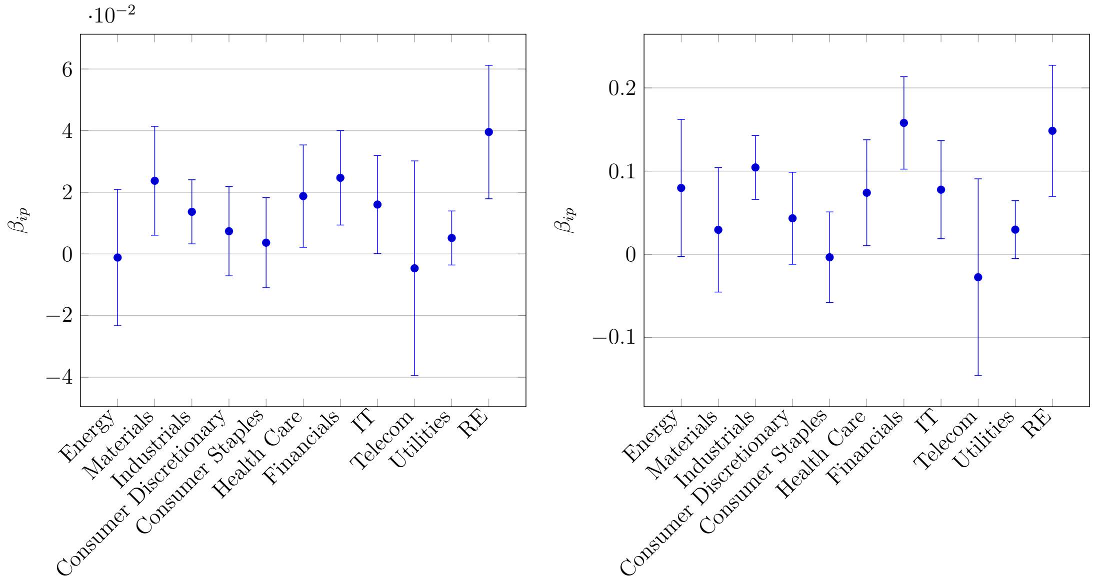
(a) Naked CAPM (b) Fama-French CAPM

Figure (3.6) News mean β_{ip}'s by size (95% CI). "Small" means in the lowest quintile of market cap, "big" means in the highest quintile.

3.4.4. Alternative Measures

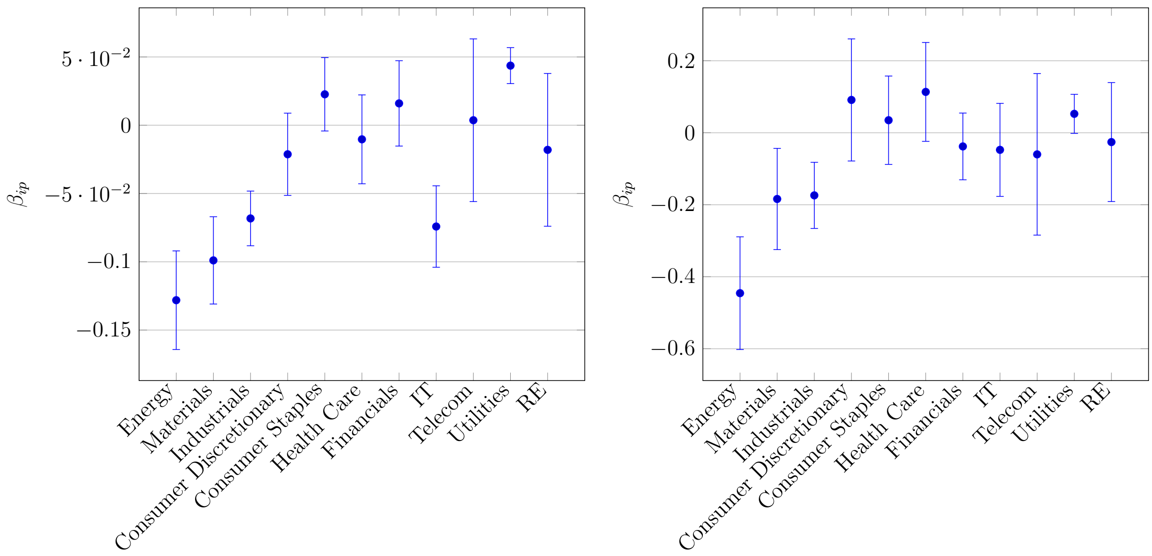
Because of concerns over how confounding factors might be weakening our analysis, we are thinking about new measures of policy uncertainty that might not present the same problems. One possibility is to use "Google Trends" data on search intensity for particular keywords. The idea is that this would allow us to get more specific measures of uncertainty in policy relevant to certain sectors. We attempt a first pass at this approach by looking at Google searches for "oil drilling" and "healthcare". Figures C.2 and C.2 depict graphical representations of these variables in the data appendix. Searches for oil drilling spike once in 2008, when the Bush administration lifted a ban on offshore drilling, and again during the summer of 2010, during the oil spill in the Gulf of Mexico; searches for healthcare tend to be high throughout the Obamacare debate.

Results are presented by sector in figures 3.7 and 3.8. What we'd expect to see here is that these two measures will be relevant for two particular sectors (respectively, Energy and Health Care), and that they preserve the sign of the coefficient found on EPU. Instead, we find quite the opposite. In both cases, the sector of interest seems to be insensitive to its associated measure of uncertainty. We believe this might be due to the fact that these measures are even more sensitive to confounding factors. The Gulf spill in 2010 certainly did not represent a "mean preserving spread" in the distribution of oil-drilling regulation.



(a) Naked CAPM (b) Fama-French CAPM

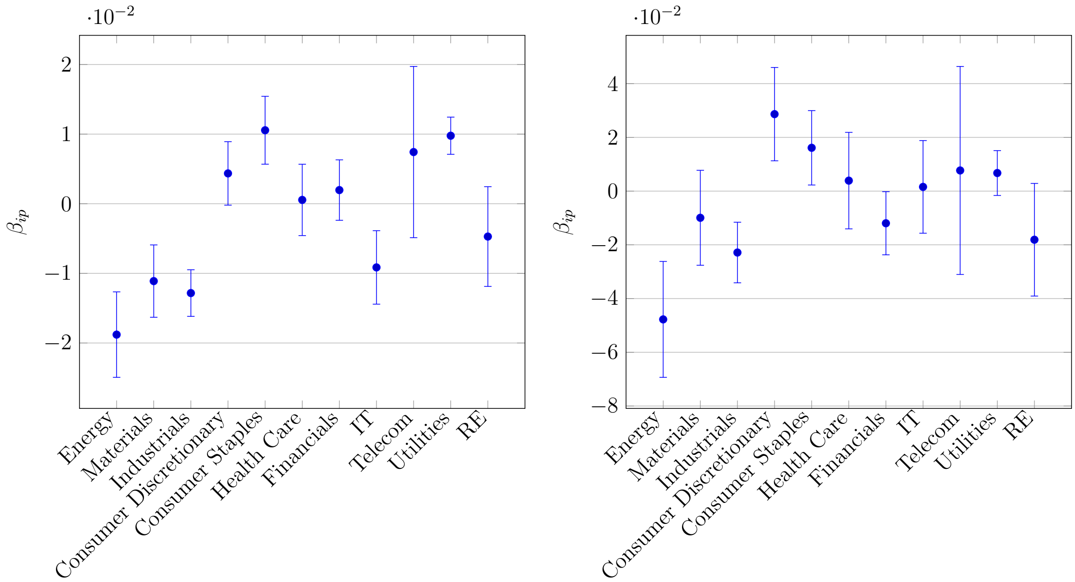
Figure (3.7) Google Trends: "oil drilling". Mean β_{ip}'s by sector (95% CI)



(a) Naked CAPM (b) Fama-French CAPM

Figure (3.8) Google Trends: "healthcare". Mean β_{ip}'s by sector (95% CI)

Another thing that might be affecting results is that Google Trends are only available starting in 2004, so the new analysis are effectively using a different sample period. To address this concern, we run the specification with EPU on the reduced sample. As we can see in figure 3.9, our results do not drastically change in the shorter sample when we use the EPU measure.



(a) Naked CAPM (b) Fama-French CAPM

Figure (3.9) GT reduced sample with EPU. Mean β_{ip}'s by sector (95% CI)

3.5. Conclusion

The question of how policy uncertainty affects investment has been raised in economics long ago. In this chapter, we show how the classical result that the deductibility of expenses leaves the choice of input undistorted is not valid whenever tax policy is correlated with other uncertain determinants of input productivity. Such a correlation might be induced by the tax system itself, or might arise due to how the political process reacts to changes in productivity.

Our theoretical model considers an arbitrary tax system and shows that whenever there is a covariance between the determinants of the firm pre-tax productivity and its marginal tax rate, this will act as an implicit tax on investment. This tax can discourage investment, if the covariance is positive; but it can also encourage investment (i.e., act as a subsidy), if the covariance is negative. The intuitive mechanism at work here is strikingly simple, and really can be boiled down to the numerical example we provided in the introduction: it is bad for expected profits if taxes are high precisely when times are bountiful.

In our empirical work, we use stock market data to see how returns on different firms react to jumps in policy uncertainty, and then use this relationship to infer the sign of the correlation between their productivity and tax policy. We do this not only to give suggestive evidence that our mechanism is at work, but also to document how different correlations are related to other firm characteristics.

We find that firms in the energy sector tend to be harmed by policy uncertainty, while firms in consumer-related activities tend to benefit from it. This is suggestive evidence that government policy tends to crack down on companies when times are good in the former, and when times are bad in the latter. One possible explanation for this could be that technological discoveries that make oil and gas companies more productive (like, say, fracking) tend to be accompanied by crackdowns on drilling regulations.

In conclusion, this chapter studies how a firm behaves when it faces uncertainty regarding both its tax bill and the productivity of its inputs. Given our preliminary results, we have reason to believe that this carries significant consequences for a number of classical results in the economics of business taxation.

Bibliography

Andrew B. Abel. Optimal Investment Under Uncertainty. *American Economic Review*, 73(1):228–233, 1983.

Ufuk Akcigit, John Grigsby, Tom Nicholas, and Stefanie Stantcheva. Taxation and Innovation in the 20th Century. *NBER Working Paper No. 24982*, 2018.

Hunt Allcott and Dmitry Taubinsky. Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review*, 105(8):2501–38, 2015.

Kenneth J. Arrow, Robert Forsythe, Michael Gorham, Robert Hahn, Robin Hanson, John O. Ledyard, Saul Levmore, Robert Litan, Paul Milgrom, Forrest D. Nelson, et al. The Promise of Prediction Markets. *Science*, 320:877, 2008.

Wiji Arulampalam, Michael P. Devereux, and Giorgia Maffini. The Direct Incidence of Corporate Income Tax on Wages. *European Economic Review*, 56(6): 1038–1054, 2012.

Alan J. Auerbach. Who Bears the Corporate Tax? A Review of What We Know. *Tax Policy and the Economy*, 20:1–40, 2006.

Alan J. Auerbach and James R. Hines. Investment Tax Incentives and Frequent Tax Reform. *American Economic Review Papers and Proceedings*, 78(2), 1988.

Alan J. Auerbach and Laurence J. Kotlikoff. Evaluating Fiscal Policy with a Dynamic Simulation Model. *American Economic Review*, 77(2):49–55, 1987.

David Autor, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. The Fall of the Labor Share and the Rise of Superstar Firms. *Working Paper, Massachussetts Institute of Technology*, 2017.

Kerry E. Back. *Asset Pricing and Portfolio Choice Theory.* Oxford University Press, 2010.

Scott R. Baker, Nicholas Bloom, and Steven J. Davis. Measuring Economic Policy Uncertainty. *Quarterly Journal of Economics*, 131(4):1593–1636, 2016.

B. Douglas Bernheim and Antonio Rangel. Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *The Quarterly Journal of Economics*, 124(1):51–104, 2009.

Tom Blake, Sarah Moshary, and Steve Tadelis. Price salience and product choice. *NBER Working Paper N. 25186*, 2017.

Olivier Blanchard, Christopher G. Collins, Mohammad R. Jahan-Parvar, Thomas Pellet, Beth Anne Wilson, et al. A Year of Rising Dangerously? The US Stock Market Performance in the Aftermath of the Presidential Election. *Journal of Policy Modeling*, 40(3):489–502, 2018.

Nicholas Bloom. The Impact of Uncertainty Shocks. *Econometrica*, 77(3):623–685, 2009.

Robin Boadway, Jean-Denis Garon, and Louis Perrault. Optimal mixed taxation, credit constraints, and the timing of income tax reporting. *Journal of Public Economic Theory*, 21:708–737, 2018.

Steven Bradley and Naomi Feldman. Hidden baggage: behavioral responses to changes in airline ticket tax disclosure. *Finance and Economics Discussion Series 2018-057. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2018.057*, 2018.

Christophe Chamley. Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives. *Econometrica*, pages 607–622, 1986.

Raj Chetty. The simple economics of salience and taxation. *NBER Working Paper N. 15246*, 2009.

Raj Chetty, Adam Looney, and Kory Kroft. Salience and taxation. *NBER Working Paper N. 13330*, 2007.

Raj Chetty, Adam Looney, and Kory Kroft. Salience and taxation: Theory and evidence. *The American Economic Review*, 99(4):1145–1177, 2009.

David M. Cutler. Tax Reform and the Stock Market: An Asset Price Approach. *American Economic Review*, 78(5):1107–1117, 1988.

Carl Davidson and Lawrence W. Martin. General Equilibrium Tax Incidence Under Imperfect Competition: A Quantity-Setting Supergame Analysis. *Journal of Political Economy*, 93(6):1212–1223, 1985.

Jan De Loecker and Jan Eeckhout. The Rise of Market Power and the Macroeconomic Implications. *NBER Working Paper No. 23687*, 2017.

Jan De Loecker, Jan Eeckhout, and Gabriel Unger. The Rise of Market Power and the Macroeconomic Implications. *Working Paper*, 2018.

Evsey Domar and Richard Musgrave. Proportional Income Taxation and Risk-Taking. *Quarterly Journal of Economics*, 58(3):388–422, 1944.

David Dorn, Lawrence F. Katz, Christina Patterson, John Van Reenen, et al. Concentrating on the Fall of the Labor Share. *American Economic Review*, 107 (5):180–85, 2017.

Eugene F. Fama and Kenneth R. French. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33(1):3–56, 1993.

R. Alison Felix and James R. Hines. Corporate Taxes and Union Wages in the United States. *NBER Working Paper No. 15263*, 2009.

Amy Finkelstein. Ez-tax: Tax salience and tax rates. *Quarterly Journal of Economics*, 124(3):969–1010, 2009.

Jeremy T. Fox. A note on nonparametric identification of distributions of random coefficients in multinomial choice models. *NBER Working Paper No. 23621*, 2017.

John N. Friedman. The Incidence of the Medicare Prescription Drug Benefit: Using Asset Prices to Assess its Impact on Drug Makers. *Working Paper*, 2009. URL `http://www.brown.edu/Departments/Economics/Faculty/John_Friedman/partd.pdf`.

Clemens Fuest, Andreas Peichl, and Sebastian Siegloch. Do Higher Corporate Taxes Reduce Wages? Micro Evidence from Germany. *American Economic Review*, 108(2):393–418, 2018.

Don Fullerton and Gilbert E. Metcalf. Tax Incidence. *Handbook of Public Economics*, 4:1787–1872, 2002.

Xavier Gabaix. A sparsity-based model of bounded rationality. *Quarterly Journal of Economics*, 129(4):1661–1710, 2014.

Fabio B. Gaertner, Jeffrey L. Hoopes, and Braden Williams. Making Only America Great? Non-US Market Reactions to U.S. Tax Reform. *Management Science*, forthcoming.

Kelly Sims Gallagher and Erich Muehlegger. Giving green to get green? incentives and consumer adoption of hybrid vehicle technology. *Journal of Environmental Economics and Management*, 61(1):1–15, 2011.

Jacob Goldin. Optimal tax salience. *Journal of Public Economics*, 131:115–123, 2015.

Jacob Goldin and Tatiana Homonoff. Smoke gets in your eyes: cigarette tax salience and regressivity. *American Economic Journal: Economic Policy*, 5(1): 302–336, 2013.

Roger H. Gordon and James R. Hines. International Taxation. *Handbook of public economics*, 4:1935–1995, 2002.

Alejandro Graziano, Kyle Handley, and Nuno Limão. Brexit Uncertainty and Trade Disintegration. *NBER Working Paper No. 25334*, 2018.

Kyle Handley and Nuno Limão. Policy Uncertainty, Trade, and Welfare: Theory and Evidence for China and the United States. *American Economic Review*, 107(9):2731–83, 2017.

Michelle Hanlon, Jeffrey L. Hoopes, and Joel Slemrod. Tax Reform Made Me Do It! *NBER Working Paper No. 25283*, 2018.

Arnold C. Harberger. The Incidence of the Corporation Income Tax. *Journal of Political economy*, 70(3):215–240, 1962.

Richard Hartman. The Effects of Price and Cost Uncertainty on Investment. *Journal of Economic Theory*, 5(2):258–266, 1972.

Kevin A. Hassett and Gilbert E. Metcalf. Investment with Uncertain Tax Policy: Does Random Tax Policy Discourage Investment? *Economic Journal*, 109(457): 372–393, 1999.

Kenneth L. Judd. Redistributive Taxation in a Simple Perfect Foresight Model. *Journal of Public Economics*, 28(1):59–83, 1985.

Patrick J. Kehoe. Policy Cooperation Among Benevolent Governments May Be Undesirable. *Review of Economic Studies*, 56(2):289–296, 1989.

Li Liu and Rosanne Altshuler. Measuring the Burden of the Corporate Income Tax Under Imperfect Competition. *National Tax Journal*, 66(1):215–238, 2013.

Charles F. Manski. Interpreting the Predictions of Prediction Markets. *Economics Letters*, 91(3):425–429, 2006.

Matthew A. Masten. Random coefficients on endogenous variables in simultaneous equations models. *The Review of Economic Studies*, 85(2):1193–1250, 2017.

Kyle C. Meng. Using a Free Permit Rule to Forecast the Marginal Abatement Cost of Proposed Climate Policy. *American Economic Review*, 107(3):748–84, 2017.

Andrea Morone, Francesco Nemore, and Simone Nuzzo. Experimental evidence on tax salience and tax incidence. *Journal of Public Economic Theory*, 20(4): 582–612, 2018.

Torsten Persson and Guido Tabellini. Political Economics and Macroeconomic Policy. *Handbook of Macroeconomics*, 1/C:1397–1482, 1999.

Robert S. Pindyck. Irreversible Investment, Capacity Choice, and the Value of the Firm. *American Economic Review*, 78:969–85, 1988.

James M. Poterba. Tax Reform and the Market for Tax-Exempt Debt. *NBER Working Paper No. 2900*, 1989.

Dani Rodrik. Policy Uncertainty and Private Investment in Developing Countries. *Journal of Development Economics*, 36(2):229–242, 1991.

Harvey S. Rosen. Tax illusion and the labor supply of married women. *The Review of Economics and Statistics*, 58(2):167–172, 1976.

Jonathan Skinner. The Welfare Cost of Uncertain Tax Policy. *Journal of Public Economics*, 37(2):129–145, 1988.

Joel Slemrod and Timothy Greimel. Did Steve Forbes Scare the U.S. Municipal Bond Market? *Journal of Public Economics*, 74(1):81–96, 1999.

Erik Snowberg, Justin Wolfers, and Eric Zitzewitz. How Prediction Markets Can Save Event Studies. *NBER Working Paper No. 16949*, 2011.

Juan Carlos Suárez Serrato and Owen Zidar. Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms. *American Economic Review*, 106(9):2582–2624, 2016.

Lawrence H. Summers. Capital Taxation and Accumulation in a Life Cycle Growth Model. *American Economic Review*, 71(4):533–544, 1981.

Dmitry Taubinsky and Alex Rees-Jones. Attention variation and welfare: theory and evidence from a tax salience experiment. *Review of Economic Studies*, 85 (4):2461–2496, 2018.

Alexander F. Wagner, Richard J. Zeckhauser, and Alexandre Ziegler. Company Stock Price Reactions to the 2016 Election Shock: Trump, Taxes, and Trade. *Journal of Financial Economics*, 130(2):428–451, 2018a.

Alexander F. Wagner, Richard J. Zeckhauser, and Alexandre Ziegler. Unequal Rewards to Firms: Stock Market Responses to the Trump Election and the 2017 Corporate Tax Reform. *AEA Papers and Proceedings*, 108:590–96, 2018b.

Matthias Weber and Arthur Schram. The non-equivalence of labour market taxes: A real-effort experiment. *The Economic Journal*, 127(604), 2016.

Justin Wolfers. Point Shaving: Corruption in NCAA Basketball. *American Economic Review*, 96(2):279–283, 2006.

Justin Wolfers and Eric Zitzewitz. Interpreting Prediction Market Prices as Probabilities. *NBER Working Paper No. 12200*, 2006.

APPENDIX A

Appendix for Chapter 1

A.1. Theoretical Foundations

In the theoretical framework that lays the ground for my empirical analysis, firm i's value originates from two different sources, V_{i1} and V_{i2}. In what follows, I give an example of a setting that would generate such a scenario.

Let us begin begin with a simple model of Cournot oligopoly. There are N identical firms indexed by i, each producing quantity q_i. They face an inverse aggregate demand $p(Q)$, where $Q = \sum_{i=1}^{N} q_i$. Each firm picks how much to produce to maximize profits, taking the output of the other firms, $Q_{-i} = \sum_{j \neq i} q_j$, as given:

$$\max_{q_i \geq 0} (1-\tau)(p(q_i + Q_{-i})q_i - c(q_i)),$$

where $c(q_i)$ is the cost function and $\tau \in (0,1)$ is the corporate tax rate. Let $q^*(N)$ denote the Cournot-Nash equilibrium of this game, and respectively define equilibrium pre-tax profits $\pi(N) \equiv p(Nq^*(N))q^*(N) - c(q^*(N))$. Now let us imagine an infinitely repeated version of the same game. Since $q^*(N)$ is the equilibrium of the stage-game, it must also be an equilibrium of the repeated game, so assume that firms simply play $q^*(N)$ every period. If future earnings are discounted at rate r, then the equilibrium value of a firm is, at any point in time, $\frac{(1-\tau)\pi(N)}{r}$.

Now imagine that each firm operates in two unrelated markets, each a copy of the one described above. In the first market, new firms can enter as long as they pay an entry cost c_e. The number of firms in the first market, N_1, must then satisfy the condition

$$\frac{(1-\tau)\pi(N_1)}{r} = c_e.$$

This condition implicitly defines the equilibrium number of firms, $N_1(\tau)$. Letting V_1 denote the value of the firm's operations in the first market, we therefore have that in equilibrium

$$V_1(\tau, N_1) = V_1\big(\tau, N_1(\tau)\big) = \frac{(1-\tau)\pi\big(N_1(\tau)\big)}{r} = c_e.$$

Since c_e is an exogenous constant, this makes it immediately evident that V_1 will not change upon a change in τ – in the language of section 3.3, we are in a case where $\alpha = 1$. We could also imagine a version of this same model where entry is limited. Say, for example, as long as $\frac{(1-\tau)\pi(N_1)}{r} > c_e$, one firm can enter every period. Depending on the length of the period and the discounting between periods, we will obtain positive values of α: the longer each entrant has to wait, and the more valued the present is relative to the future, the higher the implied value of α.

In the second market, instead, let us suppose that the number of firms is given and constant at N_2, say thanks to the presence of intellectual property rights or licensing. Then the value of firm operations in this market are

$$V_2(\tau, N_2) = \frac{(1-\tau)\pi(N_2)}{r},$$

and after-tax profits go down one-for-one with quasi-rents upon a marginal increase in the tax:

$$\frac{\partial V_2}{\partial \tau} = -\frac{\pi(N_2)}{r}.$$

A.2. Summary Statistics and Descriptive Graphs

Variable	Mean	S.D.	25th percentile	Median	75th percentile
Excess return (%)	0.09	3.53	-0.879	-0.001	0.969
Abnormal return (%)	0.009	3.49	-0.894	-0.019	0.813
Market Cap (billion USD)	6.54	26.33	0.15	0.76	3.188634
Lerner index (2016)	-10.74	235.7	-84.24	0.088	0.207
Market share (2016)	.018	.063	0.0001	0.001	0.009
Herfindahl index (2016)	.093	0.9	0.032	0.059	0.125
Proportion of foreign profits (2016)	0.137	2.499	0	0	0.06
$\bar{\tau}_i^{FOR}$ (2016)	0.072	1.34	0	0	0.01

Table (A.1) Summary statistics of firm fundamentals. Note that the Lerner Index as well as the proportion of foreign profits and the average foreign tax rate can be negative because they are fractions whose numerator (and in some cases whose denominator) can be negative in the data. For instance, about a third of firms on Compustat report operating losses. The Lerner index will be negative for those firms.

A.2.1. Data on the bet asking "Will the corporate tax rate be cut by the end of 2017?"

Figure (A.1) Average price of a "Buy Yes" contract.

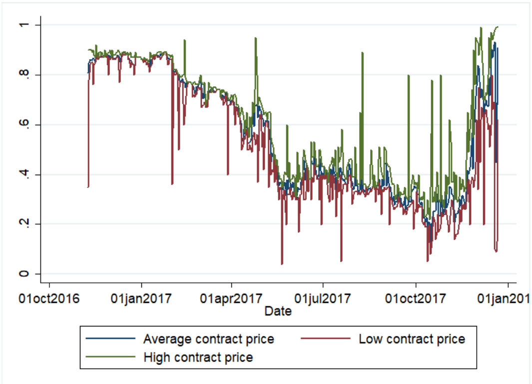


Figure (A.2) Average, low, and high prices.

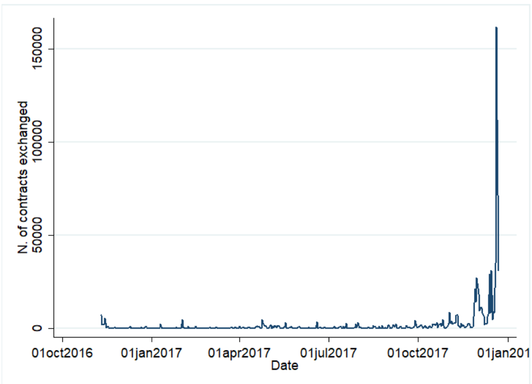


Figure (A.3) Number of contracts exchanged.

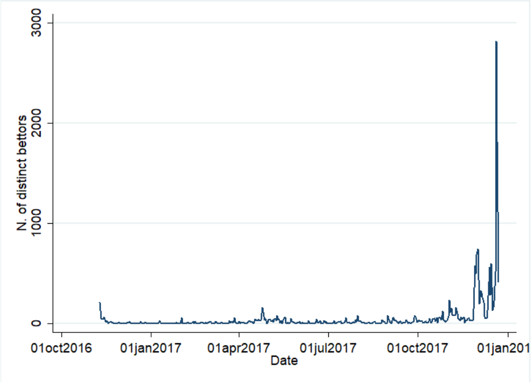


Figure (A.4) Number of individual bettors.

A.2.2. Data on the best asking "Will the corporate tax rate be cut by March 31, 2018?"

Figure (A.5) Average price of a "Buy Yes" contract.

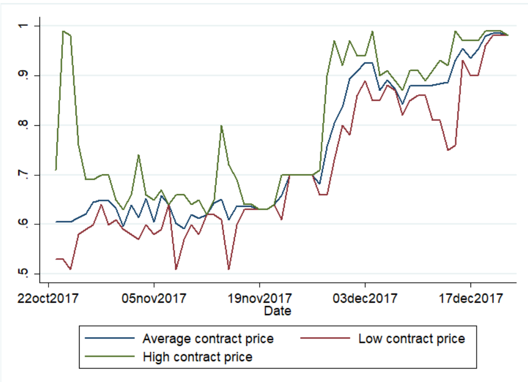


Figure (A.6) Average, low, and high prices.

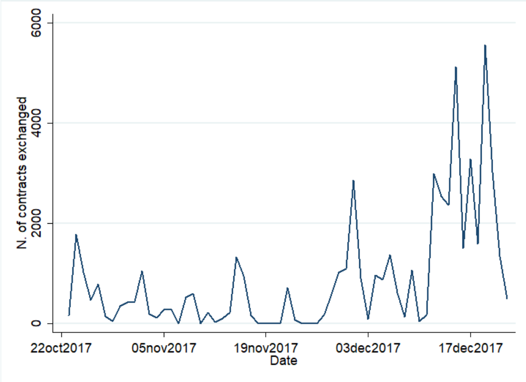


Figure (A.7) Number of contracts exchanged.

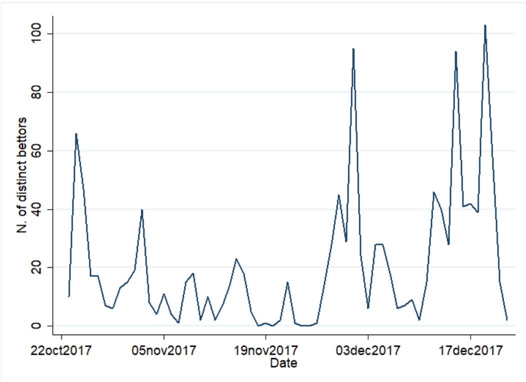


Figure (A.8) Number of individual bettors.

A.3. Date-by-date Results

	Dependent variable: abnormal returns		
	Coefficient estimate of:		
$\cdots \times \ldots$	Market Cap	HHI × s	Lerner
September 27	0.0032***	-0.0792	0.0006309***
	(0.00007)	(0.0567)	(.0000241)
November 2	0.0012***	0.2177***	0.0001389***
	(0.00007)	(0.0567)	(.0000241)
November 16	0.0013***	0.3791***	-0.0000397
	(0.00007)	(0.0567)	(.0000241)
December 2	0.0003***	3.3104***	-0.0000542**
	(0.00007)	(0.0566)	(.0000241)
December 15	0.0021***	0.9912***	0.0002194***
	(0.00007)	(0.0567)	(.0000241)
December 20	-0.0004***	0.5666***	0.0008203***
	(0.00007)	(0.0567)	(.0000241)
Time FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Foreign Profits Controls	Yes	Yes	Yes
N	296,437	1,274,778	1,166,039

Standard errors in parentheses

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

Table (A.2) Each column corresponds to an estimation of equation 1.3 with each one of our three measures of productivity and market power, disaggregated by date. Standard errors, in parentheses, are clustered by day.

A.4. Alternative Asset Pricing Models

A.4.1. Raw Excess Returns

In this case, there is no asset model. Simply, $R_{i,t} - R_t^f$ replaces $AR_{i,t}$ in specification 1.3.

	Dependent Variable: Excess Returns	
	Herfindahl	Lerner
Market share	-0.0593	
	(0.163)	
$HHI\times$ market share	0.0834	
	(0.260)	
$HHI\times$ market share $\times TCJA\ dates$	0.383	
	(0.588)	
Lerner index		0.000000781
		(0.0000233)
Lerner index $\times TCJA\ dates$		0.000283**
		(0.000133)
Proportion of foreign profits	-0.00123	-0.00143*
	(0.000808)	(0.000806)
Proportion of foreign profits $\times TCJA\ dates$	-0.00810	-0.00820
	(0.00636)	(0.00609)
Avg. for. tax rate	-0.000776	-0.000865
	(0.00166)	(0.00164)
Avg. for. tax rate $\times TCJA\ dates$	0.0208	0.0216
	(0.0145)	(0.0139)
N	1,274,778	1,166,039
Industry FE	Yes	Yes
Time FE	Yes	Yes

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

Table (A.3) Continuous diff-in-diff regressions of raw excess returns. Standard errors, in parentheses, are clustered by day.

Table (A.4) Continuous diff-in-diff regressions of raw excess returns. Standard errors, in parentheses, are clustered by day.

	Dependent Variable: Abnormal Returns						
	All Firms	Quartile I	Quartile II	Quartile III	Quartile IV	Top decile	Top 1%
Market Cap.	-0.0000718	-0.388	0.0112	-0.0158*	-0.0000519	0.0000124	0.000195*
	(0.000168)	(0.434)	(0.0419)	(0.00924)	(0.000109)	(0.0000747)	(0.000106)
Market Cap. $\times TCJA\ dates$	-0.00269**	-4.995	-0.292	-0.137**	-0.000339	-0.000111	0.000370
	(0.00118)	(4.900)	(0.489)	(0.0689)	(0.000289)	(0.000297)	(0.001000)
Proportion of foreign profits	-0.00115	-0.000639	-0.000923	0.00120	-0.00235**	0.00218	-0.00755
	(0.000815)	(0.00227)	(0.00134)	(0.00226)	(0.00118)	(0.00237)	(0.0167)
Proportion of foreign profits $\times TCJA\ dates$	-0.00772	0.00864	-0.0215	-0.00344	0.00560	-0.0188	-0.164
	(0.00627)	(0.0113)	(0.0160)	(0.0201)	(0.00907)	(0.0364)	(0.182)
Avg. for. tax rate	-0.00142	0.000569	0.000371	-0.00519	-0.00434	-0.00634	0.0156
	(0.00166)	(0.00708)	(0.00192)	(0.00319)	(0.00507)	(0.00614)	(0.0446)
Avg. for. tax rate $\times TCJA\ dates$	0.0205	-0.0211	0.0333***	0.0132	0.0409*	0.0285	-0.0503
	(0.0146)	(0.0503)	(0.00991)	(0.0224)	(0.0231)	(0.0366)	(0.205)
N	112,6540	281,600	281,371	281,920	282,215	112,982	11,037
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

A.4.2. Market Model

Now, abnormal returns are being computed using the following asset pricing model:

$$R_{it} = \alpha_i + \beta_i R_t^m + \epsilon_i \tag{A.1}$$

	Dependent Variable: Abnormal Returns	
	Herfindahl	Lerner
Market share	-0.0245	
	(0.167)	
$HHI\times$ market share	0.0116	
	(0.266)	
$HHI\times$ market share $\times TCJA\ dates$	0.396	
	(0.585)	
Lerner index		-0.0000141
		(0.0000233)
Lerner index $\times TCJA\ dates$		0.000294^{**}
		(0.000130)
Proportion of foreign profits	-0.000346	-0.000459
	(0.000815)	(0.000813)
Proportion of foreign profits $\times TCJA\ dates$	-0.00919	-0.00926
	(0.00595)	(0.00573)
Avg. for. tax rate	0.000338	0.000315
	(0.00165)	(0.00164)
Avg. for. tax rate $\times TCJA\ dates$	0.0195	0.0205
	(0.0139)	(0.0133)
N	1,274,778	1,166,039
Industry FE	Yes	Yes
Time FE	Yes	Yes

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

Table (A.5) Continuous diff-in-diff regressions of abnormal returns, computed using equation A.1. Standard errors, in parentheses, are clustered by day.

Table (A.6) Continuous diff-in-diff regressions of abnormal returns, computed using equation A.1. Standard errors, in parentheses, are clustered by day.

	Dependent Variable: Abnormal Returns						
	All Firms	Quartile I	Quartile II	Quartile III	Quartile IV	Top decile	Top 1%
Market Cap.	-0.000146	-0.766*	-0.0683	-0.0183*	-0.0000685	-0.0000252	0.0000562
	(0.000170)	(0.421)	(0.0425)	(0.00933)	(0.000102)	(0.0000715)	(0.000105)
Market Cap. $\times TCJA\ dates$	-0.00274**	-5.430	-0.429	-0.136**	-0.000145	-0.0000134	0.000215
	(0.00119)	(4.887)	(0.492)	(0.0689)	(0.000319)	(0.000341)	(0.000967)
Proportion of foreign profits	-0.000276	-0.000455	-0.000169	0.00241	0.000280	0.00289	0.0176
	(0.000822)	(0.00227)	(0.00134)	(0.00226)	(0.00114)	(0.00231)	(0.0166)
Proportion of foreign profits $\times TCJA\ dates$	-0.00867	0.00805	-0.0226	-0.00647	0.00728	-0.0192	-0.146
	(0.00592)	(0.0115)	(0.0158)	(0.0189)	(0.00935)	(0.0362)	(0.184)
Avg. for. tax rate	0.000183	0.00349	0.00194	-0.00296	0.000541	-0.00273	0.0494
	(0.00165)	(0.00711)	(0.00191)	(0.00318)	(0.00508)	(0.00611)	(0.0453)
Avg. for. tax rate $\times TCJA\ dates$	0.0192	-0.0276	0.0318**	0.0155	0.0511**	0.0417	-0.0950
	(0.0140)	(0.0471)	(0.00963)	(0.0232)	(0.0247)	(0.0343)	(0.228)
N	112,6540	281,600	281,371	281,920	282,215	112,982	11,037
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

A.4.3. Naked CAPM

Now, abnormal returns are being computed using the following asset pricing model:

$$R_{it} - R_t^f = \beta_i(R_t^m - R_t^f) + \epsilon_i \tag{A.2}$$

	Dependent Variable: Abnormal Returns	
	Herfindahl	Lerner
Market share	-0.189	
	(0.167)	
$HHI\times$ market share	0.274	
	(0.266)	
$HHI\times$ market share $\times TCJA\ dates$	0.396	
	(0.585)	
Lerner index		0.000000623
		(0.0000233)
Lerner index $\times TCJA\ dates$		0.000294**
		(0.000131)
Proportion of foreign profits	-0.00145*	-0.00168**
	(0.000814)	(0.000813)
Proportion of foreign profits $\times TCJA\ dates$	-0.00923	-0.00925
	(0.00596)	(0.00573)
Avg. for. tax rate	-0.00118	-0.00128
	(0.00165)	(0.00164)
Avg. for. tax rate $\times TCJA\ dates$	0.0194	0.0204
	(0.0139)	(0.0133)
N	1,274,778	1,166,039
Industry FE	Yes	Yes
Time FE	Yes	Yes

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

Table (A.7) Continuous diff-in-diff regressions of abnormal returns, computed using equation A.2. Standard errors, in parentheses, are clustered by day.

Table (A.8) Continuous diff-in-diff regressions of abnormal returns, computed using equation A.2. Standard errors, in parentheses, are clustered by day.

	Dependent Variable: Abnormal Returns						
	All Firms	Quartile I	Quartile II	Quartile III	Quartile IV	Top decile	Top 1%
Market Cap.	-0.0000938	-0.629	-0.0277	-0.0179*	0.0000115	0.0000644	0.000180*
	(0.000170)	(0.420)	(0.0425)	(0.00933)	(0.000102)	(0.0000715)	(0.000105)
Market Cap. $\times TCJA\ dates$	-0.00275**	-5.550	-0.433	-0.137**	-0.000146	-0.0000159	0.000213
	(0.00119)	(4.888)	(0.491)	(0.0689)	(0.000319)	(0.000340)	(0.000966)
Proportion of foreign profits	-0.00134	-0.000844	-0.00120	0.00111	-0.00181	0.00306	-0.00547
	(0.000821)	(0.00227)	(0.00134)	(0.00226)	(0.00114)	(0.00231)	(0.0166)
Proportion of foreign profits $\times TCJA\ dates$	-0.00870	0.00783	-0.0226	-0.00647	0.00731	-0.0193	-0.146
	(0.00593)	(0.0115)	(0.0158)	(0.0189)	(0.00935)	(0.0362)	(0.184)
Avg. for. tax rate	-0.00190	-0.000491	-0.000150	-0.00478	-0.00327	-0.00482	0.00376
	(0.00165)	(0.00711)	(0.00191)	(0.00318)	(0.00508)	(0.00611)	(0.0453)
Avg. for. tax rate $\times TCJA\ dates$	0.0192	-0.0276	0.0319**	0.0155	0.0515**	0.0418	-0.0937
	(0.0140)	(0.0472)	(0.00963)	(0.0232)	(0.0247)	(0.0343)	(0.227)
N	112,6540	281,600	281,371	281,920	282,215	112,982	11,037
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

A.4.4. 5-Factor Fama-French Model

Now, abnormal returns are being computed using the following asset pricing model:

$$R_{i,t} - R_t^f = \beta_i^m (R_t^m - R_t^f) + \beta_i^{HML} HML_t + \beta_i^{SMB} SMB_t + \beta_i^{RMW} RMW_t + \beta_i^{CMA} CMA_t + u_{i,t}, \tag{A.3}$$

where the two additional factors RMW (Robust Minus Weak) and CMA (Conservative Minus Aggressive) represent respectively the premium on robust relative to weak operating profitability portfolios and the premium on conservative relative to aggressive investment portfolios.

	Dependent Variable: Abnormal Returns	
	Herfindahl	Lerner
Market share	-0.203*	
	(0.115)	
$HHI\times$ market share	0.297	
	(0.206)	
$HHI\times$ market share $\times TCJA\ dates$	0.403	
	(0.356)	
Lerner index		0.00000298
		(0.0000234)
Lerner index $\times TCJA\ dates$		0.000212
		(0.000176)
Proportion of foreign profits	-0.00156*	-0.00180**
	(0.000827)	(0.000825)
Proportion of foreign profits $\times TCJA\ dates$	-0.00667	-0.00654
	(0.00567)	(0.00541)
Avg. for. tax rate	-0.00111	-0.00124
	(0.00162)	(0.00161)
Avg. for. tax rate $\times TCJA\ dates$	0.0122	0.0143
	(0.0129)	(0.0125)
N	1,274,778	1,166,039
Industry FE	Yes	Yes
Time FE	Yes	Yes

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

Table (A.9) Continuous diff-in-diff regressions of abnormal returns, computed using equation A.3. Standard errors, in parentheses, are clustered by day.

Table (A.10) Continuous diff-in-diff regressions of abnormal returns, computed using equation A.3. Standard errors, in parentheses, are clustered by day.

	Dependent Variable: Abnormal Returns						
	All Firms	Quartile I	Quartile II	Quartile III	Quartile IV	Top decile	Top 1%
Market Cap.	-0.000239**	-0.636	-0.0394	-0.0167**	-0.000112*	-0.0000504	0.0000173
	(0.0000757)	(0.411)	(0.0457)	(0.00840)	(0.0000670)	(0.0000705)	(0.000137)
Market Cap. $\times TCJA\ dates$	-0.000708	-7.397	-0.809*	-0.0441	0.00116**	0.000462	0.000352
	(0.000497)	(4.852)	(0.445)	(0.0722)	(0.000427)	(0.000533)	(0.00101)
Proportion of foreign profits	-0.00145*	-0.000797	-0.00118	0.000763	-0.00197*	0.00208	-0.00347
	(0.000833)	(0.00229)	(0.00137)	(0.00230)	(0.00118)	(0.00229)	(0.0166)
Proportion of foreign profits $\times TCJA\ dates$	-0.00647	0.00947	-0.0211	-0.0133	0.0181**	-0.0245	-0.0532
	(0.00546)	(0.0116)	(0.0155)	(0.0214)	(0.00793)	(0.0290)	(0.178)
Avg. for. tax rate	-0.00184	0.000130	-0.00000613	-0.00470	-0.00374	-0.00520	-0.00856
	(0.00163)	(0.00714)	(0.00186)	(0.00320)	(0.00512)	(0.00609)	(0.0494)
Avg. for. tax rate $\times TCJA\ dates$	0.0117	-0.0496	0.0234**	0.0193	0.0621**	0.0428	-0.328**
	(0.0129)	(0.0422)	(0.00749)	(0.0226)	(0.0298)	(0.0352)	(0.164)
N	112,6540	281,600	281,371	281,920	282,215	112,982	11,037
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

A.5. Robustness Checks Regarding Specification 1.3

A.5.1. "Parallel Trends"

In the absence of a subsample that is not affected by the corporate tax cut (which is, of course, impossible to obtain in a sample of publicly traded corporations) one cannot run a true parallel trends robustness test for the main diff-in-diff specification in equation 1.3. Instead, what follows shows that the coefficient on the interaction between different measures of M_i and each individual date tends to move with changes in the betting prices. Because stock market prices should follow a random walk, and because innovations in contract price should be exogenous, these results suggest that the results obtained using our 6 dates of interest are at least not entirely spurious.

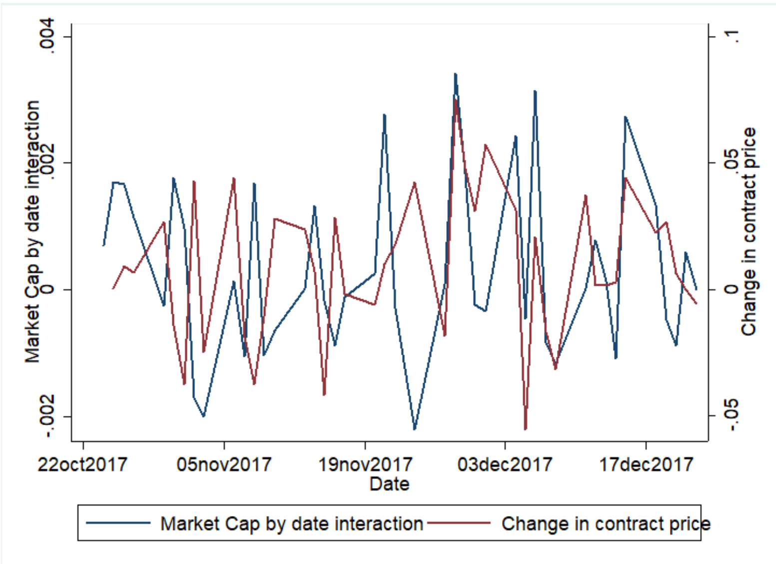


Figure (A.9) How the coefficient on Market Cap. and date dummies changes with ΔP. Correlation between the two series is 0.207.

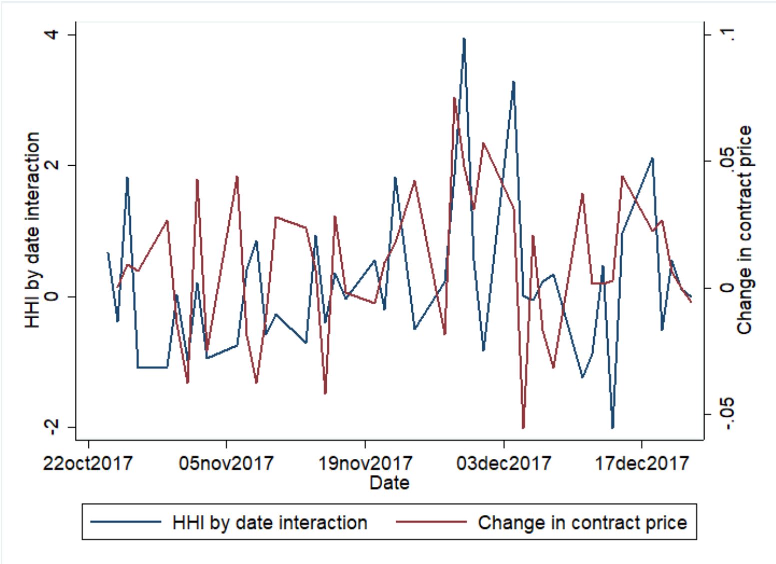


Figure (A.10) How the coefficient on $HHI \times s$ and date dummies changes with ΔP. Correlation between the two series is 0.2372.

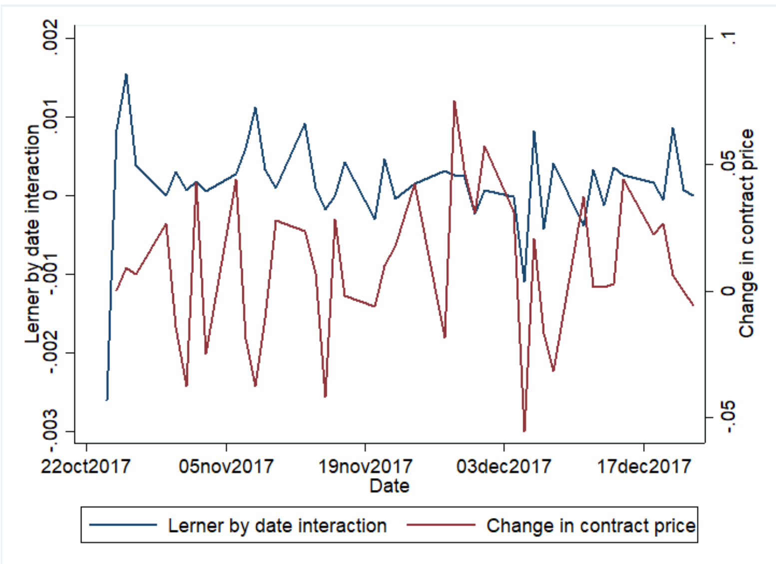


Figure (A.11) How the coefficient on L and date dummies changes with ΔP. Correlation between the two series is 0.0579.

A.5.2. No Industry Fixed Effects

Here I show results of specification 1.3 without industry fixed effects.

	Dependent Variable: Abnormal Returns	
	Herfindahl	Lerner
Market share	-0.187*	
	(0.113)	
$HHI\times$ market share	0.210	
	(0.129)	
$HHI\times$ market share $\times TCJA\ dates$	0.899*	
	(0.465)	
Lerner index		0.00000300
		(0.0000241)
Lerner index $\times TCJA\ dates$		0.000286**
		(0.000137)
Proportion of foreign profits	-0.00111	-0.00139*
	(0.000809)	(0.000802)
Proportion of foreign profits $\times TCJA\ dates$	-0.00604	-0.00557
	(0.00588)	(0.00575)
Avg. for. tax rate	-0.000613	-0.000835
	(0.00158)	(0.00157)
Avg. for. tax rate $\times TCJA\ dates$	0.0154	0.0173
	(0.0133)	(0.0127)
N	1,274,778	1,166,039
Industry FE	No	No
Time FE	Yes	Yes

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

Table (A.11) Continuous diff-in-diff regressions of abnormal returns, without industry fixed effects. Standard errors, in parentheses, are clustered by day.

Table (A.12) Continuous diff-in-diff regressions of abnormal returns, without industry fixed effects. Standard errors, in parentheses, are clustered by day.

	Dependent Variable: Abnormal Returns						
	All Firms	Quartile I	Quartile II	Quartile III	Quartile IV	Top decile	Top 1%
Market Cap.	-0.000143*	-0.545	-0.0363	-0.0229**	-0.00000487	0.00000741	0.0000354
	(0.0000732)	(0.417)	(0.0535)	(0.00829)	(0.0000672)	(0.0000709)	(0.0000909)
Market Cap. $\times TCJA\ dates$	0.000158	-7.992	-0.390	-0.0386	0.00130**	0.000752*	0.000560
	(0.000747)	(5.330)	(0.536)	(0.0593)	(0.000484)	(0.000424)	(0.000794)
Proportion of foreign profits	-0.00106	-0.000806	-0.00130	0.000626	-0.00116	0.00392	-0.00290
	(0.000811)	(0.00226)	(0.00134)	(0.00283)	(0.00113)	(0.00254)	(0.0188)
Proportion of foreign profits $\times TCJA\ dates$	-0.00556	0.0115	-0.0224	-0.0112	0.0170**	-0.0171	-0.0843
	(0.00578)	(0.0112)	(0.0155)	(0.0211)	(0.00829)	(0.0298)	(0.184)
Avg. for. tax rate	-0.00143	-0.000247	-0.0000341	-0.00410	0.00313	0.000273	-0.0482
	(0.00158)	(0.00696)	(0.00173)	(0.00320)	(0.00610)	(0.00718)	(0.0393)
Avg. for. tax rate $\times TCJA\ dates$	0.0153	-0.0490	0.0295**	0.0161	0.0674**	0.0394	-0.292
	(0.0136)	(0.0447)	(0.0107)	(0.0250)	(0.0250)	(0.0477)	(0.228)
N	1,126,540	266,018	279,477	284,891	296,437	119,008	12,452
Industry FE	No	No	No	No	No	No	No
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

A.6. Hernfindahl-Hirschmann Index in Manufacturing

While the market shares (and resulting measurements of HHI) are computed using only Compustat data. This runs the risk of omitting potentially relevant competitors in each market which are not publicly traded and thus not on Compustat. The U.S. Census data provides concentration data by NAICS code, but unfortunately only for manufacturing sectors. If we repeat the analysis using these data, we obtain the following estimates:

	Dependent Variable: Abnormal Returns	
Market share	-0.0178	-0.00228
	(0.0179)	(0.00600)
$HHI\times$ market share	0.142	-0.000666
	(0.132)	(0.0389)
$HHI\times$ market share $\times TCJA$ *dates*	0.116	0.119
	(0.202)	(0.202)
Proportion of foreign profits	-0.000481	-0.000828
	(0.00106)	(0.000994)
Proportion of foreign profits $\times TCJA$ *dates*	-0.0114	-0.0114
	(0.0110)	(0.0110)
Avg. for. tax rate	-0.00222	-0.000856
	(0.00261)	(0.00257)
Avg. for. tax rate $\times TCJA$ *dates*	0.0128	0.0121
	(0.0195)	(0.0196)
N	452,163	452,163
Industry FE	Yes	No
Time FE	Yes	Yes

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

Table (A.13) Continuous diff-in-diff regressions of abnormal returns. Market shares and HHI computed using Census data on the manufacturing sector. Standard errors, in parentheses, are clustered by day.

A.7. Robustness Checks Regarding the Distribution of Excess Returns by Firm

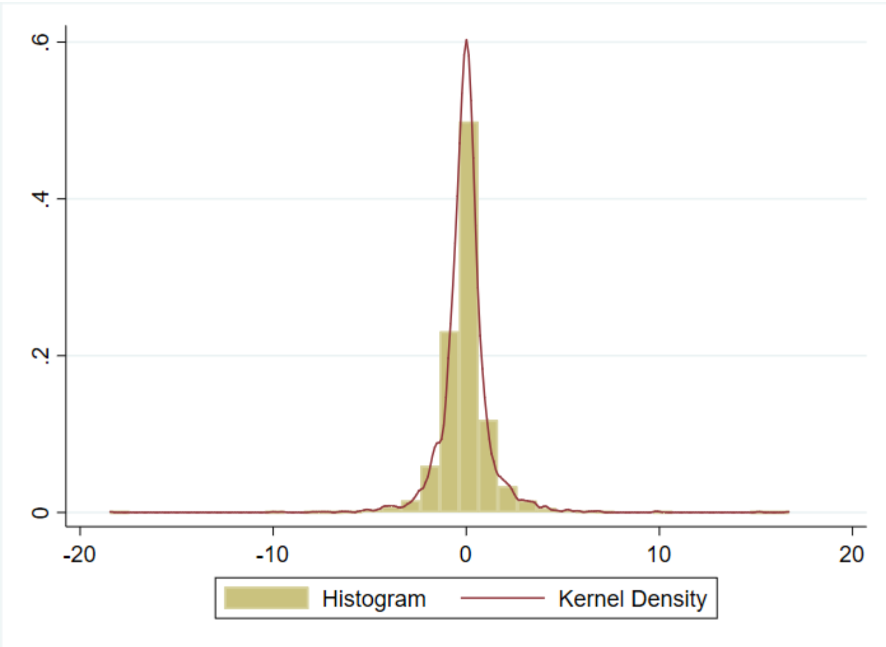


Figure (A.12) Distribution of excess returns due to six random dates aroung August 2017. The specific dates picked for this regression were July 31, 2017; August 8, 2017; August 16, 2017; August 23, 2017; August 25, 2017;and September 4, 2017.

APPENDIX B

Appendix for Chapter 2

B.1. Additional Results and Proofs

B.1.1. Additional results and proofs from section 2.2

Proposition 2. *Suppose we can extend u to $\mathbb{R}^2$ such that u is continuous and quasi-concave. Then for any $\bar{p}$, p^{NT}, τ, and W on which $\boldsymbol{q}$ is defined, there exist scalar values p^s and W^s such that:*

$$\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W) \in \underset{p^s q + p^{NT} q^{NT} \leq W^s}{\arg\max} u(q, q^{NT})$$
$$\boldsymbol{q}(\bar{p}, p^{NT}, \tau, W^s) * (\bar{p} + \tau, p^{NT}) = W$$

We demonstrate a generalization of proposition 2, in which multiple goods may be taxed. We consider a general setting with N goods, consumption set $X = X^T \times X^{NT} \subseteq \mathbb{R}^N_+$, with consumption vector $\boldsymbol{q} = (\boldsymbol{q^T}, \boldsymbol{q^{NT}}) \in X$. Here X^T is the consumption set for taxed goods, while X^{NT} is the consumption set for non-taxed goods. We assume that either $X^{NT} \subseteq \mathbb{R}_+$ or X^{NT} is convex.

The agent has preferences $\succeq$ on X. Informally, we want to assume preferences such that agents smoothly prefer moderation. To say that they prefer moderation, one generally assumes convex preferences. However, we do not want to assume a convex consumption set X. We might alternatively assume that preferences are *pseudo-convex*, in that for any $\boldsymbol{q} \in X$ and any finite n:

$$\boldsymbol{q}_k \in X, \boldsymbol{q}_k \succ \boldsymbol{q}, \lambda_k \geq 0 \; \forall k = 1, ..., n, \sum_{k=1}^{n} \lambda_k = 1, \sum_{k=1}^{n} \lambda_k \boldsymbol{q}_k \in X \Rightarrow \sum_{k=1}^{n} \lambda_k \boldsymbol{q}_k \succ \boldsymbol{q}$$

However, we also want some smoothness to preferences. More formally, we want to figure that if $\boldsymbol{q'} \succ \boldsymbol{q}$, then there is an epsilon ball around $\boldsymbol{q'}$ such that the agent would prefer any element in that epsilon ball to $\boldsymbol{q}$ if that element were also in the consumption set. Furthermore, any convex combination of points in these epsilon balls should yield a point that, if contained in X, is also strictly preferred to $\boldsymbol{q}$. We refer to this assumption on preferences as *continuous pseudo-convexity* (CPC).

Assumption 2. *For any $\boldsymbol{q} \in X$, define the set of strictly preferred allocations:*

$$\mathcal{A} \equiv \{\boldsymbol{q'} \in X | \boldsymbol{q'} \succ \boldsymbol{q}\}$$

There exists some function $\epsilon : \mathcal{A} \to \mathbb{R}_{++}$ such that for any $n \in \mathbb{N}$, for any $\lambda_1, ..., \lambda_n \geq 0$ and $\boldsymbol{q}_1, ..., \boldsymbol{q}_n \in \mathcal{A}$, if $\sum_{k=1}^{n} \lambda_k = 1$, then :

$$\exists \boldsymbol{q}'_1, ..., \boldsymbol{q}'_n \in \mathbb{R}^n : ||\boldsymbol{q}'_k - \boldsymbol{q}_k|| < \epsilon(\boldsymbol{q_k}) \ \forall k, \sum_{k=1}^{n} \lambda_k \boldsymbol{q_k}' \in X \Rightarrow \sum_{k=1}^{n} \lambda_k \boldsymbol{q_k}' \succ \boldsymbol{q}$$

We provide this description of CPC preferences to facilitate intuition, but our main result for this section comes from an equivalent, yet more geometric, expression of this description.

Lemma 1. *Preferences $\succeq$ are CPC if and only if for every $\boldsymbol{q} \in X$ with corresponding set of strictly preferred bundles $\mathcal{A}$ there is an open and convex set $\mathcal{O} \subseteq \mathbb{R}^N$ such that $\mathcal{O} \cap X = \mathcal{A}$.*[1]

Proof: For one direction, the convex hull of the union of open $\epsilon(\boldsymbol{q}')$ balls around $\boldsymbol{q}' \in \mathcal{A}$ is open, and by assumption does not contain any elements of $X \setminus \mathcal{A}$. For the other direction, for any $\boldsymbol{q}' \in \mathcal{A}$, define $\epsilon(\boldsymbol{q}')$ as a positive value such that $\boldsymbol{q}'' \in \mathbb{R}^N_+ : ||\boldsymbol{q}'' - \boldsymbol{q}'|| < \epsilon(\boldsymbol{q}') \Rightarrow \boldsymbol{q}'' \in \mathcal{O}$. We can do so because $\mathcal{O}$ is open. For any such $\boldsymbol{q}''$, if $\boldsymbol{q}'' \in X$, then $\boldsymbol{q}'' \succ \boldsymbol{q}$. □

Let $\boldsymbol{p} = (\boldsymbol{p^T}, \boldsymbol{p^{NT}}) \in \mathbb{R}^N_+$ denote a generic price vector, where $\boldsymbol{p^T}$ and $\boldsymbol{p^{NT}}$ are price vectors for taxed and non-taxed goods respectively. In particular, let $\bar{\boldsymbol{p}} = (\bar{\boldsymbol{p}}^{\boldsymbol{T}}, \bar{\boldsymbol{p}}^{\boldsymbol{NT}})$ denote the vector of sticker prices.

Let $\boldsymbol{\tau}$ denote the vector of taxes for taxed goods, so that $\boldsymbol{q^T}$, $\boldsymbol{q^{NT}}$, and $\boldsymbol{\tau}$ all have the same number of elements. The consumption vector $\boldsymbol{q}(\bar{\boldsymbol{p}}, \boldsymbol{\tau}) = (\boldsymbol{q}^T(\bar{\boldsymbol{p}}, \boldsymbol{\tau}), \boldsymbol{q}^{NT}(\bar{\boldsymbol{p}}, \boldsymbol{\tau}))$ satisfies the following properties:

$$\begin{aligned}
&\bar{p}^{NT} * \tilde{\boldsymbol{q}}^{NT} \leq W - \bar{\boldsymbol{p}}^{\boldsymbol{T}} * \boldsymbol{q}^T \\
&(\boldsymbol{q}^T, \tilde{\boldsymbol{q}}^{NT}) \succ (\boldsymbol{q}^T, \hat{\boldsymbol{q}}^{NT}) \ \ \forall \hat{\boldsymbol{q}}^{NT} \in X^{NT} : \bar{\boldsymbol{p}}^{\boldsymbol{NT}} * \tilde{\boldsymbol{q}}^{NT} \leq W - \bar{\boldsymbol{p}}^{\boldsymbol{T}} * \boldsymbol{q}^T \\
&\quad \boldsymbol{q}(\bar{\boldsymbol{p}}, \boldsymbol{0}) \in \underset{\tilde{\boldsymbol{q}} \in X : \bar{\boldsymbol{p}} * \tilde{\boldsymbol{q}} \leq W}{\arg\max} \succeq
\end{aligned}$$

In words, consumption of the non-taxed goods is always optimally determined upon choosing consumption of the taxed goods, and consumption is optimally determined when the agent correctly perceives prices, i.e. when there are no taxes. We also restrict the domain of sticker prices and taxes so that expenditure on non-taxed goods is positive, i.e.:

$$\bar{\boldsymbol{p}}^{NT} * \boldsymbol{q}^{NT}(\bar{\boldsymbol{p}}, \boldsymbol{\tau}) > 0$$

The claim is that for any $\bar{\boldsymbol{p}}$ and $\boldsymbol{\tau}$ in this domain, there is a (p^s, W^s) that explains $\boldsymbol{q}(\bar{\boldsymbol{p}}, \boldsymbol{\tau})$.

Proof of Generalization of Proposition 2: Define $\boldsymbol{q} = (\boldsymbol{q^T}, \boldsymbol{q^{NT}}) = \boldsymbol{q}(\bar{\boldsymbol{p}}, \boldsymbol{\tau})$ and:

$$\mathcal{A}^e \equiv \{(\boldsymbol{q^{T'}}, e^{NT'}) | \boldsymbol{q^{T'}} \in X^T, \exists \boldsymbol{q^{NT'}} \in X^{NT} : \bar{\boldsymbol{p}}^{NT} * \boldsymbol{q^{NT'}} = e^{NT'}, (\boldsymbol{q}^{T'}, \boldsymbol{q}^{NT'}) \in \mathcal{O}\}$$

Suppose for the sake of contradiction that $(\boldsymbol{q^T}, \bar{\boldsymbol{p}}^{\boldsymbol{NT}} * \boldsymbol{q^{NT}}) \in Co(\mathcal{A}^e)$, i.e. that

[1] Note that $\mathcal{O}$ is open in $\mathbb{R}^N$.

$\exists n \in \mathbb{N}$, $(\boldsymbol{q}_k^T, e_k^{NT}) \in \mathcal{A}^e$, and $\lambda_k \geq 0\ \forall k = 1, ..., n$ such that $\sum_{k=1}^n \lambda_k = 1$ and:

$$\sum_{k=1}^{n} \lambda_k(\boldsymbol{q}_k^T, e_k^{NT}) = (\boldsymbol{q^T}, \boldsymbol{\bar{p}^{NT}} * \boldsymbol{q^{NT}})$$

Since $(\boldsymbol{q}_k^T, e_k^{NT}) \in \mathcal{A}^e\ \forall k$, that means that:

$$\forall k\ \exists \boldsymbol{q}_k^{NT} : e_k^{NT} = \boldsymbol{\bar{p}^{NT}} * \boldsymbol{q}_k^{NT}, \boldsymbol{q}_k \equiv (\boldsymbol{q}_k^T, \boldsymbol{q}_k^{NT}) \Rightarrow \boldsymbol{q}_k \in \mathcal{O}$$

If $X^{NT} \subseteq \mathbb{R}_+$, then $\sum_{k=1}^n \lambda_k \boldsymbol{\bar{p}^{NT}} * \boldsymbol{q}_k^{NT} = \boldsymbol{\bar{p}^{NT}} * \boldsymbol{q^{NT}}$ implies that $\sum_{k=1}^n \lambda_k \boldsymbol{q}_k^{NT} = \boldsymbol{q^{NT}}$ because positive non-tax expenditure requires that $\boldsymbol{\bar{p}^{NT}} \neq 0$. In that case:

$$\sum_{k=1}^{n} \lambda_k \boldsymbol{q}_k = \boldsymbol{q} \Rightarrow\Leftarrow \sum_{k=1}^{n} \lambda_k \boldsymbol{q}_k \in \mathcal{O}$$

This is a contradiction arising from $\boldsymbol{q} \notin \mathcal{O}$.

If X^{NT} is not a subset of $\mathbb{R}_+$, then X^{NT} is convex. This means $\sum_{k=1}^n \lambda_k \boldsymbol{q}_k \in X$. Pseudo-convexity of preferences implies that:

$$\sum_{k=1}^{n} \lambda_k \boldsymbol{q}_k \succ \boldsymbol{q}$$

Yet the weighted average of taxed goods is the desired taxed good consumption bundle, whereas the weighted average of non-taxed goods is affordable:

$$\sum_{k=1}^{n} \lambda_k \boldsymbol{q_k^T} = \boldsymbol{q}^T$$

$$\boldsymbol{\bar{p}^{NT}} * \sum_{k=1}^{n} \boldsymbol{q_k^{NT}} = \sum_{k=1}^{n} e_k^{NT} = \boldsymbol{\bar{p}^{NT}} * \boldsymbol{q}^{NT}$$

Thus, the agent could not have optimally chosen $\boldsymbol{q}^{NT}$, another contradiction. We conclude that $(\boldsymbol{q}^T, \boldsymbol{\bar{p}}^{NT} * \boldsymbol{q}^{NT}) \notin Co(\mathcal{A}^e)$.

Now, we can apply the Separating Hyperplane Theorem to say that there is a vector $(\boldsymbol{p}^{Ts}, 1)$, where $\boldsymbol{p}^{Ts}$ has as many elements as $\boldsymbol{q}^T$, such that:

$$(\boldsymbol{p}^{Ts}, 1) * (\boldsymbol{q}^T, \boldsymbol{\bar{p}}^{NT} * \boldsymbol{q}^{NT}) \leq (\boldsymbol{p}^{Ts}, 1) * (\boldsymbol{q^{T'}}, e^{NT'})\ \forall (\boldsymbol{q^{T'}}, e^{NT'}) \in Co(\mathcal{A}^e)$$

Defining $\boldsymbol{p}^s \equiv (\boldsymbol{p}^{Ts}, \boldsymbol{\bar{p}}^{NT})$, this implies that for any bundle $\boldsymbol{q'} = (\boldsymbol{q}^{T'}, \boldsymbol{q}^{NT'}) \in \mathcal{O}$:

$$\boldsymbol{p}^s * \boldsymbol{q}^{'} \geq \boldsymbol{p}^s * \boldsymbol{q}$$

Since $\mathcal{O}$ is open, the above expression can never be satisfied with equality. To see this, suppose otherwise, i.e. that $\exists \boldsymbol{q'} \in \mathcal{O}$ such that:

$$\boldsymbol{p}^s * \boldsymbol{q}^{'} = \boldsymbol{p}^s * \boldsymbol{q}$$

Note that $\boldsymbol{\bar{p}}^{NT} > \boldsymbol{0}$ implies that we can choose $\boldsymbol{q}''$ within $\epsilon(\boldsymbol{q'})$ of $\boldsymbol{q'}$ by slightly reducing a component of $\boldsymbol{q'}$ for which the corresponding perceived price is positive.

Thus, $\boldsymbol{q}'' \in \mathcal{O}$, yet $\boldsymbol{p}^s * \boldsymbol{q}'' < \boldsymbol{p}^s * \boldsymbol{q}$. This yields our desired contradiction. Therefore:

$$\boldsymbol{p}^s * \boldsymbol{q}' > \boldsymbol{p}^s * \boldsymbol{q} \ \forall \boldsymbol{q}' \in \mathcal{O}$$

We conclude by defining $W^s \equiv \boldsymbol{p}^s * \boldsymbol{q}$ and noting that $\forall \boldsymbol{q}' \in X$:

$$\boldsymbol{q}' \succ \boldsymbol{q} \Rightarrow \boldsymbol{q}' \in \mathcal{O} \Rightarrow \boldsymbol{p^s} > W^s$$

Therefore, the model has rationalized consumption because no preferred consumption bundle is perceived to be affordable. □

Now that we've gone through the proof, we can make a couple of observations. One, the assumption of CPC preferences is satisfied when preferences are represented by a lower semi-continuous and quasi-concave function u on $\mathbb{R}^N$, so that:

$$\forall x, y \in X : x \succeq y \Leftrightarrow u(x) \geq u(y)$$

This makes it clear that we have, in fact, generalized proposition 2. Also, note that it may be easier in practice to check to see that preferences have such a utility representation than to check that they satisfy continuous pseudo-convexity.

Two, it may appear strange that we needed to assume that X^{NT} is concave specifically if it has dimension greater than one. This is because a discrete grid for consumption of non-taxed goods can create a lumpy evaluation of non-tax expenditure, thwarting the existence of a separating hyperplane. For example, consider a consumption set $\mathbb{R}_+ \times \{0, 1\}^2$, where there is one taxed good chosen continuously and two non-taxed goods chosen from $\{0, 1\}$. The sticker price vector is $\bar{p} = (1, 1, 1)$. The consumer have preferences rationalized by the function:

$$u(\boldsymbol{q}) = q_1 + \min\{q_2, q_3\}$$

In words, the taxed good is perfect substitutes with the minimum consumption of the two non-taxed goods, which are perfect complements with each other. Consider the consumption bundle:

$$\boldsymbol{q} = (0, 1, 0)$$

If the agent perceived income $W^s \geq 2$, they could do better by consuming $(0, 1, 1)$. Supposing otherwise, if the agent perceives a positive tax-inclusive price of the taxed good, then optimally $q_1 > 0$ and $q_2 = q_3 = 0$. Finally, there is no optimal consumption bundle if $p_1^s \leq 0$. Thus, the consumption bundle cannot be rationalized.

Next, we derive our expression for the change in consumer surplus due to the tax:

Proposition 3. *Let $e(p)$ and $h(p)$ denote the expenditure function and compensated demand for the taxed good respectively at price p for the taxed good and price p^{NT} for the other good, so that the agent is minimally compensated so as to achieve utility of at least $u(\boldsymbol{q}(\bar{p}, p^{NT}, 0, W))$; formally,*

$$e(p) = \min\{W' | u(d(p, p^{NT}, W'), d^{NT}(p, p^{NT}, W)) \geq u(\boldsymbol{q}(\bar{p}, p^{NT}, 0, W))\},$$

which is well-defined by continuity of u and connectedness of the choice set. Then *compensating variation due to the tax satisfies:*

$$\Delta CS = (\bar{p} + \tau - p^s)h(p^s) + e(p^s) - e(\bar{p}) \tag{B.1}$$

Proof: Letting W^s denote conjectured wealth when facing tax τ, local non-satiation of preferences implies that:

$$(p^s, \bar{p}^{NT}) * \boldsymbol{q}(\bar{p}, p^{NT}, \tau, W + \Delta CS) = W^s = e(p^s)$$

In words, total perceived expenditures equal perceived wealth, which must be exactly the wealth the agent would need under perceived prices to achieve the utility from before the tax. Plugging in and using the fact that $h(p^s) = q(\bar{p}, p^{NT}, \tau, W + \Delta CS)$ yields:

$$(\bar{p} + \tau - p^s)h(p^s) = [(\bar{p} + \tau, \bar{p}^{NT}) - (p^s, \bar{p}^{NT})] * \boldsymbol{q}(\bar{\boldsymbol{p}}, \tau, W + \Delta CS)$$
$$(\bar{p} + \tau - p^s)h(p^s) = W + \Delta CS - e(p^s)$$

Rearranging and again using local non-satiation yields:

$$\Delta CS = e(p^s) - W + (\bar{p} + \tau - p^s)h(p^s) = e(p^s) - e(\bar{p}) + (\bar{p} + \tau - p^s)h(p^s)$$

□

The following lemma establishes the Compensated Law of Demand (CLD) in our setting:

Lemma 2. *For any agent i with type θ_i and any two prices p and p':*

$$p < p' \Rightarrow q(p'; \theta_i, h) \leq q(p; \theta_i, l)$$

Proof: Note that there must be values q^{NT} and $q^{NT'}$ such that:

$$(q(p; \theta_i, l), q^{NT}) \sim_{\theta_i} (q(p'; \theta_i, h), q^{NT'})$$

From local non-satiation:

$$p * q(p; \theta_i, l) + p^{NT} * q^{NT} \leq p * q(p'; \theta_i, h) + p^{NT} * q^{NT'}$$

$$p' * q(p'; \theta_i, h) + p^{NT} * q^{NT'} \leq p' * q(p; \theta_i, l) + p^{NT} * q^{NT}$$

Rearranging yields:

$$p * [q(p; \theta_i, l) - q(p'; \theta_i, h)] \leq p^{NT} * [q^{NT'} - q^{NT}] \leq p' * [q(p; \theta_i, l) - q(p'; \theta_i, h)]$$

Thus, $p' > p \Rightarrow q(p; \theta_i, l) \geq q(p'; \theta_i, h)$. □

Proposition 4. *Assume a continuously differentiable and strictly increasing aggregate supply function Q^{supply}, as well as continuously differentiable compensated demand functions h_i and subjective price functions p_i^s $\forall i$. Subjective price functions change one-for-one with sticker prices, so that:*

$$p_i^s(\bar{p}, \tau) = \bar{p} + p_i^s(0, \tau) \ \forall \bar{p} \ \forall \tau \ \forall i$$

Subjective prices also agree with sticker prices when there is no tax:

$$p_i^s(\bar{p}, 0) = \bar{p} \ \ \forall \bar{p} \ \forall i$$

We implicitly define the pre-tax sticker price $\bar{p}^{old}$ by:[2]

$$Q^{supply}(\bar{p}^{old}) = \sum_i h_i(\bar{p}^{old}, \nu_i)$$

and the new sticker price $\bar{p}^{new}$ after the imposition of the tax τ when agents are compensated by:

$$Q^{supply}(\bar{p}^{new}) = \sum_i h_i((p_i^s(\bar{p}^{new}, \tau)), \nu_i)$$

Defining deadweight loss by:[3]

$$DWL \equiv \sum_i \Delta CS_i + \int_{\bar{p}^{new}}^{\bar{p}^{old}} Q^{supply}(p)dp - \tau \sum_i q_i^c$$

where

$$\Delta CS_i = (\bar{p}^{new} + \tau - p_i^s(\bar{p}^{new}, \tau))q_i^c + \int_{\bar{p}^{old}}^{p_i^s(\bar{p}^{new}, \tau)} h_i(p, \nu_i)dp \ \forall i$$

$$q_i^c \equiv h_i(p_i^s(\bar{p}^{new}, \tau), \nu_i) \ \forall i$$

then aggregate deadweight loss has second order approximation around $\tau = 0$:

$$DWL \approx -\frac{1}{2}[\sum_i m_i \frac{\partial h_i}{\partial p} - \frac{(\sum_i m_i \frac{\partial h_i}{\partial p})^2}{\sum_i \frac{\partial h_i}{\partial p} - \frac{\partial Q^{supply}}{\partial p}}]\tau^2$$

Proof:

$$DWL = \sum_i \int_{\bar{p}^{old}}^{p_i^s(\bar{p}^{new}, \tau)} h_i(p, \nu_i)dp + \int_{\bar{p}^{new}}^{\bar{p}^{old}} Q^{supply}(p)dp + \sum_i (\bar{p}^{new} - p_i^s(\bar{p}^{new}, \tau))q_i^c$$

Note that $\bar{p}^{new}$ is a function of τ. One can easily confirm that $\bar{p}^{new}|_{\tau=0} = \bar{p}^{old}$, so that deadweight loss is zero when $\tau = 0$. We can find $\frac{\partial \bar{p}^{new}}{\partial \tau}$ from the Inverse Function Theorem:[4]

$$\frac{\partial Q^{supply}}{\partial p}\frac{\partial \bar{p}^{new}}{\partial \tau} = \sum_i \frac{\partial h_i}{\partial p}[\frac{\partial p_i^s}{\partial \bar{p}^{new}}\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}] = \sum_i \frac{\partial h_i}{\partial p}[\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]$$

$$\frac{\partial \bar{p}^{new}}{\partial \tau} = \frac{\sum_i \frac{\partial h_i}{\partial p}\frac{\partial p_i^s}{\partial \tau}}{\frac{\partial Q^{supply}}{\partial p} - \sum_i \frac{\partial h_i}{\partial p}}$$

[2] $\nu_i \equiv u_i(\boldsymbol{d}_i(\boldsymbol{p}, W_i)) \ \forall i$

[3] Note that $\bar{p}^{new} \leq \bar{p}^{old} \ \forall \tau \geq 0$ from the Compensated Law of Demand and the fact that supply is strictly increasing in price.

[4] This claim also uses the fact that aggregate supply is strictly increasing while aggregate compensated demand is weakly decreasing, so that there is always a unique value for $\bar{p}^{new}$.

We can then take the first derivative of deadweight loss with respect to the tax:

$$\begin{aligned}\frac{\partial DWL}{\partial \tau} &= \sum_i [\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}] h_i - \frac{\partial \bar{p}^{new}}{\partial \tau} Q^{supply}(\bar{p}^{new}) \\ &- \sum_i \left[\frac{\partial p_i^s}{\partial \tau} h_i + (p_i^s(\bar{p}^{new}, \tau) - \bar{p}^{new}) \frac{\partial h_i}{\partial p} [\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]\right] \\ &= \frac{\partial \bar{p}^{new}}{\partial \tau} \sum_i h_i - \frac{\partial \bar{p}^{new}}{\partial \tau} Q^{supply}(\bar{p}^{new}) \\ &- \sum_i \left[(p_i^s(\bar{p}^{new}, \tau) - \bar{p}^{new}) \frac{\partial h_i}{\partial p} [\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]\right] \\ &= -\sum_i \left[(p_i^s(\bar{p}^{new}, \tau) - \bar{p}^{new}) \frac{\partial h_i}{\partial p} [\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]\right]\end{aligned}$$

Since $p_i^s(\bar{p}^{new}, 0) = \bar{p}^{new}$, it follows that

$$\left.\frac{\partial DWL}{\partial \tau}\right|_{\tau=0} = 0$$

Obtaining the second derivative would be straightforward if $h_i \in \mathbb{C}^2 \ \forall i$. Instead, we find the second derivative at $\tau = 0$ from the definition:

$$\left.\frac{\partial^2 DWL}{\partial \tau^2}\right|_{\tau=0} = \lim_{\tau \to 0} -\frac{\sum_i \left[(p_i^s(\bar{p}^{new}, \tau) - \bar{p}^{new}) \frac{\partial h_i}{\partial p} [\frac{\partial \bar{p}^{new}}{\partial \tau} + \frac{\partial p_i^s}{\partial \tau}]\right]}{\tau}$$

Note that continuity of $\frac{\partial p_i^s}{\partial \tau}$ with respect to τ for all agents implies that $\frac{\partial \bar{p}^{new}}{\partial \tau}$ is continuous. Since $\frac{\partial Q^{supply}}{\partial p}$ and $\frac{\partial h_i}{\partial p}$ $\forall i$ are also continuous:

$$\begin{aligned}\left.\frac{\partial^2 DWL}{\partial \tau^2}\right|_{\tau=0} &= -\sum_i \left.\frac{\partial h_i}{\partial p}\right|_{\tau=0} [\left.\frac{\partial \bar{p}^{new}}{\partial \tau}\right|_{\tau=0} + \left.\frac{\partial p_i^s}{\partial \tau}\right|_{\tau=0}] \lim_{\tau \to 0} \frac{(p_i^s(\bar{p}^{new}, \tau) - \bar{p}^{new})}{\tau} \\ &= -\sum_i \left.\frac{\partial h_i}{\partial p}\right|_{\tau=0} [\left.\frac{\partial \bar{p}^{new}}{\partial \tau}\right|_{\tau=0} + \left.\frac{\partial p_i^s}{\partial \tau}\right|_{\tau=0}] \left.\frac{\partial p_i^s}{\partial \tau}\right|_{\tau=0}\end{aligned}$$

Using the fact that $m_i \equiv \frac{\partial p_i^s}{\partial \tau}|_{\tau=0}$, we can note that:

$$\frac{\partial \bar{p}^{new}}{\partial \tau}|_{\tau=0} = \frac{\sum_i m_i \frac{\partial h_i}{\partial p}}{\frac{\partial Q^{supply}}{\partial p} - \sum_i \frac{\partial h_i}{\partial p}}$$

and so:

$$\frac{\partial^2 DWL}{\partial \tau^2}|_{\tau=0} = -[\sum_i m_i^2 \frac{\partial h_i}{\partial p} + \frac{(\sum_i m_i \frac{\partial h_i}{\partial p})^2}{\frac{\partial Q^{supply}}{\partial p} - \sum_i \frac{\partial h_i}{\partial p}}]$$

Now we can find the second order approximation for deadweight loss:

$$DWL \approx DWL|_{\tau=0} + \frac{\partial DWL}{\partial \tau}|_{\tau=0} \tau + \frac{1}{2} \frac{\partial^2 DWL}{\partial \tau^2}|_{\tau=0} \tau^2$$

$$DWL \approx -\frac{1}{2}[\sum_i m_i^2 \frac{\partial h_i}{\partial p} + \frac{(\sum_i m_i \frac{\partial h_i}{\partial p})^2}{\frac{\partial Q^{supply}}{\partial p} - \sum_i \frac{\partial h_i}{\partial p}}]\tau^2$$

□

B.1.2. Additional results and proofs from section 2.3

The upper and lower bounds use the following lemma:

Lemma 3. *For any agent i with type θ_i and any two pairs (p, ζ_i) and (p', ζ_i'):*

$$dwl(p'; \theta_i, \zeta_i) \geq dwl(p; \theta_i, \zeta_i') - (p - \bar{p})(q(p'; \theta_i, \zeta_i') - q(p; \theta_i, \zeta_i)).$$

Proof: Note from the definition of the expenditure function and optimal compensated consumption vectors $\boldsymbol{q}$ and $\boldsymbol{q'}$ for price vectors (p, p^{NT}) and (p', p^{NT}) respectively:

$$e(p') - e(p) = (p', p^{NT}) * \boldsymbol{q'} - (p, p^{NT}) * \boldsymbol{q} \geq (p', p^{NT}) * \boldsymbol{q'} - (p, p^{NT}) * \boldsymbol{q'} = (p' - p)q(p'; \theta_i, \zeta_i')$$

Plugging in yields:

$$\begin{aligned} dwl(p'; \theta_i) &= [e(p') - e(\bar{p})] - (p' - \bar{p})q(p'; \theta_i, \zeta_i') \\ &= [e(p') - e(p)] + [e(p) - e(\bar{p})] - [(p' - p) + (p - \bar{p})]q(p'; \theta_i, \zeta_i') \\ &\geq (p' - p)q(p'; \theta_i, \zeta_i') + e(p) - e(\bar{p}) - [(p' - p) + (p - \bar{p})]q(p'; \theta_i, \zeta_i') \\ &= e(p) - e(\bar{p}) - (p - \bar{p})q(p'; \theta_i, \zeta_i') \\ &= dwl(p) + (p - \bar{p})q(p; \theta_i, \zeta_i) - (p - \bar{p})q(p'; \theta_i, \zeta_i') \\ &= dwl(p; \theta_i) - (p - \bar{p})(q(p'; \theta_i, \zeta_i') - q(p; \theta_i, \zeta_i)) \end{aligned}$$

See also appendix figure B.1 for a graphical demonstration. □

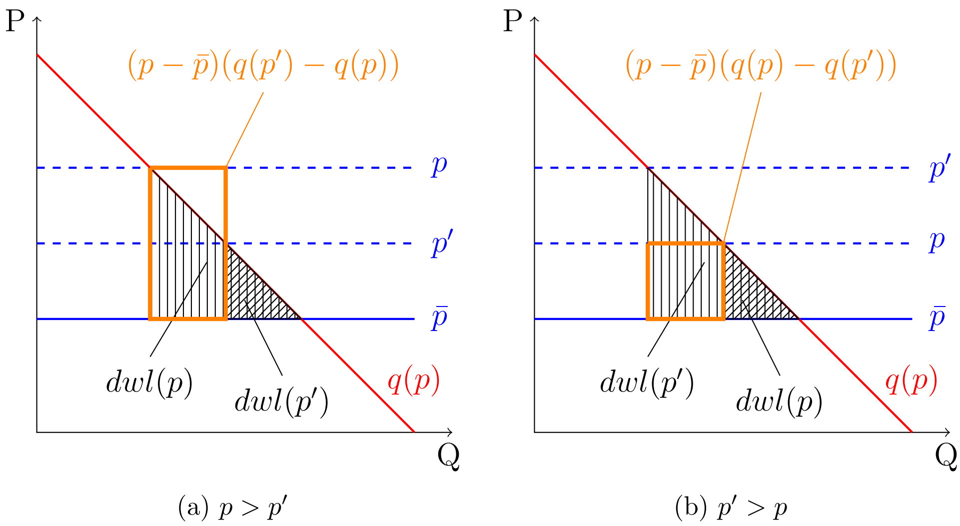


Figure (B.1) A graphical illustration of Lemma 3. As long as demand is weakly decreasing, $dwl(p')$ cannot be smaller than $dwl(p)$ minus (plus) the orange rectangle.

Proof of Proposition 1: From lemma 2 and prices being bound away from zero, we can always find a value of $\hat{p}^s$ such that:

$$\int_{\theta_i} q(\hat{p}^s;\theta_i,l)dF_\theta(\theta_i) \leq \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \leq \int_{\theta_i} q(\hat{p}^s;\theta_i,h)dF_\theta(\theta_i)$$

Pick $\lambda \in [0,1]$ such that:

$$\lambda \int_{\theta_i} q(\hat{p}^s;\theta_i,h)dF_\theta(\theta_i) + (1-\lambda)\int_{\theta_i} q(\hat{p}^s;\theta_i,l)dF_\theta(\theta_i) = \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

Define $F'_{\theta,\zeta}$ such that $F'_\theta = F_\theta$ and $\zeta = h$ with probability λ, $\zeta = l$ with probability $1-\lambda$, $\theta \perp \zeta$. Then:

$$\begin{aligned}\int_{\theta_i,\zeta_i} q(\hat{p}^s;\theta_i,\zeta_i)dF_\theta(\theta_i) &= \int_{\theta_i} [\lambda q(\hat{p}^s;\theta_i,h) + (1-\lambda)q(\hat{p}^s;\theta_i,l)]dF_\theta(\theta_i) \\ &= \int_{\theta_i,\zeta_i} q(pl_i^s;\theta_i,\zeta_i)dF'_{p^s,\theta,\zeta}(\theta_i,\zeta_i)\end{aligned}$$

□

Proof of theorem 1: From lemma 3:

$$\begin{aligned}&\int_{p_i^s,\theta_i,\zeta_i} [dwl(p_i^s;\theta_i,\zeta_i) + (\hat{p}^s - \bar{p})q(p_i^s;\theta_i,\zeta_i)]dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \\ &\qquad \geq \int_{p_i^s,\theta_i,\zeta_i} [dwl(\hat{p}^s;\theta_i,\zeta_i) + (\hat{p}^s - \bar{p})q(\hat{p}^s;\theta_i,\zeta_i)]dF'_{\theta,\zeta}(\theta_i,\zeta_i)\end{aligned}$$

But note from the rationalizability of the data that:

$$(\hat{p}^s - \bar{p})\int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) = (\hat{p}^s - \bar{p})\int_{\theta_i,\zeta_i} q(\hat{p}^s;\theta_i,\zeta_i)dF'_{\theta,\zeta}(\theta_i,\zeta_i)$$

We can thus conclude that:

$$\int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) \geq \int_{\theta_i,\zeta_i} dwl(\hat{p}^s;\theta_i,\zeta_i)dF_{\theta,\zeta}(\theta_i,\zeta_i)$$

□

Proof of theorem 2: From lemma 3 and rationalizability of the data:

$$\int\limits_{p_i^s,\theta_i} [dwl(p^b(p_i^s);\theta_i) + (p_i^s - \bar{p})q(p^b(p_i^s);\theta_i,\zeta_i)]dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$\geq \int\limits_{p_i^s,\theta_i,\zeta_i} [dwl(p_i^s;\theta_i) + (p_i^s - \bar{p})q(p_i^s;\theta_i,\zeta_i)]dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$\int\limits_{p_i^s,\theta_i} [dwl(p^b(p_i^s);\theta_i) + p_i^s q(p^b(p_i^s);\theta_i,\zeta_i)]dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$\geq \int\limits_{p_i^s,\theta_i,\zeta_i} [dwl(p_i^s;\theta_i) + p_i^s q(p_i^s;\theta_i,\zeta_i)]dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$\int\limits_{p_i^s,\theta_i} [dwl(p^b(p_i^s);\theta_i,\zeta_i) + (p_i^s - \tilde{p}^s)q(p^b(p_i^s);\theta_i,\zeta_i)]dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$\geq \int\limits_{p_i^s,\theta_i,\zeta_i} [dwl(p_i^s;\theta_i) + (p_i^s - \tilde{p}^s)q(p_i^s;\theta_i,\zeta_i)]dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

Rearranging yields:

$$\begin{aligned}\int_{p_i^s,\theta_i,\zeta_i} & dwl(p^b(p_i^s);\theta_i,\zeta_i)dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\\ &\geq \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\\ &\quad - \int_{p_i^s,\theta_i,\zeta_i} (p_i^s - \tilde{p}^s)q(p^b(p_i^s);\theta_i,\zeta_i)dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\\ &\quad + \int_{p_i^s,\theta_i,\zeta_i} (p_i^s - \tilde{p}^s)q(p_i^s;\theta_i,\zeta_i)dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)\end{aligned}$$

We can show from lemma 2 and $p_i^s \in [\bar{p}, \bar{p} + \bar{m}\tau] = \mathcal{P}\ \forall i$ that the term on the second line is non-negative. Formally for any $p_i^s \in (\bar{p}, \bar{p} + \bar{m}\tau), \theta_i, \zeta_i, \zeta_i'$:

$$p_i^s > \tilde{p}^s \Rightarrow p^b(p_i^s) > \tilde{p}^s \Rightarrow q(p^b(p_i^s);\theta_i,\zeta_i') \leq q(p_i^s;\theta_i,\zeta_i)$$

$$p_i^s \leq \tilde{p}^s \Rightarrow p^b(p_i^s) < \tilde{p}^s \Rightarrow q(p^b(p_i^s);\theta_i,\zeta_i') \geq q(p_i^s;\theta_i,\zeta_i)$$

Either way:

$$(p_i^s - \tilde{p}^s)q(p^b(p_i^s);\theta_i,\zeta_i') \leq (p_i^s - \tilde{p}^s)q(p_i^s;\theta_i,\zeta_i)$$

Thus:

$$\int_{p_i^s,\theta_i,\zeta_i} (p_i^s - \tilde{p}^s) q(p_i^s;\theta_i,\zeta_i) dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$= \int_{p_i^s \in int(\mathcal{P}),\theta_i,\zeta_i} (p_i^s - \tilde{p}^s) q(p_i^s;\theta_i,\zeta_i) dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$+ \int_{p_i^s \in \{\bar{p},\bar{p}+\bar{m}\tau\},\theta_i,\zeta_i} (p_i^s - \tilde{p}^s) q(p_i^s;\theta_i,\zeta_i) dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$\geq \int_{p_i^s \in int(\mathcal{P}),\theta_i,\zeta_i} (p_i^s - \tilde{p}^s) q(p^b(p_i^s);\theta_i,\zeta_i) dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$+ \int\limits_{p_i^s \in \{\bar{p},\bar{p}+\bar{m}\tau\},\theta_i,\zeta_i} (p_i^s - \tilde{p}^s) q(p_i^s;\theta_i,\zeta_i) dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$= \int_{p_i^s,\theta_i,\zeta_i} (p_i^s - \tilde{p}^s) q(p_i^s;\theta_i,\zeta_i) dF''_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

$$\int\limits_{p_i^s,\theta_i,\zeta_i} dwl(p^b(p_i^s);\theta_i,\zeta_i) dF_{p^s,\theta}(p_i^s,\theta_i) \geq \int\limits_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i) dF_{p^s,\theta}(p_i^s,\theta_i,\zeta_i)$$

□

Before stating and proving theorem 3, we note that deadweight loss is bounded by the product of the reduction in demand and $\bar{m}\tau$.

Lemma 4. *If* $p_i^s \in [\bar{p}, \bar{p}+\bar{m}\tau]$, *then* $dwl(p_i^s;\theta_i,\zeta_i) \leq [q(\bar{p};\theta_i,\zeta_i) - q(p_i^s;\theta_i,\zeta_i)]\bar{m}\tau \ \forall \theta_i, \zeta_i$.

Proof: Using lemma 3:

$$0 = dwl(\bar{p};\theta_i,\zeta_i) \geq dwl(p_i^s;\theta_i,\zeta_i) - (p_i^s - \bar{p})[q(\bar{p};\theta_i,\zeta_i) - q(p_i^s;\theta_i,\zeta_i)]$$

$$dwl(p_i^s;\theta_i,\zeta_i) \leq [q(\bar{p};\theta_i,\zeta_i) - q(p_i^s;\theta_i,\zeta_i)](p_i^s - \bar{p}) \leq [q(\bar{p};\theta_i,\zeta_i) - q(p_i^s;\theta_i,\zeta_i)]\bar{m}\tau.$$

□

Next, we state and prove theorem 3. It says that the maximal value of deadweight loss consistent with the data and knowledge of F_θ^* is given by having some agents perceive the highest possible price and some others perceive the lowest possible price. It achieves this by assigning the good there where it will generate the most deadweight loss, while holding aggregate demand constant. The resulting demand function, $\tilde{q}_{\Delta,\gamma}(\theta_i)$, is such that those for whom the ratio of deadweight loss[5] to change in quantity exceeds Δ perceive the highest price, those with such a ratio below Δ perceive the lowest possible price, and those with ratio equal to Δ are split between perceiving the high and low price in a way that rationalizes demand. Those who perceive the high (low) price consume the least (most) possible consistent with their perceptions.

[5]Note that when $p^s = \bar{p}$, the tax is effectively lump-sum and so there is no deadweight loss.

Theorem 3. *There exist values* $\Delta \in [0, \bar{m}\tau]$ *and* $\gamma \in [0,1]$ *such that:*

$$\int_{p_i^s,\theta_i} \tilde{q}_{\Delta,\gamma}(\theta_i) dF^*_{p^s,\theta_i} = \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

where:

$$\tilde{q}_{\Delta,\gamma}(\theta_i) =$$
$$\left[\mathbb{I}\left(\frac{dwl(\bar{p}+\bar{m}\tau;\theta_i,l)}{q(\bar{p};\theta_i,h)-q(\bar{p}+\bar{m}\tau;\theta_i,l)} > \Delta\right) + \gamma\mathbb{I}\left(\frac{dwl(\bar{p}+\bar{m}\tau;\theta_i,l)}{q(\bar{p};\theta_i,h)-q(\bar{p}+\bar{m}\tau;\theta_i,l)} = \Delta\right)\right]$$
$$\times\, q(\bar{p}+\bar{m}\tau;\theta_i,l)$$
$$+\left[\mathbb{I}\left(\frac{dwl(\bar{p}+\bar{m}\tau;\theta_i,l)}{q(\bar{p};\theta_i,h)-q(\bar{p}+\bar{m}\tau;\theta_i,l)} < \Delta\right) + (1-\gamma)\mathbb{I}\left(\frac{dwl(\bar{p}+\bar{m}\tau;\theta_i,l)}{q(\bar{p};\theta_i,h)-q(\bar{p}+\bar{m}\tau;\theta_i,l)} = \Delta\right)\right]$$
$$\times\, q(\bar{p};\theta_i,h)$$

Of course, if $q(\bar{p};\theta_i,h) = q(\bar{p}+\bar{m}\tau;\theta_i,l)$*, then* $\tilde{q}_{\Delta,\gamma}(\theta_i) = q(\bar{p};\theta_i,h)$*. Furthermore, under assumption 1, for any* $F_{p^s,\theta,\zeta}$ *that rationalizes the data such that* $F_\theta = F^*_\theta$*:*

$$\int_{p_i^s,\theta_i} \frac{\tilde{q}_{\Delta,\gamma}(\theta_i) - q(\bar{p};\theta_i,h)}{q(\bar{p}+\bar{m}\tau;\theta_i,l)-q(\bar{p};\theta_i,h)} dwl(\bar{p}+\bar{m}\tau;\theta_i,l) dF^*_{p^s,\theta}(p_i^s,\theta_i)$$
$$\geq \int_{p_i^s,\theta_i,\zeta_i} dwl(p_i^s;\theta_i,\zeta_i) dF_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

where the integrand on the left-hand side is defined as zero for any θ_i *such that* $q(\bar{p}+\bar{m}\tau;\theta_i,l) = q(\bar{p};\theta_i,h)$*.*

The intuition is for the Δ term is straightforward. The econometrician observes the reduction in aggregate demand due to the tax. In searching for the explanation of that reduction in demand that maximizes deadweight loss, one should assign the reduction in quantity demanded to those for whom that allocation yields the greatest deadweight loss. Following this procedure, there is a cutoff value Δ which describes the amount of deadweight loss obtained relative to the reduction in quantity demanded sufficient to warrant the assignment of subjective tax-inclusive price $p_i^s = \bar{p} + \bar{m}\tau$ to that agent.

The idea behind the tie-breaking provision is that those individuals who perceive the high price should reduce their consumption as much as possible to maximize deadweight loss; those who perceive the sticker price should maximize their consumption to permit even more individuals to perceive the high price.

Proof of Theorem 3: The outline of the proof is as follows. First, we use lemma 4 to show that the maximal deadweight loss consistent with aggregate demand and F^*_θ comes from a data-generating process in which agents perceiving the price $\bar{p} + \bar{m}\tau$ choose the lowest quantity consistent with preference maximization, whereas the other agents choose the largest such quantity. Then, we show that distributions satisfying such a property yield deadweight loss no larger than the proposed distribution, which exists.

First, consider an arbitrary distribution $F_{p^s,\theta,\zeta}$ (yielding well-defined aggregate demand and deadweight loss) such that $F_\theta = F_\theta^*$ and:

$$F_{p^s} = \begin{cases} 0 & p_i^s < \bar{p} \\ F_{p^s}(\bar{p}) & p_i^s \in [\bar{p}, \bar{p} + \bar{m}\tau) \\ 1 & p_i^s \geq \bar{p} + \bar{m}\tau \end{cases}$$

In words, the above expression says that the support of p^s is contained in $\{\bar{p}, \bar{p} + \bar{m}\tau\}$. By theorem 2, the maximal value of deadweight loss consistent with aggregate demand and F_θ^* must satisfy this property. Consider some value $\rho \in [0, 1]$ such that:

$$\begin{aligned} &\rho \int_{\theta_i} q(\bar{p} + \bar{m}\tau; \theta_i, l) dF_{\theta|p^s \neq \bar{p}}(\theta_i) + [1 - \rho] \int_{\theta_i} q(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i) \\ &= \int_{p_i^s, \theta_i, \zeta_i} q(p_i^s; \theta_i, \zeta_i) dF_{p^s, \theta, \zeta}(p_i^s, \theta_i, \zeta_i) \end{aligned} \tag{B.2}$$

Such a value of ρ must exist by the Intermediate Value Theorem, since by the definition of l and h and the CLD as expressed in lemma 2:

$$\begin{aligned} \int_{\theta_i} q(\bar{p} + \bar{m}\tau; \theta_i, l) dF_{\theta|p^s \neq \bar{p}}(\theta_i) &\leq \int_{p_i^s, \theta_i, \zeta_i} q(p_i^s; \theta_i, \zeta_i) dF_{p^s, \theta, \zeta}(p_i^s, \theta_i, \zeta_i) \\ &\leq \int_{\theta_i} q(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i) \end{aligned}$$

In words, we are constructing an alternative distribution that rationalizes aggregate demand such that $p^s = \bar{p} + \bar{m}\tau$ and $\zeta = l$ with probability ρ, and otherwise $p^s = \bar{p}$ and $\zeta = h$. We now show that this alternate distribution yields at least as much deadweight loss, thus showing that the maximal value of deadweight loss consistent with aggregate demand and F_θ^* must arise from a distribution in which almost surely $(p^s, \zeta) = (\bar{p}, h)$ or $(p^s, \zeta) = (\bar{p} + \bar{m}\tau, l)$.

From the definition of deadweight loss:

$$\begin{aligned} &\int_{\theta_i, \zeta_i} \bar{m}\tau[q(\bar{p}; \theta_i, \zeta_i) - q(\bar{p}; \theta_i, l)] dF_{\theta, \zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i) \\ &\qquad = \int_{\theta_i, \zeta_i} [dwl(\bar{p} + \bar{m}\tau; \theta_i, l) - dwl(\bar{p}; \theta_i, \zeta_i)] dF_{\theta, \zeta|p^s \neq \bar{p}} \end{aligned}$$

From here, the definition of l, and using the fact that $dwl(\bar{p}; \theta_i, \zeta_i) = 0 \ \forall \theta_i, \zeta_i$, we have that $\rho \geq 1 - F_{p^s}(\bar{p})$ implies that:

$$\rho \int_{\theta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, l) dF_{\theta|p^s \neq \bar{p}}(\theta_i) + (1-\rho) \int_{\theta_i} dwl(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i)$$

$$= \rho \int_{\theta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, l) dF_{\theta|p^s \neq \bar{p}}(\theta_i)$$

$$\geq [1 - F_{p^s}(\bar{p})] \int_{\theta_i, \zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$= [1 - F_{p^s}(\bar{p})] \int_{\theta_i, \zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$+ F_{p^s}(\bar{p}) \int_{\theta_i, \zeta_i} dwl(\bar{p}; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s = \bar{p}}(\theta_i, \zeta_i)$$

$$= \int_{\theta_i, \zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta}(\theta_i, \zeta_i)$$

Where the inequality follows from the fact that $\rho \geq 1 - F_{p^s}(\bar{p})$ by assumption, and the fact that $dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i)$ and the definition of l. This shows that whenever $\rho \geq 1 - F_{p^s}(\bar{p})$, the proposed alternative distribution yields at least as much deadweight loss. Now suppose instead $\rho < 1 - F_{p^s}(\bar{p})$. From lemma 4:

$$\int_{\theta_i, \zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$\geq \bar{m}\tau \int_{\theta_i, \zeta_i} [q(\bar{p}; \theta_i, h) - q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i)] \bar{m}\tau dF_{\theta_i,\zeta_i|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

In addition, we find it convenient to rewrite the aggregate demand-rationalizing equation as:

$$\int_{p_i^s, \theta_i, \zeta_i} q(p_i^s; \theta_i, \zeta_i) dF_{p^s,\theta,\zeta}(p_i^s, \theta_i, \zeta_i)$$

$$= [1 - F_{p^s}(\bar{p})] \int_{\theta_i, \zeta_i} q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$+ F_{p^s}(\bar{p}) \int_{\theta_i, \zeta_i} q(\bar{p}; \theta_i, \zeta_i) dF_{\theta|p^s = \bar{p}}(\theta_i, \zeta_i)$$

And so, using equation B.2 and rearranging terms,

$$\rho \int_{\theta_i, \zeta_i} [q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) - q(\bar{p} + \bar{m}\tau; \theta_i, l)] dF_{\theta,\zeta}(\theta_i, \zeta_i)$$

$$= (1-\rho) \int_{\theta_i} q(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i)$$

$$- [1 - F_{p^s}(\bar{p}) - \rho] \int_{\theta_i, \zeta_i} q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$- F_{p^s}(\bar{p}) \int_{\theta_i, \zeta_i} q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

Thus, plugging in and using lemma 4:

$$\rho \int\limits_{\theta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, l) dF_{\theta|p^s \neq \bar{p}}(\theta_i) + (1-\rho) \int_{\theta_i} dwl(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i)$$

$$= \rho \int_{\theta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, l) dF_{\theta|p^s \neq \bar{p}}(\theta_i)$$

$$= \rho \int_{\theta_i,\zeta_i} \left[dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) - q(\bar{p} + \bar{m}\tau; \theta_i, l) + q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) \right] dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$= \rho \int_{\theta_i,\zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$+ \rho \int_{\theta_i,\zeta_i} [q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) - q(\bar{p} + \bar{m}\tau; \theta_i, l)] dF_{\theta,\zeta}(\theta_i, \zeta_i)$$

$$= \rho \int_{\theta_i,\zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i) + (1-\rho) \int_{\theta_i} q(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i)$$

$$- [1 - F_{p^s}(\bar{p}) - \rho] \int_{\theta_i,\zeta_i} q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$- F_{p^s}(\bar{p}) \int_{\theta_i,\zeta_i} q(\bar{p}; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s = \bar{p}}(\theta_i, \zeta_i)$$

$$= \rho \int_{\theta_i,\zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i) + F_{p^s}(\bar{p}) \int_{\theta_i} q(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i)$$

$$+ [1 - F_{p^s}(\bar{p}) - \rho] \int_{\theta_i,\zeta_i} \left[q(\bar{p}; \theta_i, h) - q(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) \right] dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$- F_{p^s}(\bar{p}) \int_{\theta_i,\zeta_i} q(\bar{p}; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s = \bar{p}}(\theta_i, \zeta_i)$$

$$\geq \rho \int_{\theta_i,\zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i) + F_{p^s}(\bar{p}) \int_{\theta_i} q(\bar{p}; \theta_i, h) dF_{\theta|p^s = \bar{p}}(\theta_i)$$

$$+ [1 - F_{p^s}(\bar{p}) - \rho] \bar{m}\tau \int_{\theta_i,\zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$- F_{p^s}(\bar{p}) \int_{\theta_i,\zeta_i} q(\bar{p}; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s = \bar{p}}(\theta_i, \zeta_i)$$

$$\geq \rho \int_{\theta_i,\zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$+ [1 - F_{p^s}(\bar{p}) - \rho] \bar{m}\tau \int_{\theta_i,\zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$= [1 - F_{p^s}(\bar{p})] \int\limits_{\theta_i,\zeta_i} dwl(\bar{p} + \bar{m}\tau; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s \neq \bar{p}}(\theta_i, \zeta_i)$$

$$+ F_{p^s}(\bar{p}) \int\limits_{\theta_i,\zeta_i} dwl(\bar{p}; \theta_i, \zeta_i) dF_{\theta,\zeta|p^s = \bar{p}}(\theta_i, \zeta_i)$$

Thus, we know that the maximal deadweight loss consistent with aggregate demand and F_θ^* is generated by a distribution in which with probability one either $(p^s, \zeta) = (\bar{p}, h)$ or $(p^s, \zeta) = (\bar{p} + \bar{m}\tau, l)$. We refer to distributions of this sort as *binary* distributions.

Now, we show that the proposed distribution maximizes deadweight loss among all binary distributions, and thus among all distributions, that rationalize aggregate demand such that $F_\theta = F_\theta^*$. Towards that end, we first show that the proposed distribution exists. Note by lemma 4 and the CLD as in lemma 2:

$$\int\limits_{\theta_i} \tilde{q}_{\bar{m}\tau,1}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i) \leq \int\limits_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$
$$\leq \int\limits_{\theta_i} \tilde{q}_{0,0}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i)$$

In words, aggregate demand is contained between when all agents perceive a high price and have type h and when all agents perceive a low price and have type l. Furthermore, one can confirm that for any $\Delta, \Delta', \gamma, \gamma'$ such that $0 \leq \Delta < \Delta' \leq \bar{m}\tau$ and $0 \leq \gamma < \gamma' \leq 1$:

$$\int_{\theta_i} \tilde{q}_{\Delta,\gamma'}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i) \leq \int_{\theta_i} \tilde{q}_{\Delta,\gamma}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i)$$

$$\int_{\theta_i} \tilde{q}_{\Delta',\gamma'}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i) \geq \int_{\theta_i} \tilde{q}_{\Delta,\gamma}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i)$$

Thus, we can pick Δ such that:

$$\int\limits_{\theta_i} \tilde{q}_{\Delta,1}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i) \leq \int\limits_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$
$$\leq \int\limits_{\theta_i} \tilde{q}_{\Delta,0}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i)$$

If both sides hold with equality, we can define γ arbitrarily. Otherwise, we define γ so that the market clears:

$$\gamma \equiv \frac{\int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i) - \int_{\theta_i} \tilde{q}_{\Delta,0}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i)}{\int_{\theta_i} \tilde{q}_{\Delta,1}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i) - \int_{\theta_i} \tilde{q}_{\Delta,0}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i)}$$

We now have the values Δ and γ such that the market clears. Suppressing Δ and γ subscripts from $\tilde{q}$, we can say that:

$$\int_{\theta_i} \tilde{q}(\theta_i) dF^*_{p^s,\theta}(p_i^s,\theta_i) = \int_{p_i^s,\theta_i,\zeta_i} q(p_i^s;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p_i^s,\theta_i,\zeta_i)$$

Finally, to show that the proposed distribution maximizes deadweight loss, consider arbitrary binary distribution $F_{p^s,\theta,\zeta}$ that rationalizes aggregate demand. Defining $\mathbb{P}_F(p^s \neq \bar{p}|\theta_i) \equiv 1 - F_{p^s|\theta=\theta_i}(\bar{p} + \bar{m}\tau)$ as the probability that $(p^s,\zeta) = (\bar{p} + \bar{m}\tau, l)$ conditional on θ_i, rationalizing aggregate demand with $F_\theta = F_\theta^*$ means that:

$$\int_{\theta_i} \big[\mathbb{P}_F(p^s \neq \bar{p}|\theta_i)\ q(\bar{p}+\bar{m}\tau;\theta_i,l) + F_{p^s|\theta=\theta_i}(\bar{p})\ q(\bar{p};\theta_i,h)\big] dF^*_\theta(\theta_i)$$
$$= \int_{p^s_i,\theta_i,\zeta_i} q(p^s_i;\theta_i,\zeta_i) dF^*_{p^s,\theta,\zeta}(p^s_i,\theta_i,\zeta_i)$$

We can now write the difference in generated values of aggregate deadweight loss as:

$$\begin{aligned}
&\int_{\theta_i} \big[\frac{\tilde{q}(\theta_i) - q(\bar{p};\theta_i,h)}{q(\bar{p}+\bar{m}\tau;\theta_i,l) - q(\bar{p};\theta_i,h)} - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)\big] dwl(\bar{p}+\bar{m}\tau;\theta_i,l) dF^*_\theta(\theta_i) \\
&= \int_{\mathcal{X}} [1 - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)] dwl(\bar{p}+\bar{m}\tau;\theta_i) dF^*_\theta(\theta_i) \\
&+ \int_{\mathcal{Y}} [\gamma - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)] dwl(\bar{p}+\bar{m}\tau;\theta_i) dF^*_\theta(\theta_i) \\
&- \int_{\mathcal{Z}} \mathbb{P}_F(p^s \neq \bar{p}|\theta_i) dwl(\bar{p}+\bar{m}\tau;\theta_i) dF^*_\theta(\theta_i) \\
&\geq \Delta \int_{\mathcal{X}} [1 - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)]\ [q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)] dF^*_\theta(\theta_i) \\
&+ \Delta \int_{\mathcal{Y}} [\gamma - \mathbb{P}_F(p^s \neq \bar{p}|\theta_i)]\ [q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)] dF^*_\theta(\theta_i) \\
&- \Delta \int_{\mathcal{Z}} \mathbb{P}_F(p^s \neq \bar{p}|\theta_i) [q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)] dF^*_\theta(\theta_i)
\end{aligned}$$

Where

$$\begin{aligned}
\mathcal{X} &\equiv \{\theta_i : dwl(\bar{p}+\bar{m}\tau) > \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]\} \\
\mathcal{Y} &\equiv \{\theta_i : dwl(\bar{p}+\bar{m}\tau) = \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]\} \\
\mathcal{Z} &\equiv \{\theta_i : dwl(\bar{p}+\bar{m}\tau) < \Delta[q(\bar{p};\theta_i,h) - q(\bar{p}+\bar{m}\tau;\theta_i,l)]\}
\end{aligned}$$

We complete the proof by showing the right-hand side of the last inequality is zero. Since both distributions rationalize the same aggregate demand:

$$\begin{aligned}
&\int_{\theta_i:\mathcal{X}} q(\bar{p}+\bar{m}\tau;\theta_i,l) dF^*_\theta(\theta_i) \\
&+ \int_{\mathcal{Y}} [\gamma q(\bar{p}+\bar{m}\tau;\theta_i,l) + (1-\gamma) q(\bar{p};\theta_i,h)] dF^*_\theta(\theta_i) \\
&+ \int_{\mathcal{Z}} q(\bar{p};\theta_i,h) dF^*_\theta(\theta_i) \\
&= \int_{\theta_i} \big[\mathbb{P}_F(p^s \neq p^s|\theta_i)\ [q(\bar{p}+\bar{m}\tau;\theta_i,l) - q(\bar{p};\theta_i,h)] + q(\bar{p};\theta_i,h)\big] dF^*_\theta(\theta_i)
\end{aligned}$$

Subtracting both sides from $\int_{\theta_i} q(\bar{p};\theta_i,h)dF_\theta^*(\theta_i)$ yields:

$$\begin{aligned}&\int_{\mathcal{X}}[q(\bar{p};\theta_i,h)-q(\bar{p}+\bar{m}\tau;\theta_i,l)]dF_\theta^*(\theta_i)\\&+\int_{\mathcal{Y}}\gamma[q(\bar{p};\theta_i,h)-q(\bar{p}+\bar{m}\tau;\theta_i,l)]dF_\theta^*(\theta_i)\\&=\int_{\theta_i}\mathbb{P}_F(p^s\neq p^s|\theta_i)\ [q(\bar{p}+\bar{m}\tau;\theta_i,l)-q(\bar{p};\theta_i,h)]dF_\theta^*(\theta_i)\end{aligned}$$

Finally, subtracting the right-hand side from the left-hand size and multiplying by zero yields the desired result. Thus:

$$\int_{\theta_i}\big[\frac{\tilde{q}(\theta_i)-q(\bar{p};\theta_i,h)}{q(\bar{p}+\bar{m}\tau;\theta_i,l)-q(\bar{p};\theta_i,h)}-\mathbb{P}_F(p^s\neq\bar{p}|\theta_i)\big]dwl(\bar{p}+\bar{m}\tau;\theta_i,l)dF_\theta^*(\theta_i)=0$$

In words, deadweight loss from the proposed distribution is at least as great as the deadweight loss from any binary distribution that also rationalizes aggregate demand and with the true distribution of preference types. From the first part of the proof, any distribution that rationalized aggregate demand and had the support of perceived prices contained in $\partial d\mathcal{P}$ yielded deadweight loss no greater than what one could obtain with a binary distribution that rationalized aggregate demand with $F_\theta = F_\theta^*$. Theorem 2 noted that any distribution that rationalized aggregate demand with $F_\theta = F_\theta^*$ yielded deadweight loss no greater than that one could obtain with a distribution that had the support of perceived prices contained in $\partial d\mathcal{P}$, rationalized aggregate demand, and had $F_\theta = F_\theta^*$. Therefore, any distribution that rationalizes aggregate demand and with $F_\theta = F_\theta^*$ yields deadweight loss no greater than the proposed distribution. □

B.2. Details on Application of Linear Model

We use data gathered by CLK (2009) on the aggregate consumption of beer in U.S. states between 1970 and 2003, and cross-sectional data gathered by Goldin and Homonoff (2013) on tobacco consumption between 1984 and 2000. We translate their two models (in logs) to our linear specification. In particular, we are interested in estimating models like the one in equation 2.8.

In the case of aggregate beer consumption, we follow CLK (2009) in using a specification in first differences, and so we estimate regressions of the type:

$$\Delta y_{st}=\alpha+\beta\Delta\tau_{st}^e+\tilde{\beta}\Delta\tau_{st}^s+\gamma X_{st}+\epsilon_{st}$$

where y_{st} represents per-capita consumption of beer, in gallons, for state s at time t, τ^e represents excise taxes on beer (included in sticker price), τ^s represents sales taxes (non-salient), X is a vector of controls, and ϵ is an i.i.d. error term. All taxes are expressed in dollar amounts.

For each linear specification, we compute $\hat{m}=\frac{\tilde{\beta}}{\beta}$, which gives us the ratio of upper bound of deadweight loss to lower bound of deadweight loss (assuming that maximal attention, $\bar{m}=1$). Results are presented in table B.1. We also estimate a number of other specifications, again following CLK (2009), presented in table B.2.

These are meant to address concerns for spurious results – in particular, it could be the case that consumers react differently to the two tax rates because while sales taxes affect a variety of goods, excise taxes on beer affect only beer prices. The second last column of table B.2 shows estimates for a regression only for those states that exempt food (a likely substitute of beer) from sales tax, demonstrating that even in this restricted sample beer consumption is quite insensitive to sales tax. Finally, the last column addresses the potential concern that people might be substituting toward other alcoholic beverages when they face a beer tax increase, and not when they face a sales tax increase. As we can see, the share of ethanol people consume in the form of beer is quite insensitive to either tax rate.

We repeat the exercise for Goldin and Homonoff (2013), who have a similar set-up with individual-level, cross-sectional data on cigarette consumption. Even though this is not aggregate data, estimating a linear model that only measures average effects effectively leaves the analysis of section 2.4 unchanged. We again follow the original authors of the paper when we estimate the equation:

$$c_{ist} = \alpha + \beta\tau^e_{st} + \tilde{\beta}\tau^s_{st} + \gamma X_{st} + \delta Z_{ist} + \varepsilon_{ist}$$

where now c_{ist} stands for tobacco consumption, in average cigarettes per day, for individual i from state s in period t, τ^e, τ^s, and X_{st} should be interpreted as before, and Z_{ist} is a vector of individual-level controls. All the details can be found in the original paper. Results in table B.3 showcase a number of different specifications, including several sets of fixed-effects, all following Goldin and Homonoff (2013).

	Baseline	Business cycle	Alcohol regulations	Region trends
Δ(excise tax)	-0.966	-0.875	-0.808	-0.715
	(0.4)	(0.393)	(0.394)	(0.394)
Δ(sales tax)	-0.305	-0.113	-0.114	-0.241
	(0.708)	(0.698)	(0.699)	(.7)
Δ(population)	-0.0002	-.0002	-0.0001	-0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Δ(income per cap.)		0.0002	0.0001	0.0002
		(0.00006)	(0.00006)	(0.00006)
Δ(unemployment)		-.094	-0.093	-0.093
		(.026)	(0.026)	(0.026)
Alcohol reg. controls			X	X
Year FE	X	X	X	X
Region FE				X
$\hat{m}$	0.316	0.129	0.141	0.338
	(0.743)	(0.8)	(0.866)	(0.996)
Sample size	1,607	1,487	1,487	1,487

Table (B.1) Estimating $\hat{m}$ with several sets of controls, following the specifications in CLK (2009) in the context of a linear model. Standard errors in parentheses.

	Policy IV for excise tax	3-Year differences	Food exempt	Dep. var.: share of ethanol from beer
Δ(excise tax)	-0.808	-2.092	-1.114	0.036
	(0.395)	(0.897)	(1.174)	(0.006)
Δ(sales tax)	-0.114	-0.131	-0.449	0.018
	(0.699)	(0.826)	(0.757)	(0.011)
Δ(population)	-0.0001	-0.002	-0.00007	0.0000
	(0.0002)	(0.002)	(.0002)	(0.0000)
Δ(income per cap.)	0.0001	0.0002	0.0001	-0.0000
	(.00006)	(0.00007)	(0.00007)	(0.0000)
Δ(unemployment)	-0.094	-0.03	-0.056	-0.0001
	(.026)	(0.028)	(.032)	(0.0004)
Alcohol reg. controls	X	X	X	X
Year FE	X	X	X	X
$\hat{m}$	0.141	0.062	.403	
	(0.866)	(0.395)	(0.819)	
Sample size	1,487	1,389	937	1,487

Table (B.2) Estimating $\hat{m}$ following the strategy of CLK (2009) in the context of a linear model. As in CLK, we use the nominal excise tax rate divided by the average price of a case of beer from 1970 to 2003 as an IV for excise tax to eliminate tax-rate variation coming from inflation erosion. Next, we run the same regression in 3-year differences. Next, we run it only for states where food is exempt from sales-tax, to address concerns about whether consumers react differently to changes in the two taxes only because sales taxes apply to a broad set of goods. Finally, the last column addresses the concern that beer taxes may induce substitution with other alcoholic products, biasing the coefficient on excise tax relative to the one on sales tax. While in the log-log specification of CLK (2009) it seems to show that beer excise taxes have no discernable effect on the share of ethanol consumed from beer, we do find a significant effect. Standard errors in parentheses.

	Outcome variable: Number of cigarettes		
Specification	1	2	3
Excise Tax	-0.015	-0.015	-0.016
	(.004)	(.004)	(.004)
Sales Tax	-0.024	-0.02	-0.022
	(0.022)	(0.025)	(0.025)
Demographic controls	X	X	X
Econ. conditions controls		X	X
Income trend controls			X
State,year, and month FE	X	X	X
$\hat{m}$	1.57	1.33	1.37
	(1.65)	(1.83)	(1.82)
Sample size	274,138	274,138	274,138

Table (B.3) Estimating $\hat{m}$ based on the intensive response of cigarette consumption to sales taxes (not included in sticker price) and excise taxes (included in the sticker price). The specifications are a linearized version of the specifications in Goldin and Homonoff (2013). Standard errors in parentheses.

APPENDIX C

Appendix for Chapter 3

C.1. Adjustment Costs Example

For simplicity, we consider a tax on revenue in this section. An investment x yields return $\epsilon f(x)$, where $f(x) = 2\sqrt{x}$. Suppose that with probability one $\epsilon = 1$. In addition, $\tau = \frac{1}{2}$ with probability one. Then the firm optimally sets $x = \frac{1}{2}$. It never needs to adjust investments thereafter because τ and ϵ are known with probability one.

However, suppose instead that with probability p it turns out that $\epsilon = 1/p$, and with probability $1 - p$ it turns out that $\epsilon = 0$. In addition, we let $\tau = 1$ if $\epsilon = 0$, while $\tau = 0$ otherwise. Thus, there is now a negative covariance between τ and ϵ, whereas before there was no covariance. Finally, after the realizations of τ and ϵ, there is a fixed cost F of adjusting capital.

In this new hypothetical, so long as $p < 0.5$, the firm will never choose $x = (1/p)^2$. That option is dominated by $x = 0$, as this option makes the likelihood of adjusting and paying the fixed adjustment cost F not as likely. The other possibility is that the firm chooses $x = 1$ and does not adjust at all.[1] The firm's expected profit from the strategy of not adjusting is 1, whereas the expected profit from choosing $x = 0$ and adjusting to $(1/p)^2$ if $\epsilon = 1/p$ is $(1/p)^2 - F$. So long as $p < \min\{\frac{1}{F+1}, 0.5\}$, the firm optimizes by choosing $x = 0$. Thus, with adjustment costs, decreasing the covariance between taxes and productivity can decrease investment.

C.2. Data Appendix

In this section we provide details about the data used in our empirical analysis. Figures C.1, C.2, and C.3 plot the evolution of our main measures of policy uncertainty over the respective sample periods. Table C.1 refers summary statistics about the two samples underlying the naked CAPM and the Fama-French CAPM models estimated in our empirical analysis. Table C.2 shows correlations between our different measures of policy uncertainty.

[1] If the firm does not adjust, then it maximizes $p * \frac{2\sqrt{x}}{p} - x$.

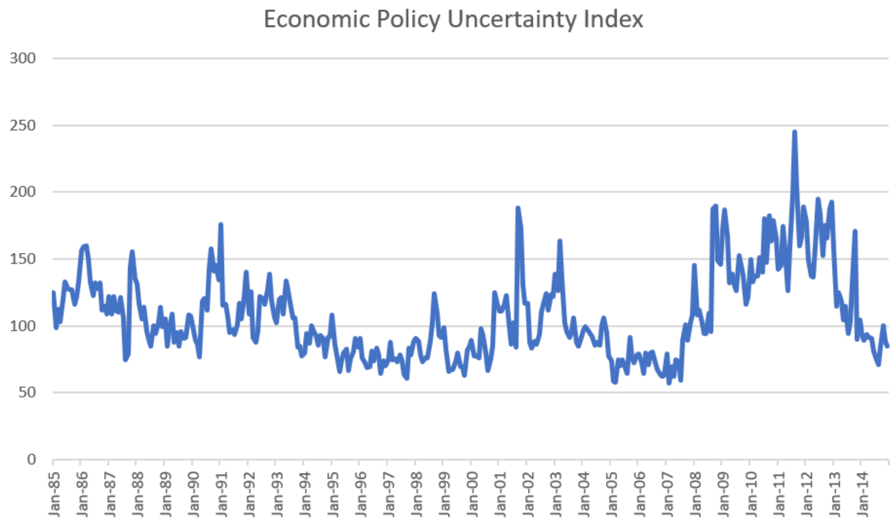


Figure (C.1) Economic Policy Uncertainty Index from Baker, Bloom and Davis (2016).

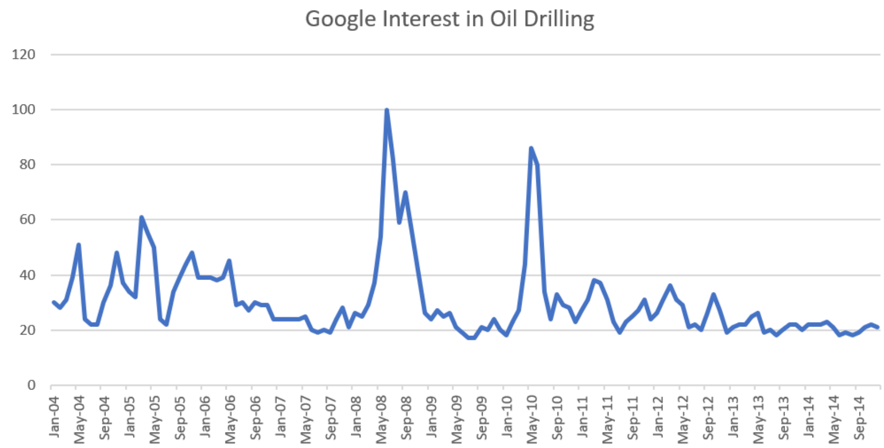


Figure (C.2) Google search interest in “oil drilling”.

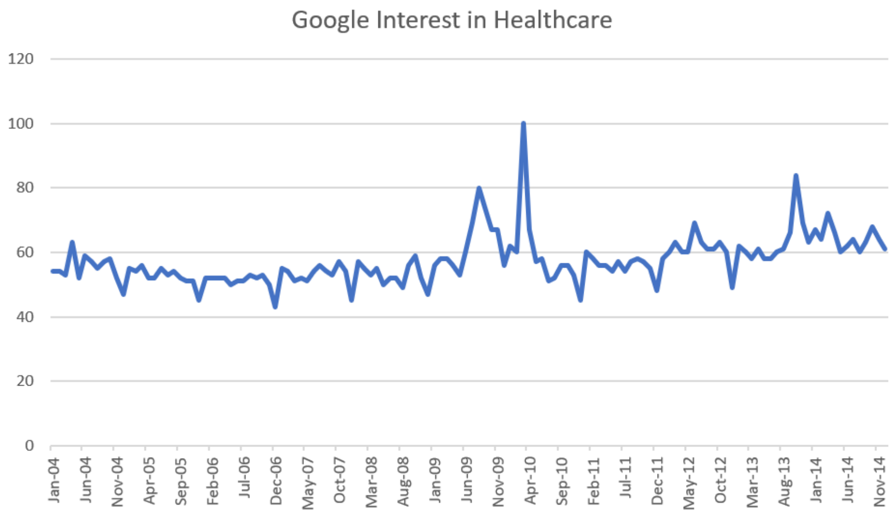


Figure (C.3) Google search interest in “healthcare”.

	Naked	Fama-French
Number of firms	690	556
Number of periods	360	120
Market value in $1000's	7310.31 (25393.46)	8442.63 (27846.81)
Multinational	0.928	0.927
Energy	0.068	0.065
Materials	0.078	0.070
Industrials	0.235	0.239
Consumer disc	0.125	0.128
Consumer staples	0.067	0.064
Health	0.070	0.077
Finance	0.109	0.088
IT	0.117	0.121
Telecom	0.007	0.009
Utilities	0.088	0.098
Real Estate	0.036	0.040

Table (C.1) Summary Statistics for the two samples corresponding to the two specifications of interest.

	oil drilling	healthcare	EPU
oil drilling	1.000		
healthcare	−0.2672	1.000	
EPU	−0.0623	0.2336	1.000

Table (C.2) Correlations between measures of policy uncertainty.