The Cost of Regulatory Compliance in the United States

Francesco Trebbi Miao Ben Zhang Michael Simkovic*

August 2025

Abstract

A key question for studying business dynamism is whether the costs of regulatory compliance fall homogeneously on small and large businesses. Using comprehensive establishment-level occupational microdata and occupation task information, we quantify a firm's compliance costs as the share of wage bill for performing regulatory compliance tasks (RegIndex). We reveal an inverted-U relation between firms' RegIndex and their size, with 500-employment firms facing compliance costs 40 percent higher as a share of total wages than small or large firms. We further develop a shift-share methodology to disentangle the influence of regulatory requirements and enforcement on driving firms' compliance costs.

Keywords: Regulation, Compliance Costs, Labor Task, Occupation, Firm Growth, Economies of Scale, Regulatory Requirement, Enforcement

*Trebbi: University of California Berkeley Haas School of Business, CEPR, and NBER and can be reached at ftrebbi@berkeley.edu; Zhang: University of Southern California Marshall School of Business and can be reached at miao.zhang@marshall.usc.edu; Simkovic: University of Southern California Gould School of Law and can be reached at msimkovic@law.usc.edu. We would like to thank the editor, the associate editor, and two anonymous referees for helpful suggestions. We also thank Matt Backus, Andrew Baker, Matilde Bombardini, Greg Buchak, Julieta Caunedo, Ernesto Dal Bó, Steven Davis, Klaus Desmet, Fred Finan, Bård Harstad, Joseph Kalmenovitz, Alessandro Lizzeri, John Matsusaka, Adair Morse, Tom Palfrey, Amit Seru, Steven Davidoff Solomon, James Spindler, Reed Walker, Jessie Wang, Joshua White, and seminar participants at the NBER Corporate Finance Fall Meeting 2023, NBER Economic Analysis of Regulation Conference, NBER SI Law and Economics 2024, Tepper-LAEF Conference, WFA, PSL Public Governance Seminar, Berkeley Political Economy Research Lab, Arizona State University, Northwestern University, Stanford GSB, University of California College of Law, University of Michigan, University of Southern California, University of Texas at Austin, Vanderbilt University, and other institutions for their comments. Niranjan Navaneethan provided excellent research assistance. This research was conducted when Miao Ben Zhang was a visiting researcher at the Bureau of Labor Statistics (BLS). The views expressed here are those of the authors and do not necessarily reflect the views of the BLS. We thank Erin Good, Jessica Helfand, Michael Horrigan, Mark Loewenstein, and especially Michael Soloy at the BLS for their excellent assistance. A separate, unpublished manuscript, "Regulation Intensity and Technology-Driven Entry" by Michael Simkovic and Miao Ben Zhang, which described some of the aggregate public BLS data, is now retired.

1 Introduction

Regulation is frequently ascribed as one of the putative drivers for rising industry concentration and underinvestment in the U.S. (Gutiérrez and Philippon, 2019, 2017) and in other countries (Aghion et al., 2021). At the core of these propositions lies an unanswered question of whether government regulation increasingly burdens small and large firms differently and acts as an obstacle to firm growth.[1] Due to limited information on firm-level incidence of regulatory compliance costs, prior studies have mostly focused on how industries' overall dynamism relates to industry compliance requirements, or at best how small and large firms' growth may respond to industry-specific regulatory policy changes. While these estimates are informative about small and large firms' elasticity to specific regulations, they cannot assess the full extent of the regulatory burden on small and large firms, which is the central metric for a complete quantitative assessment.

Indeed, how regulatory compliance costs fall on small and large firms is unclear ex ante. On the one hand, one could posit that the United States regulatory system may entail proportionally lower compliance costs for large businesses due to fixed costs or lobbying (Davis, 2017; Alesina et al., 2018; Gutiérrez and Philippon, 2019; Akcigit and Ates, 2020; Aghion et al., 2021). On the other hand, small businesses may pay lower compliance costs than others due to the many exemptions stemming from regulatory tiering (Brock and Evans, 1985; Aghion et al., 2021) or to more relaxed enforcement in the inspection of micro-establishments. Nonetheless, recent studies suggest that substantial aggregate output and productivity losses can arise in the presence of size-dependent regulatory constraints (Restuccia and Rogerson, 2008; Guner et al., 2008; García-Santana and Pijoan-Mas, 2011; Hovenkamp and Morton, 2019), yet, their extent is unknown.

While these questions are central to the ongoing policy debate on the regulatory compliance burden, measuring the size-dependency of regulatory compliance across the full spectrum of U.S. firms is challenging. For Goff et al. (1996) "*the measurement problems present such a large barrier that one could flatly assert the total amount of regulation to be unmeasurable by direct observation*" (p.87). Several factors make assessing the incidence of the regulatory burden across firms difficult from the quantitative perspective. First, regulatory requirements are different across industries, from financial services to mining, from agriculture to pharma, to name a few. Second, firms are subject to multiple sources of regulation, from federal to state and local agencies, and these overlapping regulations often intersect or complicate

[1] Gutiérrez and Philippon (2019) highlight the importance of "*emphasizing the heterogeneous impact of regulations on small and large firms*" and make a novel empirical contribution towards this direction.

each other (Agarwal et al., 2014; Kalmenovitz et al., 2025). Third, many regulators provide specific exemptions from compliance and rules are tiered to be less burdensome for certain employment classes, business forms, or ownership structures. Examples include small businesses as defined by the U.S. Small Business Administration,[2] sole proprietorship versus C-corporations, or publicly-traded versus privately-held firms. Fourth, for given regulatory requirements, small and large firms may face different stringency or frequency of inspection (Stiglitz, 2009; Kang and Silveira, 2021), which may result in heterogeneous on-equilibrium compliance efforts. Fifth, small and large firms face different technological constraints and reach different efficiency levels in complying with regulations, for example due to economies of scale or fixed costs.

A first contribution of this paper – and a necessary step for studying the relationship between regulation and firm size – is to systematically measure the costs of regulatory compliance for establishments and firms. We start by focusing on labor costs for regulatory compliance, which by many accounts are a central component of direct compliance costs.[3] Our measure is made possible by merging occupational task data from O*NET (V23.0) with the Occupational Employment and Wage Statistics (OEWS) establishment-occupation level microdata from the U.S. Bureau of Labor Statistics (BLS), a large representative survey covering about 1,200,000 establishments from all industries for 2002-2014.

We start by measuring the regulation-relatedness of a labor task in Section 3. Using a combination of keyword matching, manual assignment, and natural language processing methods, we assess the regulation-relatedness of 19,636 tasks in O*NET. We next aggregate the regulation-relatedness of tasks to the occupation level using an importance scale of each task for each occupation also provided by O*NET. Regulation-related occupations span about 1/3 of all occupation categories, covering not only compliance officers, but also compensation and benefits managers, financial examiners, nurse practitioners, insurance policy processing clerks, environmental technicians, among many others. We obtain our key measure of regulation intensity for each establishment by aggregating occupations' regulation-relatedness weighted by the establishment's labor spending on each occupation in OEWS. We label this measure *RegIndex*, a variable that indicates the percentage of an establishment's total labor spending ascribed to performing regulation-related tasks. Using the establishment's employer identification number (EIN), we further

[2]See, for example, U.S. Small Business Administration Office's advocacy for avoiding "excessive regulatory burdens on small businesses" at https://advocacy.sba.gov/about/ (Last accessed 5/1/23).

[3]For example, Jamie Dimon, CEO of J.P. Morgan Chase, in a "Letter to Shareholders" on April 9, 2014, states that "*We are applying enormous resources to the task [of regulatory compliance]... 13,000 employees will have been added since the beginning of 2012 through the end of 2014 to support our regulatory, compliance and control effort.*" We discuss later systematic evidence on the importance of labor spending for total regulatory compliance costs. As a robustness check, we also complement our measure by accounting for capital costs later.

aggregate RegIndex at the firm level (Song et al., 2018), and perform all analyses at both the firm and the establishment level, unless otherwise noted.

Compared to existing approaches for measuring the regulatory burden on industries or publicly traded firms (discussed in Section 2), our measure has several novel and important features. First, it provides a firm-specific micro measure, crucial for capturing heterogeneity in regulatory burden, as the same regulation may affect different firms differently, even within the same sector. Second, our results are based on actual compliance costs paid by firms rather than on statements, expectations, or projections by management or government agencies. Third, our measure reflects on-equilibrium effects of regulation, incorporating firms' endogenous responses to enforcement, regulatory ambiguity, and related factors. Fourth, our analysis covers both very large and very small firms—an essential feature for assessing the impact of government rules on small businesses. Fifth, our approach encompasses regulations from all relevant sources, including federal, state, local, and industry-level privately enforced rules. Finally, our methodology is operational and reproducible by government agencies for ongoing assessments and validation of rules; for instance, it can serve as a straightforward empirical measure of ex post compliance costs, complementing OIRA's ex ante projections.

Using a broad definition of regulation-related tasks which includes tax compliance, an average firm spends 3.33 percent of its total labor costs on performing regulation-related tasks per year; under our most conservative measure, the average is 1.34 percent. The aggregate compliance hours implied by RegIndex correspond to a range between 49 and 121 percent of the total compliance hours estimated by regulatory agencies from OIRA reports.[4] When aggregating RegIndex to the national level, we observe that regulatory compliance costs of U.S. businesses have grown by about 1 percent each year from 2002 to 2014 in real terms.

We next validate our RegIndex measure through a battery of tests. First, we show that our RegIndex can successfully pick up major industry-level regulations and, importantly, also deregulations which is a challenging task for many existing measures in the literature. Second, we find that the aggregated RegIndex series, which adds up establishments' actual regulatory compliance costs, closely tracks the aggregate predicted compliance costs reported by the OIRA. Third, we estimate a state-level RegIndex measure and show that states leaning towards voting for the Republican Party tend to have lower RegIndex. Fourth, we construct RegIndex for publicly-traded firms before and after IPO and show that RegIndex captures

[4] We use the most conservative version of RegIndex which we have constructed throughout our analyses. All results reported in the paper are robust to using a broad version of RegIndex that includes tax compliance and a medium version that excludes taxes, but does not down-weight regulation-related tasks based on the presence of other task objectives or keywords. These are available from the authors upon request.

the increased regulatory burden after firms become publicly traded (Ewens et al., 2024; Coates IV, 2007; De Fontenay, 2016). Moreover, our RegIndex for publicly-traded firms are also positively correlated with existing text-based firm-level measures from Kalmenovitz (2023) and Armstrong et al. (2025).

After discussing labor costs of regulatory compliance, we extend the analysis by including establishments' capital equipment and tools expenditure related to regulatory compliance. Capital equipment costs are measured following the approach presented in Caunedo et al. (2023). We use this methodology to create an extended version of regulatory costs including capital to generalize and assess the robustness of our findings. Indeed, the inclusion of capital costs produces results consistent with what we observe using labor costs alone, but the compliance cost levels are about 20 percent higher.

The paper then proceeds to address its key question. Section 4 asks how regulatory costs change with respect to firms' (or establishments') size, and it investigates the presence of increasing or decreasing returns to scale in regulatory compliance. Intuitively, a regulatory compliance cost schedule that is non-neutral to scale – as approximated by total employment – may distort incentives for producers, induce factor misallocation (Hsieh and Klenow, 2009; Huneeus and Kim, 2018), and constrain productivity growth (Parente and Prescott, 2002; Aghion et al., 2021). On the one hand, increasing regulatory costs with size implies incentives for firms to remain small, below the efficient scale of production (Guner et al., 2008; Garicano et al., 2016). This may arise, for instance, from government policies designed to support small businesses through lighter regulation (regulatory tiering), which implies increasing compliance requirements kicking in at progressively higher levels of employment (Brock and Evans, 1985; Brock et al., 1986; Alvero et al., 2023; Ewens et al., 2024).[5] On the other hand, decreasing regulatory costs with scale favors larger players vis-à-vis smaller competitors, quashing entry and fostering concentration (Gutiérrez and Philippon, 2019; Philippon, 2019). This may arise naturally from economies of scale in regulatory compliance, for example, due to the presence of fixed costs,[6] or it may even derive from regulatory capture and special deals for large players.[7] To the best of our knowledge, extant literature offers no clear consensus on the shape of the returns to scale in regulation for the entire U.S. economy.

Our main finding in Section 4 is an inverted-U relation between RegIndex and firm size. In particular,

[5] The distortive effects of regulatory tiering are well documented, especially in Europe, where firms below efficient scale appear over-represented and mid-size firms under-represented. See among the many Evans (1986); Boeri and Jimeno (2005); Schivardi and Torrini (2008); Gourio and Roys (2014); Garicano et al. (2016).

[6] Regulatory environments where increasing returns to scale prevail are frequently documented in the literature. For instance, in the case of environmental regulation, List et al. (2003) examine the effects of air quality regulation on new plant formation finding large negative impacts in New York State in 1980-90. In the case of pharmaceutical companies and the FDA, Thomas (1990) also finds large negative effects on small firms. In their analysis of compliance risk, Davis (2017) and Calomiris et al. (2020)find larger corporations being at an advantage (while focusing only on public companies).

[7] This is relevant with respect to classic research on the political economy of regulatory capture. For a recent discussion, see Lancieri et al. (2022).

firms under 500 employees experience increasing regulatory compliance costs as a share of total wages, with the percentage of labor spending on regulatory compliance sharply increasing with employment. This is consistent with the Code of Federal Regulation (CFR) having elements of tiering, but also with smaller firms having lower incentives to comply with regulations given a higher likelihood of flying under the regulator's radar (see Section 5). For firms above 500 employees, economies of scale kick in, and the percentage of labor spending on regulatory compliance progressively decreases with employment. On average, RegIndex for mid-size firms is about 47 percent greater than that of the smallest firms and 18 percent greater than that of the largest firms.

The inverted-U relation between RegIndex and firm size is a particularly robust feature of the data. First, we observe this relation after controlling for industry fixed effects, state fixed effects, and even firm fixed effects. Second, the relation remains non-monotonic after including capital costs in the RegIndex construction and when we probe the regulatory requirements of the same occupation across firms of different sizes. Third, we extensively examine the impact of outsourcing by conducting two industry-specific case studies based on reported size-specific outsourcing intensities and developing a comprehensive estimation that assigns local compliance outsourcing providers' costs to firms in different size bins, and we show that accounting for outsourced compliance costs is unlikely to overturn our finding that small firms have lower RegIndex than mid-sized firms. Fourth, we examine whether small and large firms have different requirements of regulatory compliance for the same occupation by analyzing additional data from 14 million job postings in the Burning Glass Technologies database. We show that for job postings of the same occupation, mid-sized firms require more regulatory compliance skills in their descriptions than small and large firms do, further reinforcing the finding of a baseline inverted-U shape between RegIndex and firm size.

Section 5 investigates three mechanisms behind the inverted-U shape between RegIndex and firm size: (i) fixed costs in compliance, (ii) size-dependent regulatory requirements, and (iii) differential enforcement for large and small firms. All together, the mechanisms of fixed costs and regulatory tiering appear consistent with key parts of the non-monotonic relation between compliance costs and firm size, while we find no evidence of a sizable role for differential enforcement. Specifically, in terms of (i) fixed costs, we find that large firms tend to systematically hire more specialists to comply with regulation than mid-sized and small firms. This evidence is consistent with centralization of regulatory compliance: Large firms find it economical to consolidate their compliance efforts in the hands of dedicated employees to save costs.

We propose and implement a shift-share method for understanding whether part of the relation between RegIndex and firm size may be driven by (ii) size-dependent regulatory requirements or by (iii) regulatory agencies' heterogeneous enforcement efforts. Our methodology takes two steps. First, we categorize regulation-related tasks into tasks for coping with the enactment and retirement of regulations, i.e., *requirement-sensitive tasks*, and tasks for coping with changes in regulatory enforcement, i.e., *enforcement-sensitive tasks*. To do so, we exploit the fact that each regulation-related task has a different exposure to the twelve major regulatory agencies, i.e., the "share", due to the task's presence in different industries and different industries' exposure to the agencies. For each of the major regulatory agencies, we construct a "shift" in its regulatory requirements based on the enactment and retirement of the agency's regulations and a "shift" to its enforcement stringency based on changes in its regulatory employment. Estimating business's spending on each task in response to the two shift-share measures, we obtain each regulation-related task's sensitivities to regulation requirement and enforcement and its categorization. Second, we use the two types of tasks to decompose a firm's (establishment's) RegIndex into RegIndex for requirement and RegIndex for enforcement. We show that the fact that small businesses are systematically shielded from regulatory requirements (i.e., regulatory tiering) is a mechanism matching the upward component of the inverted-U relation between regulatory compliance costs and firm size, but no evidence of differential enforcement effort by government agencies playing a role. This finding is consistent with the pervasive use of threshold-based regulations in the U.S.

Finally, we note that the inverted-U pattern can have implications for the dynamics of the U.S. firm size distribution. While our setting is not well suited to establish causality (see Garicano et al. (2016) for a proper setting), with this caveat in mind, we document two descriptive patterns: First, we find a strong negative association between RegIndex and changes in the firm size distribution within sectors. Second, at the aggregate level, we observe a rise in the share of very small firms (fewer than 10 employees) and a corresponding decline in the share of mid-sized firms between 2002 and 2014. We view these patterns as suggestive, and hope they can help guide future work on this topic.

Our paper confines its scope to the costs of regulation without addressing the benefits of regulation. We further limit our analysis to assessing the labor and equipment capital costs of regulatory compliance in terms of workers' wages and associated equipment and tools costs paid to comply with rules or standards required by the government for existing firms. Hence, our work does not cover other types of regulatory compliance costs, such as non-equipment expenditures (e.g., structure capital like reinforced concrete walls, pumping or draining infrastructure for mine water, etc.) and foregone investment opportunities and profits due to regulatory risk.[8] We also do not separately capture setup fixed costs for compliance as "*one-time costs of learning the relevant regulations, developing compliance systems and establishing relationships with regulators*," which may prove substantial barriers as discussed in Davis (2017, p.14); Alesina et al. (2018); Aghion et al. (2021), and systematically documented at least since Djankov et al. (2002). Nonetheless, we argue that labor and equipment compliance costs are a key dimension of the question. According to survey estimates from the Securities Industry Association (2006), 93.9 percent of regulatory compliance costs in the U.S. financial sector are labor related and 3.3 percent are physical capital related.[9] According to survey estimates from the National Association of Manufacturers (2014), 68.4 percent of regulatory compliance costs in the U.S. manufacturing sector are labor related and 13.4 percent capital related. Finally, while we do not directly observe firms' compliance costs paid to outsourcing service providers (e.g., external legal, compliance, and accounting services), the aforementioned surveys show that spending on outside advisers accounts for only 2.8 percent and 8.7 percent of total compliance costs for the financial sector and the manufacturing sector, respectively. We also conduct extensive assessments which suggest that accounting for outsourced compliance costs is unlikely to affect our main findings.

2 Related Literature

This paper relates to several strands of literature. The first pertains to measurement. While we are not aware of any alternative measure of regulatory costs with the exact properties of RegIndex, it is important to mention that other methods designed to obtain a comparative perspective of regulatory requirements have been proposed in the past, at least since Morrall (1986), and that many have important advantages of their own.[10] Al-Ubaydli and McLaughlin (2017)'s RegData, a data repository maintained by the Mercatus Center at George Mason University, is built using QuantGov, a library of machine learning algorithms and text analysis tools designed to collect information about the number of restrictions, rule complexity and industry incidence from the text in the Code of Federal Regulations. In another novel application, Kalmenovitz (2023) employs data from Form 83-I filed to OIRA to indicate the expectations

[8]See, for example, Ryan (2012) for a case where such omissions are of the first order. For analysis of the social costs of environmental regulation, see Hazilla and Kopp (1990).

[9]However, see Alvero et al. (2023) for a criticism of self reported regulatory costs for the banking sector.

[10]A few earlier studies attempted to use occupational data to measure the effects of regulation like we do. For example, Schaefer and Zimmer (2003) use accountants and auditors from the Current Population Surveys to examine the state regulations in the accounting profession from 1984 to 2000. Parker et al. (2009) survey 999 large Australian firms regarding their expenses on lawyers and compliance professionals to measure firms' costs of compliance with the Trade Practices Act.

of regulators about the cost of compliance with each regulation.[11] These measures shed light on regulatory burden at the more aggregate level.

More recently, a burgeoning strand of literature has developed to measure publicly-traded firms' regulation. Davis (2017) focuses on Part 1A of 10-K filings to gauge publicly traded firms' exposure to regulatory and policy risk.[12] Calomiris et al. (2020) focus on the transcripts of corporate earnings calls made by publicly traded firms and show that its measure of regulation is predictive of sales and assets growth, leverage, and other metrics of firm performance.[13] Kalmenovitz (2023) uses supervised machine learning to assign paperwork regulations (Form 83-I) to each publicly-traded firm based on the similarity between the firm's 10-K (Item 1) text and the regulation's text. A more recent development by Armstrong et al. (2025) measures each publicly-traded firm's exposure to each government regulatory agency based on 10-K texts, which further allows researchers and policy-makers to assess the impact of regulation from each agency on firms.[14]

Our RegIndex measure complements these existing firm-level measures in at least two unique ways. First, our measure can be applied to both publicly-traded firms and also privately-held firms. Hence, our measure can help understand regulatory costs across the full distribution of firms, including the very small firms with fewer than 50 employees which are frequently in the policy debate. Our measure can also help understand how publicly-traded firms' regulatory burden changes before and after IPO, as we demonstrate in this study. Second, our measure provides a *de facto* dollar amount of each firm's regulatory compliance costs, which is difficult to obtain in previous measures based on firms' 10-K or earnings call texts. Our measure and data also face limitations, as accessing our firm-level RegIndex measure requires the approval from the BLS for using their confidential microdata. BLS surveys most establishments once every three years, which further limits future research on within-establishment changes in RegIndex to a three-year horizon rather than higher frequencies. To shed light on how our measure compares with existing measures, we compute RegIndex for publicly-traded firms and show that our RegIndex is about 17%-28% correlated with the measures from Kalmenovitz (2023) and Armstrong et al. (2025), suggesting

[11] The 83-I forms include estimates of how many responses the regulator will receive per year, how many work hours firms will be required to dedicate to complying with the regulation, and, for 20% of the regulations, the estimated dollar costs of compliance (Kalmenovitz (2023)). Relative to Al-Ubaydli and McLaughlin (2017), the method in Kalmenovitz (2023) has the major benefit of aggregating the expectations of experts regarding the compliance burden.

[12] This is also an approach preceded by Baker et al. (2016) and followed subsequently by Hassan et al. (2019), who focus on firms' exposure to a broader political risk, as expressed in the 10K-forms of publicly traded firms.

[13] An important advantage of the Davis (2017) and Calomiris et al. (2020) approaches is that they are apt at capturing future regulatory risk, both in terms of discretionary enforcement and of new rules affecting firms. Indeed, Calomiris et al. (2020) underscore compliance risk as the most relevant channel behind their finding, rather than physical compliance costs.

[14] See also Bombardini et al. (2025) for a recent review of the economics literature on measuring the costs and benefits of regulation.

that our measure captures related but also distinct dimensions of regulatory burden.[15] We refer to Bombardini et al. (2025) for more detailed discussions on recent methodologies on measuring costs and benefits of regulation and exploring their economic impacts on firms.

The second strand of literature we contribute to is on the distribution of regulatory burden across firm sizes. A large body of prior work has examined this question in specific settings and showed inconclusive findings regarding which group is more burdened by regulations. On the one hand, larger firms may bear lower percentage regulatory burden because of economics of scale or regulatory capture. Prior studies supporting this view usually draw conclusions by examining small and large firms' economic responses to regulatory changes instead of directly measuring the heterogeneous regulatory costs by firm size. For instance, List et al. (2003) and Thomas (1990) show specific regulations more negatively impact smaller firms, Gutiérrez and Philippon (2019) and Philippon (2019) more systemically show the employment of smaller firms are more negative associated with industry regulation using RegData, Davis (2017) and Calomiris et al. (2020) find larger corporations are an advantage in responding to compliance risk than while focusing on publicly-traded firms, among others. Among studies supporting this view, they also make different predictions on how regulation shapes firm size distribution: Klapper et al. (2006) show that regulation forces successful entry firms to be larger, whereas Singla (2023) shows that increases in regulatory costs contribute to small firms getting smaller. Reconciling this debate is challenging as it is difficult to directly observe regulatory costs for small firms. On the other hand, small firms may bear lower regulatory burden due to regulatory tiering or lax inspection by regulatory agencies. Prior studies supporting this view usually draw conclusion by examining firm size distribution above and below the tiers rather than directly showing firms compliance costs (Brock and Evans, 1985; Brock et al., 1986; Alvero et al., 2023; Ewens et al., 2024).

Our study contributes to this literature by directly showing regulatory compliance costs across the full distribution of firm sizes, with some novel implications. We show a more nuanced picture relative to extant literature by highlighting a non-monotonic relationship between regulatory burden and firm size. Our finding that mid-sized firms are more burdened than small and large firms offers a potential reconciliation of the prior studies which tend to focus on firms from the two ends of the size distribution. Our finding also has important implications for the equilibrium size distribution of firms as it indicates that the heaviest regulatory hurdle for firm growth is when firms become mid-sized. We further analyze the mechanisms behind these results by highlighting the differential roles of regulatory requirements

[15]We thank Joseph Kalmenovitz, Daphne Armstrong, Stephen Glaeser, and Jeffrey Hoopes for making their firm-level measures publicly available.

versus endogenous compliance to enforcement.

Finally, this paper is related to a newly revived literature on regulation and its political economy. Part of this literature focuses on the role of firms in influencing regulation, via lobbying and political influence (Blanes i Vidal et al., 2012; Bombardini and Trebbi, 2025; Bertrand et al., 2014; Kang, 2016; Bombardini and Trebbi, 2020) or, more recently, via direct contest between different groups of firms (Singla, 2023). Another branch explores the political and policy risks to which firms are exposed (Julio and Yook, 2012; Baker et al., 2014, 2016; Hassan et al., 2019). A burgeoning recent literature on the interaction between market power and political power centers on regulation as the main tool employed by large business entities to generate rents (Callander et al., 2021; Cowgill et al., 2021).

3 Data, Measure, and Validation

3.1 Data

Our main source of information is the establishment-occupation level microdata from 2002 to 2014 provided by the OEWS program of the BLS. This data set covers surveys that track employment and wage rates for over 800 detailed occupations in approximately 1.2 million establishments over the course of a three-year cycle (Zhang, 2019). The sample of establishments covers, on average, 62 percent of the non-farm employment in the U.S. Within a three-year cycle, 400,000 establishments are surveyed during each year and therefore the same establishment is surveyed at most every three years (e.g., in t and $t+3$). The microdata provides each establishment's sampling weight (to recover economy-wide aggregates), NAICS 6-digit industry code, county code, government ownership indicators, and parent firm's employer identification number (EIN).

Defining the appropriate regulatory compliance entity is nontrivial. A firm may be considered the ultimate entity, in that it pays the fines and penalties if inspected at any constituting establishment and found in violation of agency rules. However, regulation varies across industries and states, inducing a firm to spend different amounts of resources on regulatory compliance across different constituting establishments. For these reasons, we conduct our main analysis at both the firm level and the establishment level.

An EIN will define the boundary of a firm in our analysis. This is because the EIN identifies a firm for tax purposes, and also because the EIN is commonly used to define a firm both in the academic literature (e.g., Chodorow-Reich, 2014; Song et al., 2018) and by government agencies, such as the IRS and the

BLS.[16] Following the convention recommended by the OEWS program, we aggregate establishments of an EIN surveyed in years $t-2$ to t to represent the occupational composition of the EIN in year t.[17] If a firm has establishments spanning multiple industries, we define the firm's major industry as the NAICS 6-digit code with the highest employment share.[18]

Our research makes use of the task statements for each occupation from the O*NET (V23.0) database. Each task statement is a single sentence and is pertinent to a unique occupation. An occupation is described by an average of 22 different task statements. O*NET also provides an average rating indicating how important the task is for the occupation rated by incumbent workers working in the occupation, occupational experts, and analysts. Occupations are categorized at the 8-digit standard occupational code (SOC) level. We match the 8-digit SOC in O*NET to the 6-digit SOC occupations in the OEWS microdata, creating a characterization of all tasks performed at the establishment level, which allows for the construction of the measures described in the following subsections.

3.2 A Simple Theoretical Framework for the Measure

Our approach to measuring regulatory compliance costs for firms (or establishments) is as follows. Let us assume that a firm's labor includes workers performing regulatory compliance related tasks and workers performing other production tasks. Let L_{it} be the total number of production workers employed in firm i at time t and w_{it} the average wage paid to the workers. Define R_{it} the total number of workers occupied in regulatory compliance related tasks. We allow the average wage of compliance workers to differ from that of production workers (due to specialization) and indicate it with w^r_{it}. We assume all wages to be taken as given from the firm's perspective. R_{it} can be derived as the result of optimization on the part of the firm owners, who maximize profits:

$$Y_{it} - w_{it}L_{it} - w^r_{it}R_{it} - p_{it} \times f_{it}$$

with respect to R_{it} and L_{it}, where firm i faces a probability of inspection p_{it} at period t and fines $f_{it} = \frac{\tilde{R}_{it}}{R_{it}}$ are levied in case the firm is found in less than full compliance with regulatory requirements $\tilde{R}_{it}$. We

[16] See U.S. Department of Labor, Bureau of Labor Statistics, "Business Employment Dynamics Size Class Data: Questions and Answers," http://www.bls.gov/bdm/sizeclassqanda.htm, questions 3 and 5.

[17] The BLS OEWS program uses a similar aggregation approach to publicize industry or geographical level statistics for each year. See technical notes at https://www.bls.gov/oes/oes_doc_arch.htm.

[18] Because the OEWS survey does not cover all establishments of a firm, our firm-level measures contain measure errors and our firm-level employment is under-estimated. Given that over 80 percent of establishments in our sample are standalone single-unit firms (Ayyagari and Maksimovic, 2017), the establishment-level and firm-level results are qualitatively the same for most analyses, as we will show.

further assume a constant return to scale production function $Y_{it} = \phi_{it} L_{it}$, where ϕ_{it} indicates a firm-specific productivity shock, and that $\tilde{R}_{it} = \tilde{R}(L_{it}) = \rho L_{it}^{\alpha}$ and $p_{it} = \pi L_{it}^{\beta}$, where α, β, ρ, π are policy parameters governing regulatory requirements and enforcement effort targeting establishment i. Simple static profit maximization implies

$$\begin{aligned} R_{it}^{*} &= \left(\frac{p_{it} \tilde{R}_{it}}{w_{it}^{r}} \right)^{\frac{1}{2}} = \left(\frac{\phi_{it} - w_{it}}{(\alpha + \beta) \left(w_{it}^{r} \pi \rho \right)^{\frac{1}{2}}} \right)^{\frac{\alpha+\beta}{\alpha+\beta-2}} \left(\frac{\pi \rho}{w_{it}^{r}} \right)^{\frac{1}{2}} \\ L_{it}^{*} &= \left(\frac{\phi_{it} - w_{it}}{(\alpha + \beta) \left(w_{it}^{r} \pi \rho \right)^{\frac{1}{2}}} \right)^{\frac{2}{\alpha+\beta-2}} . \end{aligned} \tag{1}$$

Note that, intuitively, equilibrium compliance is an increasing function of regulatory requirements $\tilde{R}_{it}$, an increasing function of the enforcement activity of regulators via inspections, p_{it}, and a decreasing function of regulatory compliance labor wages, w_{it}^{r}.

We define the index of regulatory compliance costs as:

$$\text{RegIndex}_{it} = \frac{w_{it}^{r} R_{it}}{W_{it}}, \tag{2}$$

that is, the share of the total wage bill $W_{it} = w_{it}^{r} R_{it} + w_{it} L_{it}$ allocated to compliance labor related to regulation.

3.3 Construction of the Regulation Index

The construction of the regulatory index, RegIndex, starts with identifying which tasks are related to regulatory compliance. We achieve this goal by analyzing the texts of task statements in the O*NET data in two phases: a *keyword matching phase* and an *annotation phase*. In what follows, we briefly describe the procedure for estimating our preferred (conservative) version of the RegIndex, which we use for our analysis, and we relegate more details and further description of two broader versions of the RegIndex to the Appendix A.

In the keyword matching phase we identify regulation-related tasks by matching the task statements to two tiers of keywords to balance the rate of false positives and false negatives in identifying regulation-related tasks. The first tier of keywords includes the words *regulation*, *regulations*, and *regulatory*. These matches exhibit a low rate of false positives, but may miss some regulation-related tasks when regulation is described using synonyms. For this reason, a second tier of keywords includes 34 keywords that identify alternative references to regulation, such as *statutes, ordinances,* and *code*, references to government agencies, and actions of compliance. Appendix A details the 34 keywords, and Appendix Table A.1 lists ten examples of regulation-related tasks identified by the first-tier keywords.

In the annotation phase we manually annotate each matched task statement to exclude further false positives, such as tasks in which the word "*codes*" refers to computer programming. This procedure results in a final list of 829 regulation-related tasks out of a total of 19,636 tasks in the O*NET database. Next, we assign a *regulation-relatedness value* between 0 and 1 to account for the heterogeneity among regulation-related tasks. In particular, tasks identified by the second-tier keywords, which maybe less informative about the tasks' relevance for government regulation, and tasks with multiple focuses in their statements are down-weighted to have a regulation-relatedness value less than 1.

Having measured the regulation-relatedness of each task, we next compute the regulation-task intensity (RTI) for each SOC 6-digit occupation by averaging its tasks' regulation-relatedness values weighted by the importance ratings of each task for that occupation based on O*NET importance weights. Appendix Table A.2 lists the top 25 SOC 6-digit occupations with the highest RTI. For instance, Compliance Officers have the highest RTI of 0.343, meaning that conservatively Compliance Officers on average spend 34.3 percent of their work hours on directly performing government regulation-related tasks (Compliance Officers also perform tasks such as "*maintain and repair materials, worksites, and equipment*", which we do not classify as regulation). Finally, we merge each SOC 6-digit occupation's RTI to the relevant establishments in the OEWS data, which provides each establishment's labor cost (employment times wage rate) for each occupation.

We define an establishment's regulation index (RegIndex) as the share of its total labor cost spent on performing regulation-related tasks. In particular, an establishment i's RegIndex is its occupations' average RTI weighted by its labor cost on each occupation j at time t:

$$\text{RegIndex}_{i,t} = \frac{\sum_j RTI_j \times emp_{i,j,t} \times wage_{i,j,t}}{\sum_j emp_{i,j,t} \times wage_{i,j,t}} \times 100. \tag{3}$$

Using regulation-relatedness values from the conservative, medium, and broad approaches, we obtain three versions of establishment-level RegIndex. The three versions of RegIndex are highly correlated. Hence, we perform much of our analysis using the conservative RegIndex.[19]

A firm's RegIndex is similarly obtained based on occupational labor costs aggregated from those of

[19]Notice that the definition of RegIndex can also be applied to a single regulation related task o by assessing the ratio of the task's associated labor costs as a share of the total wage bill of the firm or establishment. We indicate it as $\text{RegIndex}_{i,o,t}$.

establishments with the same EIN. Appendix Table A.3 shows the top 25 NAICS 3-digit industries whose establishments have the highest average RegIndex. This study focuses on firms' (or establishments') in-house regulatory compliance cost. Hence, it excludes firms in industries which provide legal or compliance work as their main function including legal services, accounting services, central banks, and public administration.[20]

Capital Our baseline RegIndex is based on the firms' labor costs. As noted earlier, another important component of firms' regulatory compliance spending is capital expenditure. Unfortunately, due to the lack of microdata on firm-level detailed capital spending, a precise estimation of firms' regulatory compliance costs accounting for capital expenditure is not available. Nevertheless, alternatives have been devised in the literature based on tools and capital equipment use. Here we follow Caunedo et al. (2023) and construct an approximate measure of capital cost per employee for each occupation, based on the occupation's use of tools and equipment from the O*NET Tools and Technology module and from the NIPA quantity and price tables for each equipment item. With this additional information available, one is able to sum up capital user costs per employee and annual wages per employee for each occupation, obtaining the total input cost (in terms of labor and capital) for each employee in each occupation. From this, it is possible to proceed to construct a more general version of RegIndex.

The capital augmented $\text{RegIndex}_{i,t}^{tot}$ is the share of a firm i's total input costs at t spent on performing regulation-related tasks:

$$\text{RegIndex}_{i,t}^{tot} = \frac{\sum_j RTI_j \times emp_{i,j,t} \times (wage_{i,j,t} + kcost_{i,j,t})}{\sum_j emp_{i,j,t} \times (wage_{i,j,t} + kcost_{i,j,t})} \times 100. \tag{4}$$

Focusing on $\text{RegIndex}_{i,t}^{tot}$, we can derive additional statistics about the the composition of firms' regulatory compliance spending in terms of labor and capital that are highly consistent with industry survey findings.[21]

[20]Following Song et al. (2018), educational institutions are also excluded due to high government ownership.

[21]For instance, National Association of Manufacturers (2014) report shows that labor spending on regulatory compliance cost as a fraction of the sum of labor and capital spending on regulatory compliance cost is 84% for manufacturing industries in 2012 (see Chart 15 of the report). Aggregating regulatory compliance costs based on $\text{RegIndex}_{i,t}^{tot}$ for the manufacturing sector in 2012, we obtain a corresponding fraction of 78%. Similarly, Securities Industry Association (2006) report shows that labor spending on regulatory compliance cost as a fraction of the sum of labor and capital spending on regulatory compliance cost is 97% for financial industries in 2006 (see figure 3b of the report). We using$\text{RegIndex}_{i,t}^{tot}$ for the financial industries in 2006, we obtain a fraction of 88%.

3.4 Summary Statistics

The sample for the analysis in this paper includes 3.03 million U.S. firm-year observations and 3.36 million establishment-year observations surveyed by the OEWS program from 2002 to 2014. Table 1 provides the key summary statistics. To begin with, the unweighted average firm in our sample has 92 employees and the unweighted average establishment has 48 employees, both appear above the national average based on Census statistics. If one applies the sampling weights to the establishments assigned by the OEWS, the weighted average employment falls to 14, which is closer to the establishment-level average employment of 15.6 employees reported by the Census Statistics of U.S. Businesses (SUSB) during 2002-2014. The average annual wage per employee in our sample is $43,733 ($2.09 million divided by 47.79), which is in line with the average annual wage per employee of $41,974 from SUSB.

The key summary statistics in Table 1 focus on RegIndex. The average firm in our sample spends 1.34 percent of its total labor costs on regulation-related tasks in any given year. The average establishment spends 1.31 percent.[22] To be further noted, RegIndex varies substantially across firms and establishments, with the 0.5 percentile at 0, the median at 0.8-0.9 percent, and the 99.5 percentile above 10 percent. Further decomposition shows that year fixed effects explain merely 0.01 percent of the variation in establishments' RegIndex, state fixed effects explain 0.10 percent, NAICS 6-digit industry fixed effects explain 36.13 percent (for coarser categories, NAICS 2-digit, 3-digit, and 4-digit fixed effects explain 8.69 percent, 12.16 percent, and 22.15 percent, respectively), and the residual 63.63 percent of the RegIndex variation is unexplained by the above.

To illustrate the recent time trends in regulatory compliance costs, Figure 1 plots the aggregate time-series of RegIndex. Following the aggregation method explicit in the OEWS program, Figure 1 aggregates RegIndex of all establishments in our sample surveyed between $t-2$ and t, weighted by a product of the establishments' sampling weights and their total labor cost, to obtain the average RegIndex of the U.S. economy at t.

From Figure 1, we observe an increase in aggregate RegIndex from 1.49 percent in 2002 to 1.59 percent in 2014, reaching $79 billion in regulatory compliance costs. This level of regulatory compliance costs is a conservative estimate. If we use all private establishments (including establishments that earn revenue from regulatory compliance services, such as legal services), the aggregate regulatory compliance costs

[22]These statistics are based on the conservative RegIndex, our main measure. Using the broad measure, the average establishment in our sample spends 3.28 percent of its total labor costs on regulation-related tasks in any given year, and the average firm spends 3.33 percent.

are \$103 billion in 2014. If we apply our broad version of RegIndex to all private establishments, the aggregate regulatory compliance costs are \$239 billion. Finally, adding capital equipment brings this upper bound to \$289 billion in 2014. For comparison, U.S. gross business income taxes amounted to \$353 billion in the same year.[23]

In real terms, from 2002 to 2014 the yearly growth rate of aggregate regulatory compliance costs averaged around 1 percent, which is about half of the 1.92 percent average yearly growth rate of the U.S. Real Gross Domestic Product over the same period. Overall, these patterns indicate the substantial economic magnitude of regulatory compliance, but a burden that has been growing at a slower rate relative to the rate of the growth of the U.S. economy.

3.5 Validation

3.5.1 Time-Series Relation with Agency-Estimated Compliance Hours

As a first validation exercise, we examine whether our approach for constructing RegIndex captures the time-series variation in the regulatory costs in the United States. Federal regulatory agencies are required to file Form 83-I to the OIRA in which the agency estimates firms' or individuals' compliance time for each regulation. We collect the estimates for each year from 2002 to 2014 from the "Information Collection Budget of the United States Government" from the White House website.[24]

We next compute counterpart compliance hours based on our approach for constructing RegIndex. Specifically, we regard an occupation's regulation-task intensity as the fraction of time an employee spends on regulation-related tasks in an hour. Assuming that all regulation-related occupations work 2,080 full-time hours in a year (noting that part-time workers are concentrated in the retail and restaurant industries), we estimate U.S. establishments' aggregate *de facto* regulatory compliance hours.

Figure 2 plots the time-series of annual aggregate regulatory compliance hours based on our approach and the compliance hours estimated by regulatory agencies from OIRA. OIRA compliance hours include a non-trivial amount of estimated hours for households to comply with individual income taxes. To produce a proper comparison of OIRA with our measure, which focuses on businesses' regulatory compliance hours, we exclude estimated hours for individual income taxes from the OIRA total compliance hours.[25] Panel

[23]See https://www.irs.gov/statistics/soi-tax-stats-irs-data-book.

[24]See https://www.whitehouse.gov/omb/information-regulatory-affairs/reports/. The filling for Form 83-I is mandated by the Paperwork Reduction Act (44 U.S.C. 3501). Federal agencies estimate three burden metrics for each regulatory regarding how many responses it will receive per year, how many hours it will take the public to comply with the regulation, and what would be the dollar costs of compliance. Only estimated compliance time is consistently reported by the "Information Collection Budget" report by OIRA each year.

[25]Compliance hours for individual income taxes account for about 40 percent of the Department of Treasury's to-

A shows that our estimates based on the conservative RegIndex account for about half of the hours estimated by federal agencies through OIRA. Using a less conservative weighting of tasks and accounting for all tax compliance efforts, the compliance hours based on our medium and broad versions of RegIndex amount to 106 percent and 121 percent of the OIRA compliance hours respectively.

Figure 2 also shows that compliance hours based on our RegIndex track the estimated compliance hours by agencies robustly over time in terms of changes. The two time-series exhibit a statistically significant correlation of 67 percent using our main conservative definition and 71 percent using our medium and broad definitions. Appendix Figure A.1 shows that the results are similar for all three versions of RegIndex.

3.5.2 Establishment RegIndex and Industry Regulatory Policy Changes

As a second validation exercise, this subsection illustrates the response of RegIndex to four major industry-level regulatory policy changes. Before and after salient regulatory policy changes, we examine the RegIndex response by establishments within a treated industry relative to appropriately matched control industries. These case studies not only help demonstrate the effectiveness of our RegIndex in tracing major industry-level regulatory policy changes, but also suggest that our RegIndex is able to overcome several limitations of existing regulation measures.

Case 1: Regulation of the Credit Card Industry The Credit Card Accountability Responsibility and Disclosure Act of 2009 (CARD Act) was drafted to "implement needed reforms and help protect consumers by prohibiting various unfair, misleading and deceptive practices in the [U.S.] credit card market" (U.S. Senate, 2009).[26] A key aspect of the CARD Act, in particular, was to impose regulatory limits on the ability of credit card issuers to charge certain types of credit card fees on customers, which became effective in February and August of 2010. Accordingly, as the treatment group for this analysis we use establishments in the credit card issuing industry (NAICS 52221). As the control group we use establishments in all other nondepository credit intermediation industries narrowly defined (NAICS 5222x, except for 52221, including sales financing, consumer lending, and real estate credit).

tal regulatory compliance hours. For example, in 2011, The U.S. individual income tax return imposes an estimated 2.70 billion compliance hours out of 6.74 billion compliance hours for Department of Treasury's total regulations, see https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/inforeg/inforeg/icb/icb_2013.pdf. We thus exclude compliance hours for individual income taxes in Figure 2 by subtracting 40 percent of Department of Treasury's estimated compliance hours from the total estimated compliance hours by all agencies. See the breakdown of IRS tax filing hours at https://taxfoundation.org/tax-compliance-costs-irs-regulations/.

[26]Congress enacted the Card Act of 2009 after research (e.g., (Simkovic, 2009)), as cited by a Senate committee, suggests that credit card pricing was less than perfectly competitive.

Case 1 in Figure 3 plots the weighted average RegIndex for the treated credit card issuing industry and for the control industries in the years around the policy change (from 2005 to 2014). In the figure, establishment RegIndex is weighted by the product of each establishment's sampling weight and each establishment's total labor costs. We can clearly observe in this figure that before the enactment of the CARD Act in 2009 the 95 percent confidence interval of RegIndex for the treated and the control groups appear to overlap and that they are statistically indistinguishable from one another. This feature of the data suggests the validity of the parallel trends assumption necessary for the consistency of the simple difference-in-differences estimator underlying this case study. We can see further that after the policy change, RegIndex in the credit card issuing industry rises dramatically, while the RegIndex average for the control group remains basically flat. While this graphical evidence underscores the ability of RegIndex to trace the heightened regulatory burden associated with the CARD Act on credit card issuing establishments during the post period of the analysis, measuring regulation intensity at such a granular level is proven to be challenging for supply-side approaches. Indeed, we show in the Appendix Figure A.2 that the popular RegData measure, which counts restrictive words in the Code of Federal Regulations (CFR), cannot identify a similar effect to ours.

Case 2: Deregulation and Re-regulation of the Oil and Gas Industry In our second case study, we explore both (i) the deregulation of the oil and gas extraction industry by the Energy Policy Act of 2005 (EPAct) under President George W. Bush, which exempted oil and gas facilities from environmental regulations to boost production, as well as (ii) the re-regulation of the industry by President Barack Obama's executive orders following the British Petroleum (BP) Deepwater Horizon oil spill in 2010. These orders led to several new regulatory policies, which were regarded as "the most aggressive and comprehensive reforms to offshore oil and gas regulation and oversight in U.S. history", according to the Bureau of Ocean Energy Management.[27]

Our treated industry for this case study is oil and gas extraction (NAICS 2111). We select downstream industries which use a significant amount of the treated industry's output as their inputs as the control industries. We assume that downstream industries share close economic ties with the upstream treated industry, but face a sufficiently different set of regulations to be not as strongly affected by EPAct and the Obama orders.

In fact, both control and treated industries may be affected by similar economic and regulatory forces

[27] See https://www.boem.gov/regulatory-reform/.

prior to the regulatory policy changes.[28] Using the BEA input-output table from 2007, we select the following three industries as the control group: petroleum and coal products manufacturing (NAICS 3241), natural gas distribution (NAICS 2212), and basic chemical manufacturing (NAICS 3251).[29]

Case 2 in Figure 3 shows parallel trends of RegIndex for the oil and gas extraction industry and the control industries before the enactment of the EPAct in 2005. After 2005, there appears to be a dramatic decline in the RegIndex for oil and gas extraction relative to the control industries, in line with the EPAct being a deregulatory policy change. This evidence highlights an important advantage of our measure: RegIndex can distinguish between regulation and deregulation. By contrast, separating statutory text that increases regulatory burdens from statutory text that deregulates an industry is challenging for supply-side language-based measures. Such measures tend to increase with both regulation and deregulation. Consistent with this limitation, we show in the Appendix Figure A.2 that the supply-side RegData measure exhibits an increase in regulation of oil and gas extraction after the EPAct was signed into law in 2005.

After the BP oil spill in 2010, consistent with the contemporary understanding that the industry faced heavy re-regulation, we observe a rapid increase in the RegIndex for the oil and gas extraction industry both in absolute terms and relative to control industries.

Case 3: Regulation of Pharmaceutical Manufacturing Our third case study focuses on pharmaceutical manufacturing which has experienced significant regulatory changes since 2008. The U.S. Food and Drug Administration Amendments Act signed into effect on September 27, 2007 increased the FDA's regulatory authority over drug safety and required companies to implement a risk evaluation and mitigation strategy for many drugs. Following the legislative change, FDA significantly amended the current good manufacturing practice (CGMP) regulations for human drugs in 2008 on quality control, risk management, record-keeping, facility design and maintenance, and employee training.[30]

We choose establishments from the pharmaceutical and medicine manufacturing industry (NAICS 3254) as the treatment group for this analysis. As the control group, we use establishments in all other

[28] While intuitive, choosing control industries based on input-output relations offers no guarantee that the treated and control industries will exhibit parallel trends in the Regulation Index during the pre-treatment periods. We thus examine the parallel trends empirically when analyzing each regulatory shock. Another challenge with this approach is that the control group may also be affected by the new regulatory policy changes. When treatment and control groups are both affected by the regulatory shock, we will be less likely to detect significant differences between the treated and control groups post-treatment. In this sense, our selection of control industries is conservative.

[29] The input-output account data from BEA provides information at the detailed industry level for only 2007 and 2012. See https://www.bea.gov/industry/input-output-accounts-data.

[30] https://www.federalregister.gov/documents/2008/07/15/E8-16011/current-good-manufacturing-practice-and-investigational-new-drugs-intended-for-use-in-clinical.

chemical manufacturing industries (NAICS 325x except for 3254) including the manufacturing of basic chemical, synthetic rubber, pesticide, paint, soap, etc. Case 2 in Figure 3 shows that the RegIndex of pharmaceutical and medicine manufacturing experience a similar slow upward trend with that of other chemical manufacturing before 2008. However, RegIndex of pharmaceutical and medicine manufacturing increased dramatically after FDA taking regulatory actions in 2008, while RegIndex for for other chemical manufacturing is essentially flat in the post treatment period. In contrast, RegData cannot distinguish the treated versus the control groups, as can be seen from Appendix Figure A.2. The likely reason is that FDA guidance may not be immediately reflected in the CFR.

Case 4: Regulation of Hospitals The fourth and final case study used for the validation of RegIndex focuses on the healthcare industry. The Patient Protection and Affordable Care Act (ACA) is a landmark U.S. federal statute enacted in 2010 under President Barack Obama. The ACA represents the U.S. healthcare system's most significant regulatory overhaul and expansion of coverage since the enactment of Medicare and Medicaid in 1965. At the time of its passing, hospitals faced significant pressure in coping with the new regulatory changes and dealing with insurers, all features that should correspond with higher levels of RegIndex.

For this case study we regard hospitals (NAICS 622) as the treated group industry and use animal hospitals (NAICS 54194) as the control industry establishments. Case 4 in Figure 3 validates the parallel trends assumption for RegIndex in hospitals and animal hospitals prior to 2010, which is once again suggestive of this exercise being informative. As can be seen from the figure, after 2010, the RegIndex of hospitals increases in both absolute terms and relative to the control group, appropriately tracing the heightened regulations imposed on the treated group relative to the controls. In this case, where ACA is clearly mapped in the CFR, imposes heightened regulation, and separates treated and control industries very differently, supply-side measures are likely to also identify the regulation well. Indeed, we observe in the Appendix Figure A.2 that the RegData measure can also identify increased regulatory restrictions for hospitals from animal hospitals.

Finally, Appendix Table A.4 reports the statistical significance of the graphical evidence of the cases in Figure 3, and Appendix Figures A.3 and A.4 show that the results are robust to using the medium and broad versions of RegIndex.

3.5.3 State RegIndex and State Voting for Republican Party

More indirect dimensions of the data may also be informative of the validity of the RegIndex methodology. Political parties support different approaches to regulation and intervention (Peltzman, 1998; Mian et al., 2010). In the U.S., Republican administrations make limiting the government burden on firms an explicit goal. Given the presence of substantial leeway at the state level in creating state-specific regulatory environments (e.g., in the case of the insurance industry or state banking for instance, see Agarwal et al. (2014)), one would expect to see lower levels of RegIndex in Republican-controlled or Republican-leaning constituencies. This sanity check is illustrated below.

We begin by estimating state-specific RegIndex averages, conditional on the state's industry composition. That is, we extract state fixed effects in each year controlling for industry fixed effects, and recover the conditional mean RegIndex for each one of the 50 states and the District of Columbia. All establishments are weighted by their total wage payment and the sampling weights assigned by the OEWS survey. Industry is defined at the NAICS 6-digit level.

Figure 4 reports the heat map of state-specific RegIndex averages in 2014. States with the highest RegIndex include Democratic party leaning Vermont, Connecticut, Delaware, Massachusetts, while states with the lowest RegIndex include Republican strongholds Alabama, Louisiana, North Dakota, Mississippi. Comparing our state RegIndex with the state RegData which counts for restrictive words in state regulatory texts since 2017,[31] a notable difference is that state RegData is heavily related to the number of businesses in the state. For instance, states with the highest RegData are California, New York, New Jersey, Ohio, Illinois, and Texas, while states with the lowest RegData are South Dakota, Idaho, North Dakota, Alaska, and Montana (see Appendix Figure A.5). States' number of establishments (from the Census SUSB) in 2017 explains 62 percent of the variation in the 2017 state RegData, where the coefficient has a t-statistics of 8.74. In contrast, states' number of establishments explains only 1 percent of our state RegIndex where the t-statistics of the coefficient is -0.74.

More systematically, state RegIndex averages correlate negatively with state political inclination to vote for the Republican Party. As an illustration, we consider states' Republican vote shares in the 2016 Presidential Election (Donald Trump vs. Hillary Clinton), the 2016 House elections, and the 2018 Senate

[31] RegData starts to count restrictive words in state regulatory texts for 16 states in 2017, 9 additional states in 2018, 18 additional states in 2018, 3 additional states in 2020, and 3 additional states in 2022. Analyzing RegData that covers the same state in multiple years reveals that state RegData is extremely stable over time, as state fixed effects explain over 99.6 percent of the variation of the pooled state RegData sample. Hence, for each state, we use the earliest available state RegData to represent the state's RegData in 2017. We download state RegData at https://quantgov-bulk-downloads.s3.amazonaws.com/State-RegData-Definitive-Edition_Regulations.zip.

elections. Table 2 shows that state-specific RegIndex is significantly and negatively related to the state Republican vote share in all three elections.[32]

3.5.4 Evidence on Publicly-Traded Firms' RegIndex

Lastly, we validate our RegIndex measure using publicly-traded firms. Motivated by the large literature examining the "disappearing IPO" puzzle (e.g. Gao et al. (2014)) and recent work by Ewens et al. (2024) examining publicly-traded firms bunching near Sarbanes-Oxley Act (SOX)'s thresholds on public float, this subsection investigates whether RegIndex can capture changes in corporate regulatory environment after initial public offerings.[33]

To perform this exercise, we merge BLS establishments to publicly-traded firms in the Compustat-CRSP database using a combination of EIN matching and fuzzy name matching.[34] Following BLS's suggested aggregation practices and prior literature, we represent each publicly-traded firm's occupational employment using all of its matched establishments surveyed in the current year and the previous two years.[35] Importantly, using the establishment unique identifier, we are able to obtain publicly-traded firms' occupational employment before they become public and enter the Compustat-CRSP merged database. Following equation (3), we then measure each publicly–traded firm's RegIndex as the percentage of its labor spending on regulation-related tasks, and track the firm's RegIndex before and after the IPO.[36] Our final sample requires firms to be incorporated in the U.S. and have a RegIndex measure. The sample ends up covering 1,024 IPO events from 2002 to 2014.

Our specification regresses a firm's RegIndex on indicators of the IPO event year from -3 to 3, with the year prior to the IPO as the benchmark, while controlling for year and firm fixed effects. Column (1) of Table 3 shows no significant pre-trend in RegIndex before the IPO. Instead, we detect a significant increase in a firm's RegIndex after the IPO, consistently with the hypothesis that being publicly traded exposes firms to a regulatory compliance burden higher than being private.[37] The economic magnitudes are also meaningful. Two years after the IPO, a firm's RegIndex increases on average by about 8 percent (= 0.151/1.798) of the pre-IPO level at $t - 1$.

[32] In the Appendix Table A.5, we further control for state RegData and state number of establishments in the regression, and we observe very similar results to Table 2.

[33] See Coates and Srinivasan (2014) for a review.

[34] See also Zhang (2019).

[35] Because the OEWS program surveys each establishment every three years, the BLS produces and publicizes aggregate statistics of occupational employment at the industry or geographic level using surveyed establishments from $t - 2$ to t, see BLS technical notes at https://www.bls.gov/oes/oes_doc_arch.htm.

[36] We obtain firm IPO dates from Jay Ritter's website.

[37] Notice that our goal for this subsection is simply to validate whether our RegIndex measure can produce findings in line with prior literature rather than fully establishing a causal effect (see Bernstein (2015) for a discussion of identification).

One concern about the exercise may be that the raw RegIndex may not account for changes in industry-level regulation for IPO firms. For instance, it is possible that regulation may tighten for all firms in a sector, private and public, after the industry undergoes a major wave of IPOs. We mitigate this concern by subtracting from the IPO firms' RegIndex the industry-year average of all publicly-traded firms' RegIndex.[38] Column (2) shows that this de-meaned RegIndex exhibits similar pattern as the original RegIndex, suggesting that our IPO results do not appear entirely driven by industry trends. For completeness, Appendix Table A.6 shows that the results are similar if one employs the medium and broad versions of RegIndex.

Finally, we examine how RegIndex relates to two other prominent measures of publicly-traded firms regulatory burden. Appendix Table A.7 reports that our measure is moderately, but consistently correlated with the measure from Armstrong et al. (2025) (correlations of 0.240-0.281 depending on the versions of RegIndex) and the measure from Kalmenovitz (2023) (correlations of 0.165-0.182).

4 Regulatory Compliance Costs and Business Size

This section examines a crucial property of RegIndex_{it} and addresses the paper's main question. It will present evidence that regulatory compliance costs fall disproportionately on medium sized businesses relative to others.

To explore economies of scale in regulatory compliance costs, we focus on the sign and magnitude of the derivative of the regulatory index with respect to firm (or establishment) employment, $\frac{\partial \text{RegIndex}_{it}}{\partial L_{it}}$.[39] We present both estimates for the whole U.S. economy and industry-specific estimates that account for heterogeneity in regulatory regimes across different sectors.

As discussed in the Introduction, economies of scale in regulatory compliance are a key feature of any regulatory architecture. Diseconomies of scale introduce a potential deterrent to firm growth, pushing firms to operate below their efficient scale of production. Regulation may also introduce incentives toward concentration and may act as a barrier to entry, favoring large incumbents.[40] While the issue of returns to scale in regulation has received much attention in the Political Economy and Industrial Organization literature in some specific industries, to the best of our knowledge, this paper is the first to provide a

[38] We use the NAICS 4-digit code from Compustat to assign a publicly-traded firm's industry.

[39] For simplicity of exposition and with a limited abuse of notation, we will refer to L_{it} as "employment" (as opposed to the proper total employment given by the sum $L_{it} + R_{it}$). Given the magnitudes that we report in this article, this approximation is warranted, and it makes both exposition and analysis much clearer. In the empirical analysis, for accuracy, we employ $L_{it} + R_{it}$.

[40] Classic references are Stigler (1971); Peltzman (1976).

comprehensive set of facts representative of the entirety of the U.S. economy.

The simple framework in equation (1) allows one to explicit several factors driving economies (or diseconomies) of scale under a limited set of assumptions. There are at least two factors. First, in the model, fines imposed by regulators are a function of requirements imposed by the rules, $\tilde{R}_{it}$. These standards are, in turn, a function of size, $\tilde{R}_{it} = \tilde{R}(L_{it})$. Importantly, it is plausible to hypothesize $\frac{\partial \tilde{R}_{it}}{\partial L_{it}} \gtreqless 0$.[41] An example of a positive derivative is capital requirements imposed on large bank holding companies kicking in at several thresholds for total assets and as a function of the systemic importance of the financial institution (both measures correlate with employment). Another is the regulatory tiers discussed in Brock and Evans (1985). An example of a negative derivative is instead presented in Hopkins (1995), which shows that the smallest firms in his sample have paperwork and tax compliance costs (measured against turnover) about twice as large as those of the largest firms.

As a second factor, scale matters through the probability of inspection p_{it}. The inspection probability is naturally driven by government enforcement effort, which can be a function of size, $E_{it} = E(L_{it})$. One may plausibly posit $p_{it} = p(E_{it})$, with $\frac{\partial p_{it}}{\partial E_{it}} \geq 0$, and $\frac{\partial E_{it}}{\partial L_{it}} \gtreqless 0$, where this derivative may be positive if larger firms have more weight in inspection protocols or negative if, for instance, smaller plants are easier/faster to inspect.[42]

These considerations suffice to illustrate a theoretical ambiguity in the relationship between regulatory costs and scale. Using equation (1) and the discussion above, we have $R_{it} = \left(\frac{p(E_{it})\tilde{R}_{it}}{w^{r}_{it}}\right)^{\frac{1}{2}}$, where $\frac{\partial R_{it}}{\partial L_{it}} \gtreqless 0$ depending on the dominating force. Importantly, using the definition (2), it follows that the sign of $\frac{\partial \text{RegIndex}_{it}}{\partial L_{it}}$ is also ambiguous.

This ambiguity is borne out by the data. Figure 5 presents firm and establishment-level nonparametric evidence of a non-monotonic relationship between total employment size and RegIndex. The dots in the graph represent averages for employment bins of [1, 2] employees, [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. As firm or establishment employment increases regulatory costs per employee increase steadily until a firm size of around 500 workers, then regulatory compliance costs per employee start falling rapidly, indicating economies of scale. We estimate the percentage of labor costs for regulatory compliance for

[41] In addition, note that fines may be a function of the firm's covariates and size (in addition to the level of compliance exerted by the firm relative to a given standard required by the rules, $\tilde{R}_{it}$). That is, one could posit a general $f_{it} = f(L_{it}, R_{it}, \tilde{R}_{it})$ with $\frac{\partial f_{it}}{\partial L_{it}} \gtreqless 0$, where a positive derivative case can arise if fines are designed to be more than proportional to firm size "to set an example", and negative if fines are capped by statutory limits/by the threat of litigation from large firms.

[42] See Helland (1998) for a discussion related to EPA in the United States. See also Shimshack (2014) for a review of the evidence.

mid-size businesses are about 40-50 percent higher than that for the smallest businesses and about 10-20 percent higher than that for the largest ones.

The evidence from a parametric representation of the non-monotonicity is reported in Table 4, which includes the *max* and *argmax* of an inverted-U relationship between RegIndex and size for both firms and establishments. The coefficients of the parametric regression are significant and precisely estimated across all different specifications. The specifications in the table include different sets of fixed effects to assess the robustness of the finding: Year FE; Year×Industry FE; Year×Industry×State FE; Year FE + Firm FE. The table reports a range for *argmax* in panel A of about 499-511 total employees for firms at the peak of regulatory compliance costs and an *argmax* in panel B of about 309-344 for establishments. As the average firm in the United States includes only 1.26 establishments from 2002 to 2014 according to the Census SUSB, it is not completely surprising that *argmax* aligns between Panels A and B. Finally, Appendix Figure A.6 shows that our finding holds for all three versions of RegIndex.

Figure 6 reports the same information by large sector aggregates, highlighting a degree of heterogeneity in the presence of regulatory scale economies. The pictures present evidence both at the firm level in Panel A and the establishment level in Panel B. Both figures show a non-monotonic relation between RegIndex and size within industries such as finance, retail, and other services. The non-monotonicity is less pronounced in others, such as manufacturing and utilities, as one would expect from the vast differences in sector specific regulations.

Figure A.7 provides a visual representation of the RegIndex across employment bins each year from 2002 to 2014. While RegIndex increases for all size bins over time, the inverted-U shape appears highly stable over time, with some changes in the left tail of the size distribution (firms with less than 10 employees), which appears to become flatter from 2002 to 2014. A possible reason for this pattern is the relative reduction of the regulatory burden on small firms over this time period.[43]

[43]An example arises from the Small Business Administration (SBA)'s advocacy in favor of small firms. In the August 2002, Executive Order 13272 by President George W. Bush further implemented the Regulatory Flexibility Act of 1980, which is the act codifying the SBA. Over the period 2002-2009 one can find several instances in which SBA Office of Advocacy actively lobbied to reduce compliance costs on small businesses. Examples include in 2008, when the EPA finalized the Lead Renovation, Repair, and Painting Rule requiring contractors working in pre-1978 homes to be certified and follow specific lead-safe work practices. The rule was simplified in its requirement for small residential construction and renovation companies, lessening costs and complexity of compliance for small-business contractors. With the passage of the Small Business Jobs Act of 2010, the Office of Advocacy further strengthened. This, combined with the vast regulatory intervention post 2009 financial crisis (e.g. the Dodd-Frank Act of 2010, the ACA of 2010), likely contributed to a relative reduction in compliance costs for small businesses over 2010-2014. For instance, in the case of the Affordable Care Act of 2010, the threshold of 50 employees for the application of the Employer Mandate is an area where the Office of Advocacy actively intervened, including by helping simplify reporting requirements, facilitating tax compliance for firms around the threshold, and advocating for tax credits for health care for affected business just above the regulatory "cliff".

4.1 Robustness

4.1.1 Within-Occupation Variation

One may be concerned that workers in the same occupation may in practice perform more regulation-related tasks in small firms than in mid-size firms. For instance, small firms may hire non-regulation-related occupations to cover non-regulation-related tasks and also cover some regulatory compliance tasks. To assess how firms' requirements on regulation-related tasks vary with their size for a given occupation, we investigate firms' skill requirements in their job posting descriptions using over 14 million job postings during 2010-2014 from the Burning Glass Technologies (BGT) data.

The BGT data provide 17,420 skills extracted from the millions of job posting descriptions. For each skill, the data also provide a skill definition, usually one sentence like the task statements in the O*NET data. We identify a BGT skill as "regulation-related" using the same procedure that identified regulation-related tasks in Section 3.3. This procedure yields 523 regulation-related skills in the BGT data. Averaging the regulation-related dummies for all skills within a job posting generates a continuous regulation-relatedness measure for each job posting. Finally, we name match BGT firms to OEWS firms and regress the regulation-related measure of job postings on the posting firms' employment and employment squared. Because the regulation-related measure is at the job posting level and varies within an occupation, we can now explore small and large firms' requirement on regulatory compliance when hiring the same occupation in the same year. This allows, for example, one to investigate whether the regulatory skill requirements for an electrical engineer are particularly high for an engineer working in a small firm relative to a large one.

Figure 7 plots the average of the demeaned regulation-relatedness measure for job postings of occupations with positive RTI in each firm employment bin, where each job posting's regulation-relatedness measure is demeaned by subtracting the average regulation-relatedness measure of all job postings in that occupation in that year. The figure shows a similar inverted-U shape where firms with around 500 employees have greater requirements on performing regulatory compliance tasks than small and large firms when hiring the same occupation in the same year.

Table 5 shows the results more formally by regressing each job posting's regulation-relatedness measure on firm size while controlling for Year$\times$Occupation fixed effects. Column (1) provides evidence that firms' demand for regulatory compliance tasks also exhibits an inverted-U relationship with firm size, though with the peak at around 800 employees. Specifically, for the same occupation, mid-size firms appear to require more regulatory compliance skills than small and large firms. Columns (2) and (3) further show that mid-size firms require more regulatory compliance skills than small and large firms regardless of hiring regulation-related occupations (RTI > 0) or non-regulation-related occupations (RTI = 0).

Overall, these results suggest that our finding on the inverted-U relationship between RegIndex and firm size is a conservative estimate. The relationship between firms' regulatory compliance costs and their size would show an even more pronounced inverted-U shape if we were to fully account for within-occupation variation in regulatory compliance requirements.

4.1.2 Capital

Next we consider whether the findings so far may be sensitive to (or even entirely driven by) small, medium, and large firms' heterogeneous use of labor and physical capital in complying with regulation.

As discussed in Section 3, precisely estimating firms' use of capital for regulatory compliance purposes is challenging, mostly due to the lack of detailed data on various capital user costs at the micro-level in the OEWS. Nonetheless, following Caunedo et al. (2023) and focusing on O*NET tools and capital equipment requirements of occupations, we are able to construct an approximate measure of share of regulatory compliance costs out of the total labor and capital costs in Section 3. This measure, $\text{RegIndex}_{i,t}^{tot}$, is employed in replicating Figure 5.

Figure 8 shows how labor-based $\text{RegIndex}_{i,t}$ and total-input-based $\text{RegIndex}_{i,t}^{tot}$ compare in terms of relationship with firm and establishment size. We observe again a distinctive inverted-U shape for $\text{RegIndex}_{i,t}^{tot}$, proving the robustness of our main finding. There are, however, some important differences relative to $\text{RegIndex}_{i,t}$. The inverted-U shape traced by using $\text{RegIndex}_{i,t}^{tot}$ coincides with $\text{RegIndex}_{i,t}$ for small firms and establishments, as expected given the limited role of physical capital for very small units. However, the $\text{RegIndex}_{i,t}^{tot}$ curve remains consistently lower than $\text{RegIndex}_{i,t}$, indicating, as mentioned above, that labor inputs play a disproportionately important role relative to capital inputs in regulatory compliance. Further, the distance between the two curves widens as firm and establishment increase in size, indicating that capital plays an increasingly relevant role in regulatory compliance cost savings as firm grow in size.[44]

[44]It can also be possible that the regulation-related capital expenditure is more difficult to consolidate than labor costs, reducing the economies of scale for larger firms.

4.1.3 Outsourcing

Firms also hire outside compliance professionals to comply with regulation. In this subsection we explore whether outsourcing may be a major source of bias for our main findings.

How outsourcing may affect our main results is a priori unclear. On the one hand, small firms may have greater incentives to outsource compliance tasks due to limited resources and lack of economies of scale in in-house compliance. On the other hand, large firms may prefer outsourcing due to an increased focus on streamlining operations and allocating in-house workforce to "core competencies" (e.g., Goldschmidt and Schmieder (2017)). Moreover, if outsourcing itself involves fixed costs –such as those associated with searching for service providers and setting up business with external suppliers– larger firms may find even more economic than small firms to outsource compliance.

While we do not observe firm-to-firm contracts for regulatory compliance outsourcing in our data, this subsection provides two sets of facts suggesting that outsourcing is unlikely to qualitatively affect our main finding. We first leverage two major industry survey reports in order to assess how small and large firms outsource in practice. A first report is from the National Association of Manufacturers (2014) and provides the likelihood for small, medium, and large firms to hire outside advisers in complying with federal regulations for the manufacturing sector. Chart 3 of National Association of Manufacturers (2014) shows that this likelihood is smallest for small firms, 58 percent, increasing to 70 percent for medium firms, and 82 percent for large firms. Based on this monotonic relation, heterogeneous outsourcing appears unlikely to overturn our findings that RegIndex is lower for small firms than for medium and large firms in Figure 6.[45] This finding is also consistent with Chayes and Chayes (1985) showing that in-house and outsourced regulatory spending complement each other.

A second report from the Securities Industry Association (2006) shows that, while outsourcing spending accounts for only 2.8 percent of financial sector firms' total compliance-related spending (as compared to 93.9 percent of in-house labor spending), outsourcing spending accounts for a larger fraction relative to in-house labor spending in small firms than in large firms. Calculations based on figure 3b of Securities Industry Association (2006) show that outsourcing spending is about 16.4 percent, 3.2 percent, and 2.2 percent of the in-house compliance-related labor spending for small, medium, and large firms, respectively. One can further evaluate the impact of outsourcing spending by computing an outsourcing-adjusted RegIndex adding outsourcing costs to both the numerator and denominator of equation 4.

[45] One limitation is that this report only shows the extensive margin but not the intensive margin for small, medium, and large firms' compliance outsourcing.

Appendix Figure A.8 shows that such adjustment does not overturn our main findings either.

A second set of facts can be inferred from a more comprehensive assessment of how accounting for costs paid to external compliance specialists affects our main findings. We construct a measure of outsourcing costs for establishments and firms based on the following intuition. Suppose we consider a certain geographic market, say an US county, and focus on the employment distribution in the county across small, medium and large establishments. Suppose further that the market for compliance service providers is localized in nature. Under these assumptions, we can test whether small establishments utilize more outsourcing services than medium or large establishments. That is, if one were able to compare two counties similar in everything other than for the fact that one has a larger share of employment in small establishments than the other, we should observe a higher presence of outsourcing service providers in the former relative to the latter.

With this idea in mind, we develop a co-location estimator across US counties to assign outsourced compliance costs to establishments in different establishment size bin of the inverted-U shape plot. Our assumption behind this estimation is that in the period from 2002 to 2014, outsourcing is largely provided via on-site services and that establishments of different sizes have a different utilization rate of such services.[46]

Operationally, our estimates focus on five major compliance service providers: legal, accounting and tax, payroll, IT, and management consulting services. For all these outside service providers we observe detailed wage bill information. As these providers' production costs are ultimately paid by client establishments (i.e. businesses that outsource their compliance), we can conservatively use these providers' labor spending on regulation-related tasks to proxy for the total outsourced regulatory compliance client spending each year.[47] We define the outsourcing services intensity in each county, $OutSource_{c,t}$, as the total labor spending on regulation-related tasks from outsourcing providing establishments in county c divided by all other establishments total labor spending in c.[48] We then compute each county's establishments' size distribution, $Share_{j,c,t}$, as the share of labor spending from each of the 14 size bins described in Section 4, where j represents each size bin. Finally, we regress $OutSource_{c,t}$ on the 14 $Share_{j,c,t}$ with-

[46] As an example, Charoenwong et al. (2024) show that technology-driven compliance (i.e., RegTech) in finance arises after SEC amendments in 2014.

[47] The five compliance outsourcing industries are 5411 (Legal Services), 5412 (Accounting, Tax Preparation, Bookkeeping, and Payroll Services), 5415 (Computer Systems Design and Related Services), 5416 (Management, Scientific, and Technical Consulting Services), and 5419 (Other Professional, Scientific, and Technical Services).

[48] As a first pass check, we plot in Appendix Figure A.9 the relation between RegIndex and establishment size in three subsamples of counties with low, medium, and high outsourcing abundance, defined by terciles of $OutSource_{c,t}$ in the year. We observe that small establishments in counties with low outsourcing abundance counties still have lower RegIndex than mid-sized establishments in those counties, demonstrating the robustness of our inverted-U shape results to the impact of outsourcing.

out imposing a constant term, while controlling for major county characteristics including the county's total employment, state and urban fixed effects, and year fixed effects.[49] This regression generates 14 coefficients for the size bins representing the relative usage of compliance outsourcing for different establishment sizes. A higher coefficient here proxies for a higher intensity of using compliance outsourcing for establishments in that bin relative to those in other bins.[50] Figure A.10 plots these estimates. To quantify the impact of outsourcing, we estimate the outsourced compliance costs in addition to in-house compliance costs, $RegIndex^{O}$, for establishments in each size bin.

We present two versions for robustness. Our first version assumes that establishments only use outsourcing service providers within their county, as per the co-location estimation's assumption. Our calculation procedure is the following. First, denote $RegIndex^{O}_{j,c,t}$ as the average outsourced compliance costs normalized by total labor costs for establishments in size bin j and county c. Denote β_j as the relative outsourcing intensity for each size bin j estimated above, we can express $RegIndex^{O}_{j,c,t} = x_{c,t} \times \beta_j$, where $x_{c,t}$ is a benchmark value that we will solve for later. Second, note that county c's observed outsourcing intensity, $OutSource_{c,t}$, is the average of $RegIndex^{O}_{j,c,t}$ weighted by the total labor costs of establishments in size bin j and county c at t, $w_{j,c,t}$. Third, solving for $x_{c,t}$ based on the above setting results in $RegIndex^{O}_{j,c,t} = (OutSource_{c,t} \times \frac{\sum_s \beta_s \times w_{s,c,t}}{\sum_s w_{s,c,t}}) \times \beta_j$,[51] which we assign as the outsourced compliance costs for each establishment i in size bin j and county c at t. Finally, adding establishment i's estimated $RegIndex^{O}_{i,t}$ to its baseline (in-house) RegIndex results in the establishment's RegIndex with outsourced compliance costs, $RegIndex^{I+O}_{i,t}$.

Our second version assumes that all establishments use services from compliance outsourcing firms nationwide. In this case, establishments in each size bin, regardless of location, have the same $RegIndex^{O}$ in a year, which we back out through the same procedure as described above but use the nationwide outsourcing intensity (instead of the county-specific $OutSource_{c,t}$) in each year.

These two versions are based on polar assumptions about the local constraints of regulatory compliance outsourcing and thus provide bounds on the impact of outsourcing costs. We also compute the two versions at firm level (EIN) by averaging an EIN's establishments' $RegIndex^{I+O}_{i,t}$ weighted by the

[49] We obtain each county's 2010 urban population share from the Census website at https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/2010-urban-rural.html. A county is classified as urban if more than half of its population resides in urban areas, as opposed to rural areas.

[50] Establishments from the largest size bin are sparsely distributed across counties, making the outsourcing intensity estimate for that bin volatile. We keep the 14 bins in this estimation to be consistent with our baseline specification. Nevertheless, our focal test result on how outsourcing affects small versus mid-sized establishment RegIndex is very similar with and without including the largest size bin in the estimation. Given the volatile results for the largest size bin, we exclude this bin in the relevant plots.

[51] In particular, we solve for $x_{c,t}$ from $OutSource_{c,t} = \sum_j \frac{RegIndex^{O}_{j,c,t} \times w_{j,c,t}}{\sum_j w_{j,c,t}} = \sum_j \frac{x_{c,t} \times \beta_j \times w_{j,c,t}}{\sum_j w_{j,c,t}}$.

establishments total labor costs.

Figure 9 plots the two versions of firm-level $RegIndex^{I+O}_{i,t}$ and the baseline RegIndex by size bins in Panel A and similarly at the establishment level in Panel B. We observe that, while outsourcing costs disproportionately increase RegIndex for small firms, it does not change the qualitative inference that small firms pay proportionately lower regulatory compliance costs than mid-size firms. This finding mitigates the concern that compliance outsourcing may overturn the left-hand size of our inverted-U shape results.

One final potential concern is that only small outsourcing service providers may focus on serving local clients, while large outsourcing providers may serve clients nationwide, making large outsourcing providers less suitable for our co-location estimation. If small outsourcing providers serve smaller firms more, we are likely to underestimate the outsourcing intensity of small firms. To mitigate this concern, we rerun our co-location estimation using only small outsourcing providers in the county, while assuming large outsourcing providers as affecting all establishments' RegIndex equally. To the extent that large outsourcing providers serve larger establishments more, this re-estimation is likely to overestimate small establishments' outsourcing intensity relative to larger establishments, providing an upper bound for outsourcing's effect on overturning our main results. We define small outsourcing providers as those with 20 or fewer employees. Panels C and D of Figure 9 show that accounting for costs paid to small outsourcing providers, $RegIndex^{I+O}_{i,t}$ is still lower for small firms than mid-sized firms. We also experiment with different cutoffs for defining small outsourcing providers such as using 10 or 50 employees, and we find very similar results (see Appendix Figure A.11). Finally, our main findings are robust if we rerun the estimation using medium or broad versions of RegIndex throughout (see Appendix Figure A.12).

Overall, the evidence in this subsection suggests that accounting for outsourcing appears unlikely to overturn the key implications of our analysis.

4.1.4 Multi-State Regulations

As an final robustness check, we examine whether firms' heterogeneous exposure to multiple state regulations plays an outsized role in driving the inverted-U shape. In particular, one may be concerned that some of the non-monotonicity in Figure 5 may originate from a specific mixture of single-state firms, exposed to limited regulation, and medium-large multi-state firms, facing the extra complexities of complying with multiple states' regulations. Looking at single-state versus multi-state firms in isolation may help assuaging this concern, especially if similar non-monotonicities are present for all types of firms.

We classify firms as follows. First, we identify single-unit establishments (based on a flag measure provided by the BLS) as the obvious group for single-state firms. These establishments account for over 80% of all establishments in our final sample. Second, if the establishment's parent EIN spans multiple states, we classify the firm (and establishments) as obviously part of a multi-state business. Third, we consider the residual group of firms (not belonging to the first or the second group) as likely to be single-state.[52]

Figure 10 plots the relationship between RegIndex and firm size for single-unit firms, likely-single-state firms, and multi-state firms. We observe that both single-unit firms and likely-single-state firms show a strong inverted-U shape – clearly not a flat relationship. This is reassuring, as for example it rules out that our inverted-U shape result may be solely the result of omitted heterogeneity in the regulatory compliance profiles of multi-state versus single-state firms.

The RegIndex for multi-state firms shows two interesting features. First, multi-state firms show higher level of RegIndex compared to same-size single-state firms, confirming that complexity of multi-state regulation increase firms' regulatory compliance costs. Second, we observe inverted-U shape for firm-level RegIndex for multi-state firms, but a largely increasing shape for establishment-level RegIndex for multi-state firms. These differences are consistent with multi-state firms concentrating their regulatory compliance costs within few establishments rather than decentralizing their regulatory compliance, possibly with the goal of achieving economies of scale.

5 Mechanisms and Implications

This section presents empirical evidence on the mechanisms behind the key results of Section 4. We begin by investigating economies of scale, probably the most intuitive driver of the decreasing component of the inverted-U shape traced by Figure 5. Section 5.1 shows evidence consistent with the economies of scale mechanism for large firms to have a lower RegIndex than mid-sized firms. To understand why small firms have a lower RegIndex than mid-sized firms, we engage in a more complex exercise in Section 5.2, designed to decouple the different roles of regulatory requirements and of enforcement in driving compliance costs across different employment sizes, finding evidence consistent with regulatory tiering as a potential additional mechanism behind our findings.

[52] Some of them may be multi-state if a firm presents multiple EINs across states or if the OEWS sampling misses certain EINs in other states, but most appear single-state firms with multiple establishments within a state.

5.1 Fixed Costs

We begin by exploring what is intuitively a simple, but important economic mechanism behind the non-linear relation between compliance costs and firm size: fixed costs in regulatory compliance.

Evidence that part of the increasing returns to scale in regulatory compliance (i.e. the downward part of the parabola traced in Figure 5) may derive from fixed costs can be inferred from the growing presence of regulatory compliance specialists as firm size increases. Compliance specialists may in fact offer more efficient handling of regulatory requirements, but may be affordable only for firms reaching a certain scale. The OEWS data allows us to carefully perform this check through its detailed occupation information.

For each firm's RegIndex, we compute the fraction of RegIndex that is performed by regulatory specialists, where specialists are defined using the top 25 occupations with the highest regulation-task intensity (see RTI in Appendix Table A.2). The higher the share of RegIndex performed by specialists, the more the firm concentrates its compliance in a dedicated subset of its workforce.

The relationship between regulatory compliance specialization and size is reported in Figure 11, which captures the increasing firms' reliance on regulatory specialists as their size increases. The figure shows that specialization of compliance is almost monotonically increasing with employment. This indeed suggests that larger firms pay the fixed costs of hiring compliance specialists, who can then consolidate regulatory compliance tasks and perform them more efficiently. The increment is quantitatively sizable as the spending in regulatory compliance through compliance specialists appears to almost triple for firms between 50 and 5000 employees. We also verify that the evidence of increasing specialization in regulatory compliance is present across all different industries and it is not driven solely by certain sectors, such as finance or utilities.

5.2 Size-Dependent Regulatory Requirements and Enforcement

As discussed in Section 3, the share of regulatory compliance costs RegIndex_{it} may be driven by both the extent of regulatory agencies' regulatory requirements, $\tilde{R}_{it}$, and endogenous compliance due to enforcement effort, E_{it} (which affects the likelihood of inspection p_{it}). In particular, from equations (1) and (2), we know $\text{RegIndex}_{it} = \frac{\left(w^r_{it} p_{it} \tilde{R}_{it}\right)^{0.5}}{W_{it}}$, where $\tilde{R}_{i,t}$ captures the intensity of regulation requirements and $p_{i,t}$ captures the enforcement stringency of agencies. These competing drivers, which may be size-dependent, introduce a challenge in interpreting the findings in Section 4. For example, a small firm may exhibit low regulatory compliance costs because the agency rules are designed to be lighter on small firms (proxied by lower $\tilde{R}_{it}$) or because small firms comply less, as they know they are more likely to fly under the radar of inspectors (proxied by lower p_{it} for smaller firms). These considerations are relevant to decomposing the non-monotonic shape of the relationship between compliance costs and firm size in Section 4.

We present a methodology to analyze the impact of regulation requirements and enforcement on the inverted-U relation between RegIndex and firm size. Our approach takes two steps. First, we use changes in regulation requirements and enforcement to estimate each regulation-related *task*'s sensitivity to requirements and to enforcement.[53] Second, we construct two sub-components of RegIndex: the share of a firm's labor spending on requirement-sensitive tasks, $\text{RegIndex}_{i,t}^{req}$, and the share of a firm's labor spending on enforcement-sensitive tasks, $\text{RegIndex}_{i,t}^{enf}$. This decomposition is constructed so that $\text{RegIndex}_{i,t} = \text{RegIndex}_{i,t}^{req} + \text{RegIndex}_{i,t}^{enf}$. We then analyze the role of each subcomponent in driving our main results.

5.2.1 Methodology for Estimating Task Sensitivity to Requirements and Enforcement

Assume that certain regulation-related tasks may be more relevant for complying with regulatory requirements, while other regulation-related tasks may be more relevant for dealing with enforcement. Firms respond to changes in regulatory requirements ($\tilde{R}_{it}$) by adjusting spending on *requirement-sensitive tasks* and respond to changes in enforcement (p_{it}) by adjusting spending on *enforcement-sensitive tasks*.[54] To estimate each regulation-related task's sensitivity to regulation requirements and enforcement, we develop a methodology based on a shift-share approach.

For each regulatory agency k at time t, suppose being able to measure both its regulatory requirements, reg_{kt}, and the extent of the enforcement and supervision effort, enf_{kt}. Further, suppose we are also able to measure industry j's regulation originating from agency k at time t, r_{jkt}, i.e. the industry-specific *exposure* to each regulatory agency.[55] How to perform these operations is discussed below in Section 5.2.2.

[53] For instance, one would expect that the task to "*... learn about and analyze the potential impacts of proposed environmental policy regulations*" (O*Net task ID 19529) to be more sensitive to new regulatory requirement, while the task to "*... coordinate investigative efforts with law enforcement officers and attorneys*" (O*Net task ID 16055) to be more sensitive to regulatory enforcement.

[54] Some regulation-related tasks will likely be sensitive to both requirements and enforcement. Our estimation allows for such "mixed" tasks.

[55] We define the agency exposure at the industry level rather than firm level due to an empirical limitation. The only information we possess that can inform firms' exposure to regulatory agencies is the regulation-related tasks they perform. A non-trivial among small firms in our data do not perform regulation-related tasks (consistent with our inverted-U finding). However, this does not mean the small firms are not exposed to regulatory agencies' future changes in regulation requirements and enforcement. Defining the agency exposure at the industry level ensures us to capture all firms' exposure to future regulatory changes, despite that they can have different responses to the future changes as we show in Section 5.2.3.

It is possible to create two shift-share variables tracing changes in regulatory requirements and in enforcement pertinent to firm i in industry j as the interaction of shifts and shares:

$$\begin{aligned} iv(\Delta\log(\tilde{R}_{it})) &= \sum_k \Delta \log reg_{kt} \times r_{j(i)kt} \\ iv(\Delta\log(p_{it})) &= \sum_k \Delta \log enf_{kt} \times r_{j(i)kt}, \end{aligned} \tag{5}$$

where $\Delta \log reg_{kt}$ and $\Delta \log enf_{kt}$ are the agency-specific requirement (log) changes and enforcement changes from year t to $t+3$, and $r_{j(i)kt}$ is industry j's exposure to agency k's regulations in year t.

One can now categorize each regulation-related task o as *requirement-sensitive* or *enforcement-sensitive* by estimating the following regression for each task o:

$$\Delta\log(\text{RegIndex}_{iot}) = \alpha_o + \beta_o iv(\Delta\log(\tilde{R}_{it})) + \gamma_o iv(\Delta\log(p_{it})) + \delta_o X_{it} + \xi_{it}, \tag{6}$$

where $\Delta\log(\text{RegIndex}_{iot})$ is firm i's change in spending on task o from year t to $t+3$.[56] In equation (6) β_o and γ_o respectively measure task o's sensitivity to regulatory requirement changes and to enforcement changes.[57]

In practice, we run the regression (6) for each regulation-related task o on standardized shift-shares and categorize a task as follows: If $\beta_o > \gamma_o$ and is statistically significant at five percent level, we regard task o as *requirement-sensitive*. If $\beta_o < \gamma_o$ and is significant at five percent level, we regard task o as *enforcement-sensitive*. Finally, if the two coefficients are not statistically different at five percent significance level, we regard task o as *mixed* (half requirement-sensitive and half enforcement-sensitive).[58]

Following this procedure, our estimation below shows that 49% of regulation-related tasks are labeled as *requirement-sensitive*, 16% as *enforcement-sensitive*, and 35% as *mixed*. See Appendix Table A.8 for examples. Once each task is categorized, we next adopt the exact definition in equation (2) to create a requirement-specific $\text{RegIndex}^{req}_{i,t}$ by aggregating only the percentage costs for performing the

[56]Our empirical implementation adopts the Haltiwanger percentage growth rate to proxy for the log change, $\Delta\log(\text{RegIndex}_{iot}) \cong \frac{\text{RegIndex}_{iot+3} - \text{RegIndex}_{iot}}{(\text{RegIndex}_{iot+3} + \text{RegIndex}_{iot})/2}$, to avoid taking logarithms of 0 (consider, for example, an firm changed from zero spending on task o to a positive spending on task o in response to regulatory changes). The Haltiwanger growth rate is 98% correlated with log changes in RegIndex in observations where both measures are available. We choose the three-year horizon for all log changes in equation (6), as well as the two IVs, because establishments are covered by the OEWS sample in every three years.

[57]We also include firm level controls X_{it}, firms' initial RegIndex level to account firm heterogeneity, year fixed effects to account for macro conditions, and changes in firms' total labor costs and in the average wage rate of firms' regulation-related tasks, which are motivated by the framework in Section 3.2. In particular, recall that plugging equation (1) into equation (2) and taking log changes, follows $\Delta\log(\text{RegIndex}_{it}) = \frac{1}{2}\Delta\log(\tilde{R}_{it}) + \frac{1}{2}\Delta\log(p_{it}) + \frac{1}{2}\Delta\log(w^r_{it}) - \Delta\log W_{it}$.

[58]We also conduct robustness analyses using ten or one percent significance level to label tasks. The results are qualitatively similar and available from the authors upon request.

requirement-sensitive tasks with the firm (establishment), and an enforcement-specific $\text{RegIndex}_{i,t}^{enf}$ by aggregating only the percentage costs for performing the enforcement-sensitive tasks.

5.2.2 Empirical Implementation

Let us now come back to the empirical components of the two variables in equation (5). Each regulatory agency's regulatory requirement changes, $\Delta \log reg_{kt}$, enforcement changes, $\Delta \log enf_{kt}$, and each NAICS 6-digit industry's exposure to each regulatory agency, r_{jkt}. We select twelve major regulatory agencies for which we can measure changes in regulatory requirements and enforcement. These agencies together account for 82% of non-Treasury regulatory compliance burden (by hours) in the U.S. during 2002-2014. See all details for the construction in the Appendix B.

We measure changes in each major agency k's regulatory requirements, $\Delta \log reg_{kt}$, based on changes in regulations from agency k from t to $t+3$. For each fiscal year, major agencies need to file to the White House OIRA their estimates of the changes in regulatory compliance hours for regulations under their oversight (Kalmenovitz, 2023). We use 3-year log differences in reported compliance hours due to changes in enactment and retirement of regulations to measure regulation-requirement changes of the agency.[59]

We measure changes in each major agency k's enforcement, $\Delta \log enf_{kt}$, using panel data of U.S. federal government employees from 2002 to 2014.[60] This individual-level database is compiled by the U.S. Office of Personnel Management's (OPM) Enterprise Human Resources Integration System. The data are made available by BuzzFeed News through a Freedom of Information Act request. The data cover detailed information of all federal employees, except the Department of Defense, at a quarterly frequency.[61] Variables crucial for our study include the employee's agency, occupation, and full-time/part-time employment status.[62] We identify each occupation as "regulation-related" using the same list of keywords and procedure in Section 3. We use 3-year log differences in regulation-related full time employment in each agency to measure enforcement changes of the agency.

[59] To account for regulatory changes only due to enactment and retirement of regulations, we exclude changes in regulatory compliance hours due to agencies' re-estimation of the compliance hours for the existing regulations in our calculation. Data can be downloaded from the "Information Collection Budget of the United States Government" reports at https://www.whitehouse.gov/omb/information-regulatory-affairs/reports/#ICB.

[60] See Levinson (1996); Jackson and Roe (2009); Jackson (2007); Muehlenbachs et al. (2016) for examples of prior studies using regulatory agencies' employment to measure their enforcement efforts.

[61] The data includes 206 million observations and can be downloaded at https://archive.org/download/opm-federal-employment-data. We thank Joe Raffiee for introducing this data to us.

[62] Importantly, while the data adopt a different occupation classification system from SOC, we are able to obtain each federal employee occupation's task description from the "Handbook of Occupational Group and Families" on the OPM website. The OPM has its own definitions for government occupations that are different from the SOC system. The handbook for OPM occupation description can be downloaded at https://www.opm.gov/policy-data-oversight/classification-qualifications/classifying-general-schedule-positions/occupationalhandbook.pdf.

Appendix Figure A.13 traces the time series of the two changes ($\Delta \log reg_{kt}$ and $\Delta \log enf_{kt}$) for each of the twelve regulatory agencies. We observe that there is substantial independent variation in the two change series across all regulators, although for some agencies the separation is starker. Our estimation procedure in equation (6) exploits the independent variations of the two changes to assign whether a task is requirement-sensitive or enforcement-sensitive by running regressions that include both variables.

The shift-share variables in equation (5) also require measuring each industry's exposure to regulatory agencies' changes, r_{jkt}. We compute this exposure based on the textual similarity between the regulation-related tasks that the industry performs and the rules that agency's regulation specifies. This approach is similar to Al-Ubaydli and McLaughlin (2017) who measure the probability that a given regulatory restriction targets a specific industry based on industry descriptions and Armstrong et al. (2025) who measure publicly-traded firms' exposure to various government agencies based on their 10-K filings. Specifically, the approach takes two steps. First, we extract the top 50 identifying keywords for each regulatory agency using natural language processing of the Code of Federal Regulation text,[63] and we compute a Google BERT similarity between each regulation-related task o and an agency k's keywords, $\phi_{k,o}$.[64] We standardize the twelve similarities for each task to sum up to one, i.e., $\sum_k \phi_{k,o} = 1$. These standardized similarities capture a regulation-related task's exposure to the twelve agencies.[65] Second, we have each industry j's percentage labor spending on each regulation-related task o, $s_{j,o,t}$. Hence, we compute industry j's beginning-of-period exposure to agency k's regulations at t (r_{jkt} in equation (5)) as weighted average of its regulation-related task's exposure to agency k, i.e., $r_{jkt} = \sum_o s_{j,o,t} \times \phi_{k,o}$. With all elements in equation (5) now available, we can construct $iv(\Delta\log(p_{it}))$ and $iv(\Delta\log(\tilde{R}_{it}))$. Applying the two shift-share variables to equation (6), we obtain each regulation-related task's sensitivities to regulatory requirement changes and enforcement changes, which ultimately help us label each task into requirement-sensitive, enforcement-sensitive, or both following the methodology in Section 5.2.1. Once

[63] We obtain the keywords in three steps: First, we manually select all the CFR chapters that mentions the names of the 12 agencies and their sub-agencies, where the CFR texts are available at https://www.govinfo.gov/app/collection/cfr/2021/. For instance, the CFR chapters that mention Environmental Protection Agency are Title 40 Chapter I (Part 1-1099) and Chapter IV (Part 1400-1499). Second, we count the term frequency of each word in the overall selected CFR text ($count_{all}$) and also in the texts about each agency k ($count_k$). We identify keywords as specific to agency k if the keywords appear over 50 percent of the time in agency k's texts alone, i.e., $count_k/count_{all} \geq 0.5$. Third, we remove uninformative words from the list, such as abbreviations by computing the similarity of the keywords to its agency's name using Google Bidirectional Encoder Representations from Transformers (BERT). Our final keywords are the top 50 keywords that have the highest BERT similarity score. Appendix Table A.9 lists the 50 keywords for each of the 12 agencies.

[64] Google BERT similarity captures contextual and semantic meaning in text by leveraging deep language representations, making it more accurate for understanding nuanced or ambiguous language compared to standard cosine similarity. In contrast, standard cosine similarity relies on static vector representations and often fails to account for word order or context, limiting its effectiveness in complex linguistic tasks.

[65] See Kalmenovitz and Chen (2024) for an application of supervised machine-learning algorithm to compute the regulatory similarity of each pair of S&P1500 firms.

each regulation-related task is labeled, we decompose a firm's (establishment's) RegIndex into $\text{RegIndex}_{i,t}^{req}$ and $\text{RegIndex}_{i,t}^{enf}$.

5.2.3 Results of Decomposing RegIndex

Decomposing RegIndex in Levels Figure 12 plots the two components of RegIndex, RegIndex^{req} and RegIndex^{enf} for different size bins. Panel A plots the results for firms, while Panel B for establishments. For each employment bin, we observe the average share of the total wage bill of a firm or of an establishment associated with performing compliance tasks related to requirements and of those related to enforcement.

The figure demonstrates how, for both firms and establishments, the inverted-U relationship between RegIndex and size originally observed in Figure 5 can be primarily traced to regulatory requirements, i.e. RegIndex^{req}. Firms in middle-sized bins around 500 employees have 0.57% higher RegIndex than the firms in the smallest bin (1.67%-1.1%). The differences in RegIndex^{req} between the two groups of firms is 0.55%, accounting for 97% (=0.55/0.57) of the difference in RegIndex. In contrast, the relation between RegIndex^{enf} and firm size appears flat. Indeed, regulatory requirements, RegIndex^{req}, account for almost the entire difference in RegIndex between mid-sized firms and the largest size group.

Overall, Figure 12 does not support that differential enforcement for large and small firms as a quantitatively important mechanism behind the inverted-U shape relationship between RegIndex and firm size. Heterogeneous regulatory requirements on small and large firms appear to drive essentially the entirety of the size-dependence of regulatory compliance costs incidence. This is consistent with the pervasive implementation of threshold-based regulations in practice, commonly known as regulatory tiering. In Appendix Table A.18, we present a varied sample of regulatory thresholds in firm size. As can be seen, size thresholds in employment, revenues or assets are common in regulations covering a broad spectrum of industries (e.g., Agriculture, Finance & Insurance, Manufacturing, Retail, Mining, Utilities, etc.) from both federal and states.

Sensitivity to Requirements and Enforcement Changes We can further corroborate the above decomposition of RegIndex in *levels* with an analysis of RegIndex's *changes*. Consider the regression of (log) changes in RegIndex on the regulatory requirements and on the enforcement shift-shares:

$$\Delta\log(\text{RegIndex}_{it}) \quad = \quad \delta_0 + \beta_1 iv(\Delta\log(\tilde{R}_{it})) + \gamma_1 iv(\Delta\log(p_{it})) + \delta_1 X_{it} + \xi_{it}, \tag{7}$$

where all variables are as in equation (6).

Panel A of Table 6 reports our estimation of equation (7) for the full sample of firms and establishments. For comparability, coefficients are reported for standardized variables.

We begin by investigating the roles of $iv(\Delta\log(p_{it}))$ and $iv(\Delta\log(\tilde{R}_{it}))$ separately. Columns (1) and (4) examine the response of firms' and establishments' share of regulatory costs ($\Delta\log(\text{RegIndex}_{it})$) to enforcement changes, $iv(\Delta\log(p_{it}))$. In both columns, we observe positive and statistically significant coefficients, confirming the intuitive conditional correlation between increases in agency regulatory hires translating into more enforcement and consequently higher regulatory compliance expenditure. The estimated coefficients further indicate a quantitatively meaningful relationship, as a one standard deviation increase in $iv(\Delta\log(p_{it}))$ produces a 0.21 of a standard deviation increase in the change in regulatory compliance costs index for a firm (0.23 for an establishment). Columns (2) and (5) examine the response of firms' and establishments' regulatory compliance costs to regulation requirements changes, $iv(\Delta\log(\tilde{R}_{it}))$. The coefficient for $iv(\Delta\log(\tilde{R}_{it}))$ is larger than for the case of $iv(\Delta\log(p_{it}))$. A one standard deviation increase in $iv(\Delta\log(\tilde{R}_{it}))$ produces a 0.26 of a standard deviation increase in the change in regulatory compliance costs index for a firm (0.27 for an establishment). These results suggest that our shift-shares meaningfully associate with ex-post changes in firms' regulatory compliance costs.

Columns (3) and (6) include both requirements and enforcement variables in the estimation of equation (7), allowing to compare the marginal effect of changes in each component. This horse-race specification reveals, confirming our results above, a stronger role for regulatory requirements than enforcement in driving changes in compliance. The regulatory requirements coefficient is estimated at 0.23 and statistically significant, while the enforcement effort slope is 0.11 and statistically insignificant for firms (similarly for establishments).[66]

To see how the shift-share slopes may be informative about the mechanisms behind the relation between the level of RegIndex and business size, divide firms in four employment bins: between 1 and 19 employees, between 20 and 399, between 400 and 749, and above 750 employees. Panel B of Table 6 confirms that the effects of changes in regulatory requirements, $iv(\Delta\log(\tilde{R}_{it}))$, are substantially higher for medium and large firms relative to small ones, particularly for those entities employing ≤ 10 workers. The null hypotheses of equality of the coefficients on the variable $iv(\Delta\log(\tilde{R}_{it}))$ for firms below 10 employees and for those between 400-749 or those above 749 employees are rejected at standard levels of statistical significance. This is consistent with an intuitive mechanism. Under regulatory tiering, which implies

[66] Appendix Table A.12 and displays variations of the above results across broad sectors of firms and establishments.

less stringent regulatory requirements for small business, such as for example in the cases of local banks or labor laws applying to firms below certain employment thresholds, changes in requirements affecting medium and large firms have only indirect or general equilibrium effects, as small businesses are exempt. This in turn explains the smaller coefficient on $iv(\Delta\log(\tilde{R}_{it}))$. Notice that, as regulatory tiering is a feature of the design of regulatory requirements, it is important to hold enforcement of regulation constant, to cleanly isolate this mechanism in the tests of Table 6.[67]

In summary, the evidence from decomposing the level of firm RegIndex and inspecting the sensitivity of RegIndex changes to requirement and enforcement changes suggests that regulatory requirements and regulatory tiering seem important drivers of the inverted-U relationship between firms' RegIndex and and their employment size.

5.3 Potential Implications for Firm Size Distribution

Potential distortions to the establishment and firm size distributions ensue from the presence of the size-dependent regulation that we just documented.[68] To precisely quantify aggregate macroeconomic implications, however, one would need to decouple the role of regulatory requirements and endogenous enforcement from concomitant market forces (e.g., technology or international trade) and unobserved sector-specific shocks. In this study, we do not possess an appropriate identification strategy to pursue this rigorously and any evidence in terms of macroeconomic implications should be taken as exploratory, rather than conclusive. Nonetheless, one can outline some implications useful to motivate future research.

Our simple theoretical framework in Section 3.2 suggests that firms choose their optimal size in an environment where regulatory compliance costs are size-dependent, i.e., $L^*_{it} = \left(\frac{\phi_i - w_{it}}{(\alpha+\beta)\left(w^r_{it}\pi\rho\right)^{\frac{1}{2}}}\right)^{\frac{2}{\alpha+\beta-2}}$ in equation (1). This, in turn, implies that policy parameters $(\alpha, \beta, \rho, \pi)$ directly affect the optimal firm size distribution. We thus can offer a discussion of the association between the size-dependent regulatory compliance costs and the dynamics of firm/establishment size distribution in the U.S. from 2002 to 2014.

We begin by performing the simple exercise of examining the conditional correlation between changes in firm size distribution and RegIndex within the sector. To account for changes in sectoral composition in the economy, we estimate the following regression:

$$\Delta FirmShare_{l,j} = \varphi RegIndex_{l,j} + \delta_j + \epsilon_{l,j},$$

[67]This is why controlling for $iv(\Delta\log(p_{it}))$, which may also be size-dependent in principle, is a crucial step in isolating the role of regulatory tiering. For recent research using regulatory tiering for statistical identification of regulatory distortions, see Garicano et al. (2016); Aghion et al. (2021); Alvero et al. (2023); Ewens et al. (2024).

[68]For the case of French labor regulation, see Aghion et al. (2021).

where $\Delta FirmShare_{l,j}$ measures the change in the mass of firms from each of 14 different employment bins (l) within each of 8 different sectors (j) between the years 2002 and 2014, $RegIndex_{l,j}$ is the average RegIndex for the employment bin within the sector (see Figure 6), and δ_j represents sector fixed effects. Intuitively, in this regression a negative φ estimate indicates a "hollowing out" of an employment bin within the sector whenever the level of RegIndex in that employment-sector bin is higher.[69]

Indeed, Panel A of Table A.19 reports a negative and statistically significant estimate of $\varphi = -0.017$ ($t = -3.16$) for changes in firm size distribution (see Column (1)). The estimates are similar for changes in establishments' size distribution $\varphi = -0.016$ ($t = -3.67$) as shown in Panel B. The economic magnitude is sizable, though we emphasize that these results should not be interpreted causally: size bins with a one-standard-deviation higher RegIndex within a sector are associated with a 0.7-standard-deviation larger loss in firm or establishment share. Columns (2) and (3) further split the sample into the smaller and larger halves of the firm size distribution, revealing that the effects are concentrated among small and mid-sized firms rather than large firms. This finding aligns with the view that threshold-based regulations tend to benefit small firms while hollowing out mid-sized firms, since most U.S. regulatory thresholds fall within this range (see examples in Table A.18); correspondingly, RegIndex is consistently lower for small firms than for mid-sized firms across nearly all sectors (Figure 6).

To further confirm the negative association between regulatory compliance burden and firm size distribution, in Columns (4)-(6) of Table A.19, we run another regression specification based on *changes* in the employment-sector bin's RegIndex from 2002 to 2014:

$$\Delta FirmShare_{l,j} = \varphi \Delta RegIndex_{l,j} + \delta_j + \epsilon_{l,j}.$$

In this regression, φ estimates the conditional correlation between growth in the mass of firms in an employment bin and the increase in compliance costs in that bin. Consistent with our finding above, we observe a significantly negative estimate of φ in Column (4). The economic magnitude, measured in one-standard-deviation terms, is less than half that of the RegIndex level in Column (1). In Columns (5)

[69]To measure $\Delta FirmShare_{l,j}$, we first compute the share of establishments in each of the 14 size bins (Section 4) within each NAICS 1-digit sector in each year. To ensure the shares are representative of the size distribution in the economy, we weigh each establishment by their sampling weights assigned by the BLS—BLS assigns sampling weights to establishments by size, location, industry, and ownership so as to produce the aggregate statistics at the industry, geographic, and national level for public use at https://www.bls.gov/oes/tables.htm. Next, we compute the changes in the share of establishments in each employment-sector bin from 2002 to 2014. We also compute the changes in firms' size distribution using the same methodology. For single-establishment firms, their sampling weights are the same as their establishments'. Assigning sampling weights to multi-establishment firms is challenging. We use the lowest sampling weights among establishments within the firm, so as to mitigate the impact of large sampling weights due to low coverage of businesses in that establishment's sampling bucket. Nevertheless, the results are very similar if we use the average weight of establishments within the firm as the firm's weight.

and (6), we further separate the sample into the smaller and larger halves of the firm size distribution. Again, we observe that the results are driven by changes in regulation affecting small and mid-sized firms. In other words, this evidence suggests that small firms' growth is also more sensitive to changes in regulatory costs—another relevant dimension when assessing the implications of rising compliance costs for firm dynamism. Our RegIndex, which accounts for all firms—including the very smallest—is thus particularly valuable for future research on this topic.

There is a direct implication of these correlation estimates when combined with the lower regulatory burden falling on the very small firms and the higher burden on mid-size firms at the aggregate level shown in Figure 6: A loss of mass in the middle of the firm (and establishment) size distribution–a hollowing out in the middle of the size distribution over time. We observe some suggestive evidence of this phenomenon in Figure A.15. The figure plots the changes in firms' and establishments' size distribution from 2002 to 2014 in the U.S. economy from our microdata. We observe that the fraction of small and mid-sized businesses with employees between 10 and 250 declines relative to the very small or very large firms, consistent with the intuition that mid-sized firms face higher regulatory compliance costs. The biggest expansion over 2002-2014 occurs for the very small firms, where RegIndex is the lowest. While this pattern should not be uniquely ascribed to regulatory requirements, as many omitted confounding factors may affect this negative correlation, it provides suggestive evidence of the potentially broader implications of our work.

6 Conclusion

This paper presents a new approach to estimating business compliance costs of regulation in the United States. Based on micro establishment-occupation level data, we quantify the total labor costs paid by businesses to employees engaging in safety, compliance, monitoring, and other regulation-related tasks in order to meet federal, state, and local regulatory requirements. The average U.S. firm spends between 1.3 and 3.3 percent of its total wage bill on regulatory compliance. This wage bill grew at an annual rate of about 1 percent from 2002 to 2014, roughly half of the average annual GDP growth rate over the period. The total wage bill devoted to regulatory compliance workers in 2014 was between $79 billion and $239 billion, depending on the stringency of the regulatory compliance measure employed, and up to $289 billion when capital is also added, comparable to U.S. gross business income taxes.

We show that the percentage of firms' labor costs paid for regulatory compliance first increases with firm size, measured by total employment, and then decreases, exhibiting an inverted-U shape. This inverted-U relationship is a robust pattern as we show across multiple specifications—when accounting for within-occupation variation in regulatory tasks, incorporating capital expenditures, adjusting for outsourcing of compliance functions, and controlling for multi-state regulatory exposure. These consistent findings strengthen our confidence in the reliability of the inverted-U pattern as a structural feature of regulatory burden distribution.

This inverted-U shape between RegIndex and firm size suggests that for small businesses regulation is tiered and tends to be lighter, while red tape increases as employment reaches 500 workers. Beyond this threshold, regulatory costs tend to decrease, with evidence of economies of scale kicking in for regulatory compliance. Identifying the presence of increasing returns to scale in regulatory compliance for medium-large firms is an important step in the direction of assessing equilibrium firm size distortions due to the design of the U.S. administrative system. We further design and implement a shift-share method to identify firm responses to regulation requirements versus enforcement. Using this design, we argue that changes in regulatory requirements appear to contribute significantly to the enhanced inverted-U relationship between firms' regulatory compliance cost and size in our sample period, while the contribution from changes in enforcement is limited.

Future research should extend the use of our methodology to other high-income countries, where similar micro data is available, for a comparative perspective on the costs of regulation and assessing external validity. Quantifying the productivity losses to U.S. firms due to regulation also appears to be an important direction of future inquiry.

References

Acemoglu, D., Autor, D., 2011. Skills, tasks and technologies: Implications for employment and earnings, in: Handbook of labor economics. Elsevier. volume 4, pp. 1043–1171.

Acemoglu, D., Autor, D., Hazell, J., Restrepo, P., 2020. AI and jobs: Evidence from online vacancies. Technical Report. National Bureau of Economic Research.

Agarwal, S., Lucca, D., Seru, A., Trebbi, F., 2014. Inconsistent regulators: Evidence from banking. The Quarterly Journal of Economics 129, 889–938.

Aghion, P., Bergeaud, A., Van Reenen, J., 2021. The impact of regulation on innovation. Technical Report. National Bureau of Economic Research.

Akcigit, U., Ates, S.T., 2020. Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory. American Economic Journal: Macroeconomics doi:10.1257/mac.20180449.

Al-Ubaydli, O., McLaughlin, P.A., 2017. Regdata: A numerical database on industry-specific regulations for all united states industries and federal regulations, 1997–2012. Regulation & Governance 11, 109–123.

Alesina, A., Battisti, M., Zeira, J., 2018. Technology and labor regulations: theory and evidence. Journal of Economic Growth 23, 41–78.

Alvero, A., Ando, S., Xiao, K., 2023. Watch what they do, not what they say: Estimating regulatory costs from revealed preferences. The Review of Financial Studies 36, 2224–2273.

Armstrong, D.M., Glaeser, S., Hoopes, J.L., 2025. Measuring firm exposure to government agencies. Journal of Accounting and Economics 79, 101703. doi:10.1016/j.jacceco.2024.101703.

Ayyagari, M., Maksimovic, V., 2017. Fewer and less skilled? human capital, competition, and entrepreneurial success in manufacturing. Human Capital, Competition, and Entrepreneurial Success in Manufacturing (December 26, 2017) .

Baker, S.R., Bloom, N., Canes-Wrone, B., Davis, S.J., Rodden, J., 2014. Why has us policy uncertainty risen since 1960? American Economic Review P&P 104, 56–60.

Baker, S.R., Bloom, N., Davis, S.J., 2016. Measuring economic policy uncertainty. The quarterly journal of economics 131, 1593–1636.

Bernstein, S., 2015. Does Going Public Affect Innovation? The Journal of Finance 70, 1365–1403. doi:10.1111/jofi.12275.

Bertrand, M., Bombardini, M., Trebbi, F., 2014. Is it whom you know or what you know? an empirical assessment of the lobbying process. The American Economic Review 104, 3885–3920.

Boeri, T., Jimeno, J.F., 2005. The effects of employment protection: Learning from variable enforcement. European Economic Review 49, 2057–2077.

Bombardini, M., Trebbi, F., 2020. Empirical models of lobbying. Annual Review of Economics , 391–413.

Bombardini, M., Trebbi, F., 2025. The Political Power of Firms. UC Berkeley Mimeo .

Bombardini, M., Trebbi, F., Zhang, M.B., 2025. Measuring the Costs and Benefits of Regulation. Annual Review of Economics 17.

Brock, W.A., Evans, D.S., 1985. The economics of regulatory tiering. The Rand Journal of Economics , 398–409.

Brock, W.A., Evans, D.S., Phillips, B.D., 1986. The economics of small businesses: Their role and regulation in the US economy. Holmes & Meier New York.

Callander, S., Foarta, O., Sugaya, T., 2021. Market Competition and Political Influence: An Integrated Approach. Technical Report. Centre for Economic Policy Research.

Calomiris, C.W., Mamaysky, H., Yang, R., 2020. Measuring the cost of regulation: A text-based approach. Technical Report. National Bureau of Economic Research.

Caunedo, J., Jaume, D., Keller, E., 2023. Occupational exposure to capital-embodied technical change. American Economic Review .

Charoenwong, B., Kowaleski, Z.T., Kwan, A., Sutherland, A.G., 2024. RegTech: Technology-driven compliance and its effects on profitability, operations, and market structure. Journal of Financial Economics 154, 103792. doi:10.1016/j.jfineco.2024.103792.

Chayes, A., Chayes, A.H., 1985. Corporate Counsel and the Elite Law Firm. Stanford Law Review 37, 277–300. doi:10.2307/1228616, arXiv:1228616.

Chodorow-Reich, G., 2014. The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis. The Quarterly Journal of Economics 129, 1–59.

Coates, J.C., Srinivasan, S., 2014. SOX after Ten Years: A Multidisciplinary Review. Accounting Horizons 28, 627–671. doi:10.2308/acch-50759.

Coates IV, J.C., 2007. The goals and promise of the Sarbanes–Oxley Act. Journal of economic perspectives 21, 91–116.

Cowgill, B., Prat, A., Valletti, T., 2021. Political power and market power. arXiv preprint arXiv:2106.13612 .

Davis, S.J., 2017. Regulatory complexity and policy uncertainty: headwinds of our own making. Becker Friedman Institute for Research in economics working paper .

De Fontenay, E., 2016. The deregulation of private capital and the decline of the public company. Hastings LJ 68, 445.

Djankov, S., La Porta, R., Lopez-de Silanes, F., Shleifer, A., 2002. The regulation of entry. The quarterly Journal of economics 117, 1–37.

Evans, D.S., 1986. The differential effect of regulation across plant size: Comment on pashigian. The Journal of Law and Economics 29, 187–200.

Ewens, M., Xiao, K., Xu, T., 2024. Regulatory costs of being public: Evidence from bunching estimation. Journal of Financial Economics 153, 103775.

Gao, X., Ritter, J.R., Zhu, Z., 2014. Where Have All the IPOs Gone? Journal of Financial and Quantitative Analysis 48, 1663–1692. doi:10.1017/S0022109014000015.

García-Santana, M., Pijoan-Mas, J., 2011. Small scale reservation laws and the misallocation of talent .

Garicano, L., Lelarge, C., Van Reenen, J., 2016. Firm size distortions and the productivity distribution: Evidence from france. American Economic Review 106, 3439–79.

Goff, B.L., et al., 1996. Regulation and macroeconomic performance. volume 21. Springer Science & Business Media.

Goldschmidt, D., Schmieder, J.F., 2017. The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure. The Quarterly Journal of Economics 132, 1165–1217. doi:10.1093/qje/qjx008.

Gourio, F., Roys, N., 2014. Size-dependent regulations, firm size distribution, and reallocation. Quantitative Economics 5, 377–416.

Guner, N., Ventura, G., Xu, Y., 2008. Macroeconomic implications of size-dependent policies. Review of Economic Dynamics 11, 721–744.

Gutiérrez, G., Philippon, T., 2017. Declining Competition and Investment in the U.S. Working Paper 23583. doi:10.3386/w23583.

Gutiérrez, G., Philippon, T., 2019. The Failure of Free Entry. Technical Report. National Bureau of Economic Research.

Hassan, T.A., Hollander, S., Van Lent, L., Tahoun, A., 2019. Firm-level political risk: Measurement and effects. The Quarterly Journal of Economics 134, 2135–2202.

Hazilla, M., Kopp, R.J., 1990. Social cost of environmental quality regulations: A general equilibrium analysis. Journal of Political Economy 98, 853–873.

Helland, E., 1998. The enforcement of pollution control laws: Inspections, violations, and self-reporting. Review of Economics and Statistics 80, 141–153.

Hopkins, T.D., 1995. Profiles of regulatory costs. Report to the US Small Business Administration, Office of Advocacy .

Hovenkamp, H., Morton, F.S., 2019. Framing the chicago school of antitrust analysis. U. Pa. L. Rev. 168, 1843.

Hsieh, C.T., Klenow, P.J., 2009. Misallocation and manufacturing tfp in china and india. The Quarterly journal of economics 124, 1403–1448.

Huneeus, F., Kim, I.S., 2018. The effects of firms' lobbying on resource misallocation .

Jackson, H.E., 2007. Variation in the intensity of financial regulation: Preliminary evidence and potential implications. Yale J. on Reg. 24, 253.

Jackson, H.E., Roe, M.J., 2009. Public and private enforcement of securities laws: Resource-based evidence. Journal of Financial Economics 93, 207–238. doi:10.1016/j.jfineco.2008.08.006.

Julio, B., Yook, Y., 2012. Political uncertainty and corporate investment cycles. The Journal of Finance 67, 45–83.

Kalmenovitz, J., 2023. Regulatory intensity and firm-specific exposure. The review of financial studies 36, 3311–3347.

Kalmenovitz, J., Chen, J., 2024. Regulatory Similarity. Journal of Law and Economics, forthcoming .

Kalmenovitz, J., Lowry, M., Volkova, E., 2025. Regulatory Fragmentation. The Journal of Finance 80, 1081–1126. doi:10.1111/jofi.13423.

Kang, K., 2016. Policy influence and private returns from lobbying in the energy sector. Review of Economic Studies 83, 269–305.

Kang, K., Silveira, B.S., 2021. Understanding disparities in punishment: Regulator preferences and expertise. Journal of Political Economy 129, 2947–2992.

Klapper, L., Laeven, L., Rajan, R., 2006. Entry regulation as a barrier to entrepreneurship. Journal of financial economics 82, 591–629.

Lancieri, F., Posner, E.A., Zingales, L., 2022. The political economy of the decline in antitrust enforcement in the united states. Available at SSRN .

Levinson, A., 1996. Environmental regulations and manufacturers' location choices: Evidence from the Census of Manufactures. Journal of public Economics 62, 5–29.

List, J.A., Millimet, D.L., Fredriksson, P.G., McHone, W.W., 2003. Effects of environmental regulations on manufacturing plant births: evidence from a propensity score matching estimator. Review of Economics and Statistics 85, 944–952.

Mian, A., Sufi, A., Trebbi, F., 2010. The political economy of the us mortgage default crisis. American Economic Review 100, 1967–1998.

Morrall, J.F., 1986. A review of the record. Regulation 10, 25.

Muehlenbachs, L., Staubli, S., Cohen, M.A., 2016. The Impact of Team Inspections on Enforcement and Deterrence. Journal of the Association of Environmental and Resource Economists 3, 159–204. doi:10.1086/684035.

Parente, S.L., Prescott, E.C., 2002. Barriers to riches. MIT press.

Parker, C.E., Rosen, R.E., Nielsen, V.L., 2009. The two faces of lawyers: Professional ethics and business compliance with regulation. Geo. J. Legal Ethics 22, 201.

Peltzman, S., 1976. Toward a more general theory of regulation. The Journal of Law and Economics 19, 211–240.

Peltzman, S., 1998. Political participation and government regulation. University of Chicago Press.

Philippon, T., 2019. The great reversal: How America gave up on free markets. Harvard University Press.

Restuccia, D., Rogerson, R., 2008. Policy distortions and aggregate productivity with heterogeneous establishments. Review of Economic dynamics 11, 707–720.

Ryan, S.P., 2012. The costs of environmental regulation in a concentrated industry. Econometrica 80, 1019–1061.

Schaefer, J., Zimmer, M., 2003. Professional regulation and labor market outcomes for accountants: Evidence from the current population survey, 1984–2000. Research in Accounting Regulation 16, 87–104.

Schivardi, F., Torrini, R., 2008. Identifying the effects of firing restrictions through size-contingent differences in regulation. Labour Economics 15, 482–511.

Shimshack, J.P., 2014. The economics of environmental monitoring and enforcement. Annu. Rev. Resour. Econ. 6, 339–360.

Simkovic, M., 2009. The effect of BAPCPA on credit card industry profits and prices. Am. Bankr. LJ 83, 1.

Singla, S., 2023. Regulatory costs and market power. mimeo LBS .

Song, J., Price, D.J., Guvenen, F., Bloom, N., Von Wachter, T., 2018. Firming up inequality. The Quarterly Journal of Economics 134, 1–50.

Stigler, G.J., 1971. The theory of economic regulation. The Bell journal of economics and management science , 3–21.

Stiglitz, J., 2009. Regulation and Failure. New Perspectives on Regulation , 13–25.

National Association of Manufacturers, 2014. The cost of federal regulation to the u.s. economy, manufacturing and small business .

Securities Industry Association, 2006. The costs of compliance in the u.s. securities industry: Survey report .

Thomas, L.G., 1990. Regulation and firm size: Fda impacts on innovation. The RAND Journal of Economics , 497–517.

Blanes i Vidal, J., Draca, M., Fons-Rosen, C., 2012. Revolving door lobbyists. The American Economic Review 102, 3731–3748.

Zhang, M.B., 2019. Labor-technology substitution: Implications for asset pricing. Journal of Finance 74, 1793–1839.

Figure 1: **Aggregate Series of Regulation Index**

This figure plots the aggregate Regulation Index from 2002 to 2014. RegIndex is the percentage of an establishment's annual labor spending on performing regulation-related tasks (see Section 3). The RegIndex in each year is aggregated from our final sample of establishments, weighted by establishment weights designated by the BLS. Our final sample focuses on non-government/non-education establishments (Song et al., 2018) and exclude establishments from industries that provide legal or compliance services: legal and accounting firms, government administration, courts, central banking.

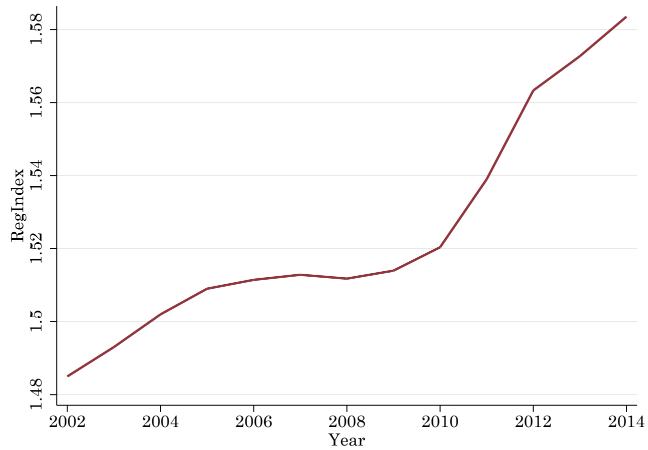


Figure 2: **Validation of Regulation Index Using Agency-Estimated Compliance Hours**

This figure plots the aggregate annual compliance hours (in billions of hours) identified by our most conservative RegIndex measure which is our baseline measure used in this study. See Section 3 for details about the three versions of RegIndex. We compare the aggregate compliance hours identified by our measure with the estimated annual compliance hours (in billions of hours) submitted by various regulatory agencies to the White House Office of Information and Regulatory Affairs (OIRA). Appendix Figure A.1 plots additional graphs using the hours identified by two broader versions of the RegIndex measure: Medium RegIndex and Broad RegIndex.

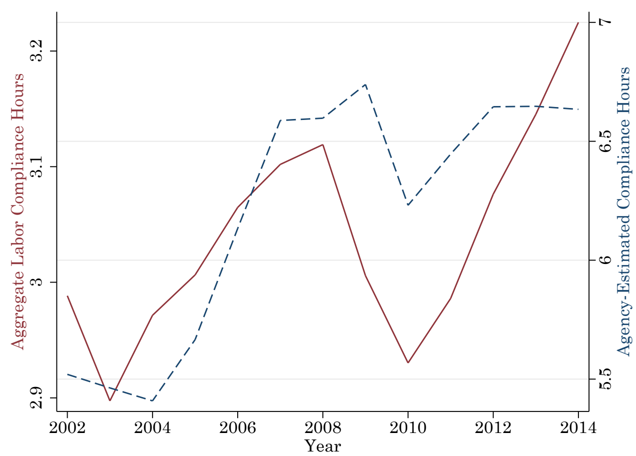


Figure 3: **Validation of Regulation Index Using Industry Regulatory Policy Changes**

This figure plots the response of industries' Regulation Index (RegIndex) to five industry-level regulatory policy changes. RegIndex is the percentage of an industry's annual labor spending on performing regulation-related tasks. Section 3 provides details of the industry regulatory policy changes and discusses the classification of treated and control groups. To ease the comparison, we shift the lines vertically so that they have the same value in the year before the treatment. The value in the year before the treatment is the average of the regulation measures across the treated and control industries in that year. The difference between the two lines after the treatment, minus the difference between the two lines before the treatment reflects the difference-in-differences estimation. The shaded areas indicate the 95% confidence interval of the industries' average RegIndex.

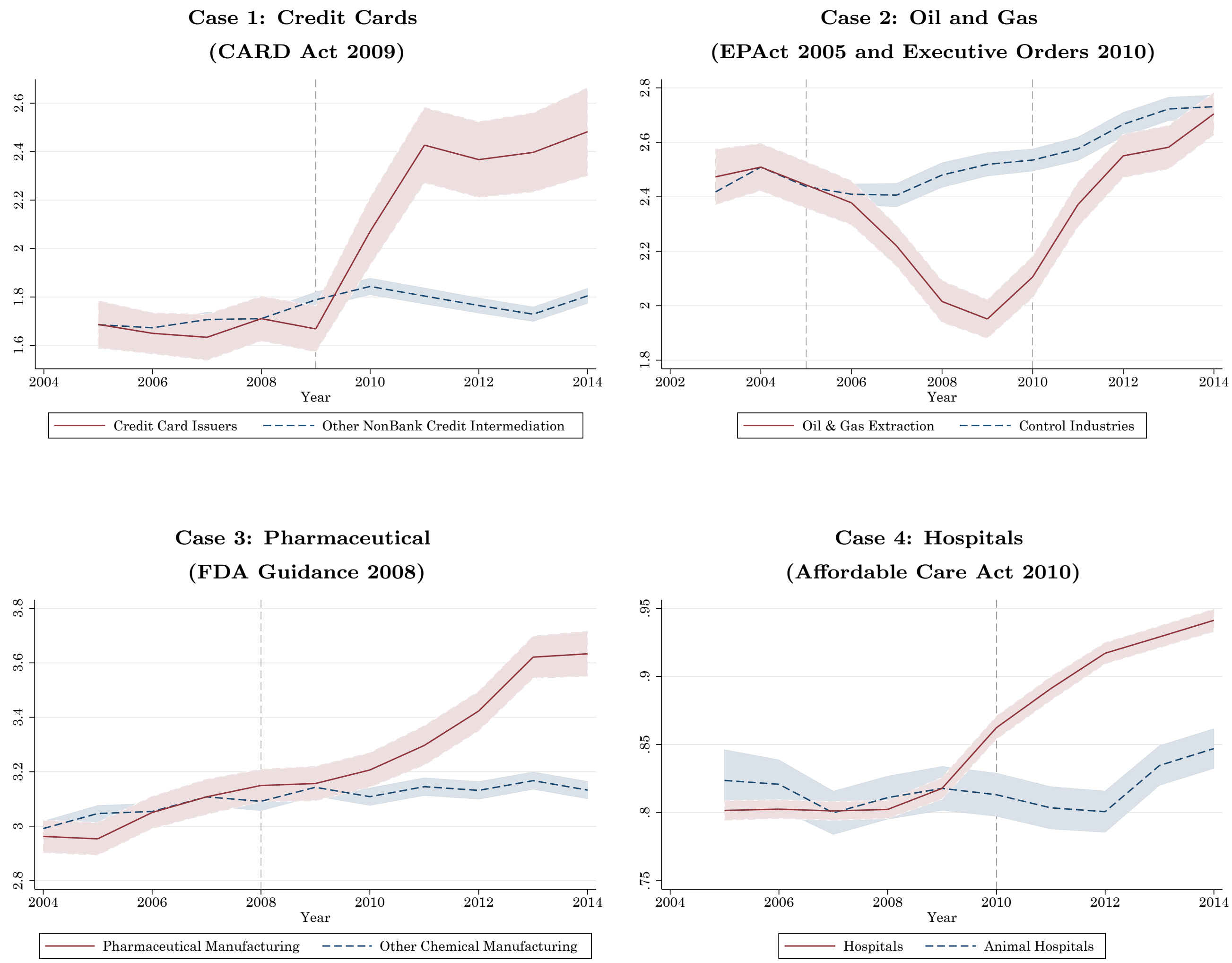


Figure 4: **Regulation Index Across States**

This figure plots the coefficients on state dummies in the following regressions based on about 1 million private establishments in the 2014 OEWS universe: $RegIndex_{i,t} = \alpha + \sum_{s \in States} \beta_s \times State_s + FE_{Ind} + \epsilon_{i,t}$ The coefficient α shows the RegIndex for the benchmark state "Alabama." The sum of coefficients $\alpha + \beta_s$ shows the RegIndex for the other 50 states (including the District of Columbia). All establishments are weighted by their total wage payment and their sampling weights assigned by the OEWS survey. Industry is at the NAICS 6-digit level.

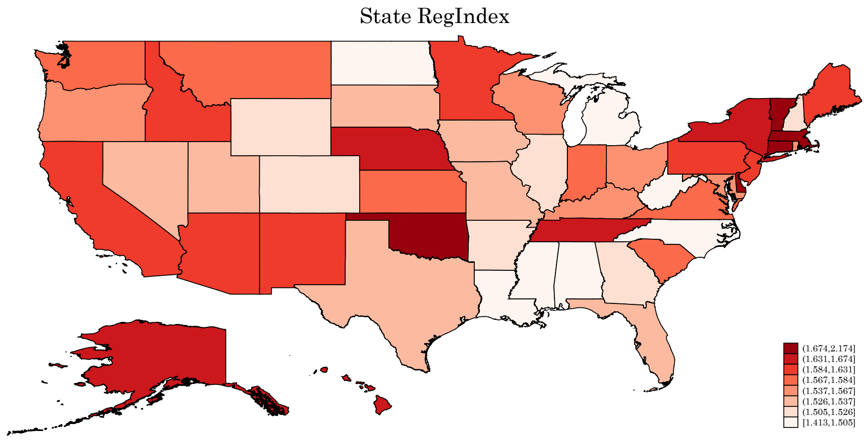
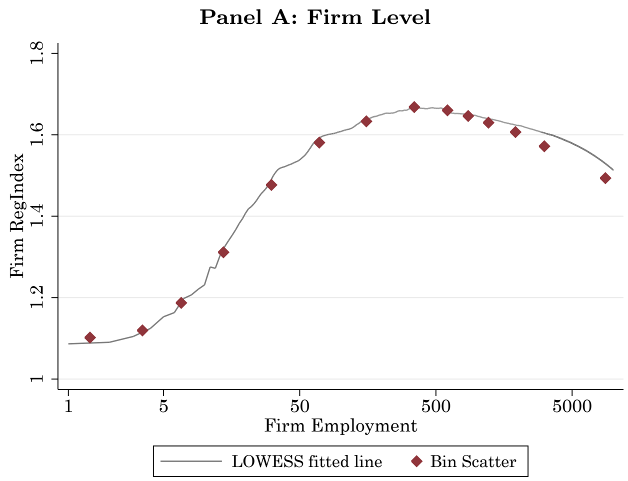


Figure 5: **RegIndex and Size**

This figure plots the relation of RegIndex and employment for firms in Panel A and establishments in Panel B. The dots represent the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The line represents the LOWESS smoothed fitted curve using the bandwidth of 0.05. The x-axis is on a log scale.



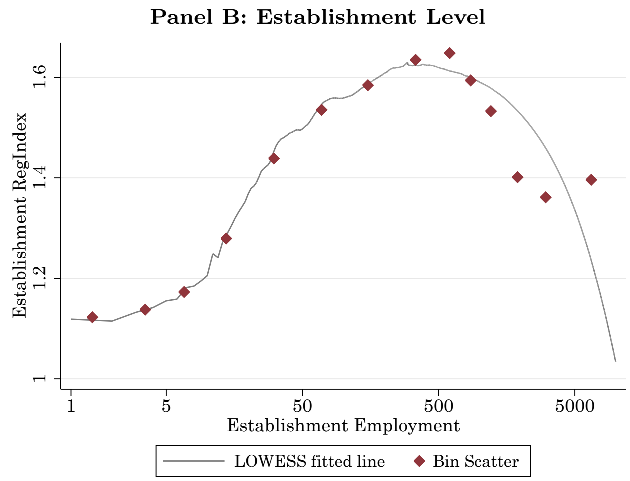


Figure 6: **RegIndex and Size by Sector**

Panel A shows the relation of RegIndex and firm employment in each NAICS 1-digit sector. Panel B shows the relation at the establishment level. Each dot represents the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

Panel A: Firm-Level

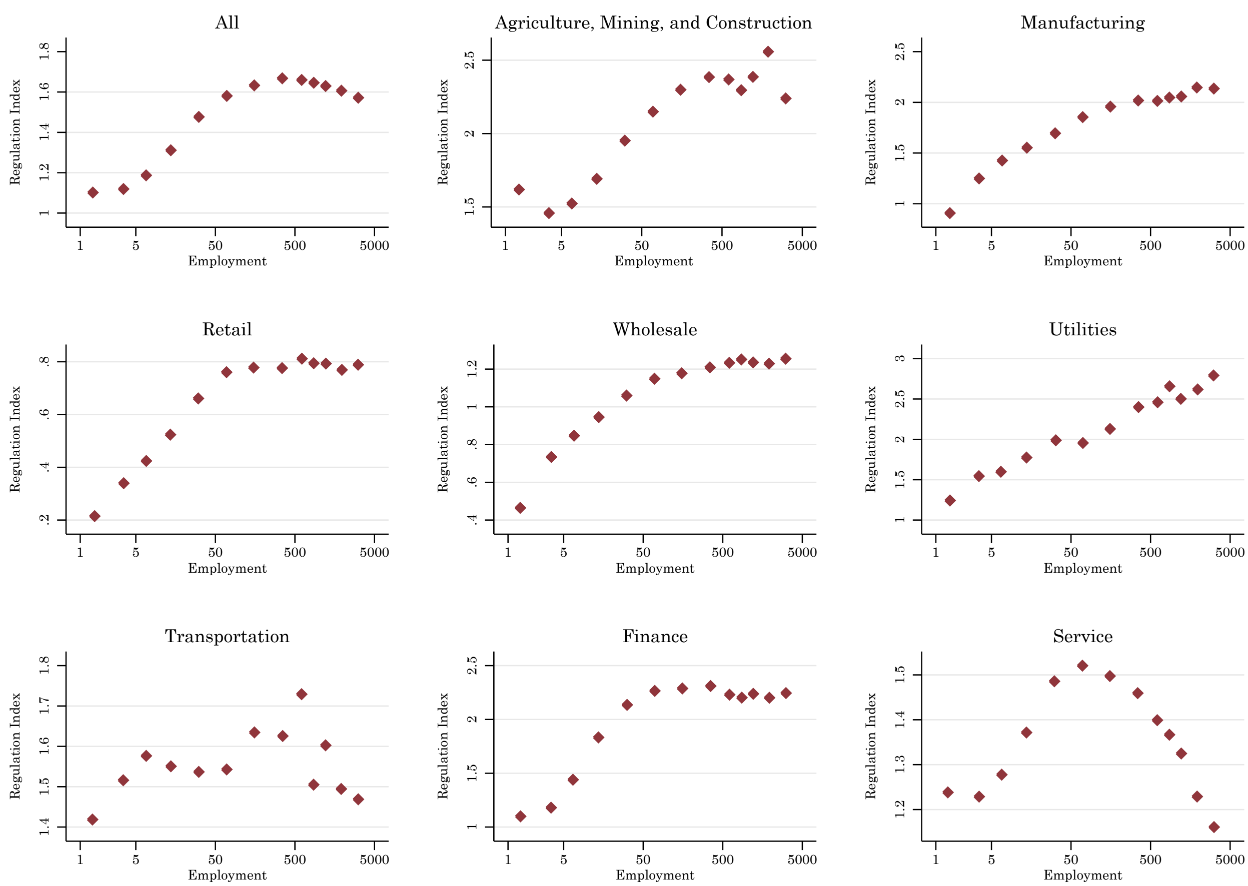


Figure 6: **RegIndex and Size by Sector**—*Continued*

Panel A shows the relation of RegIndex and firm employment in each NAICS 1-digit sector. Panel B shows the relation at the establishment level. Each dot represents the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

Panel B: Establishment-Level

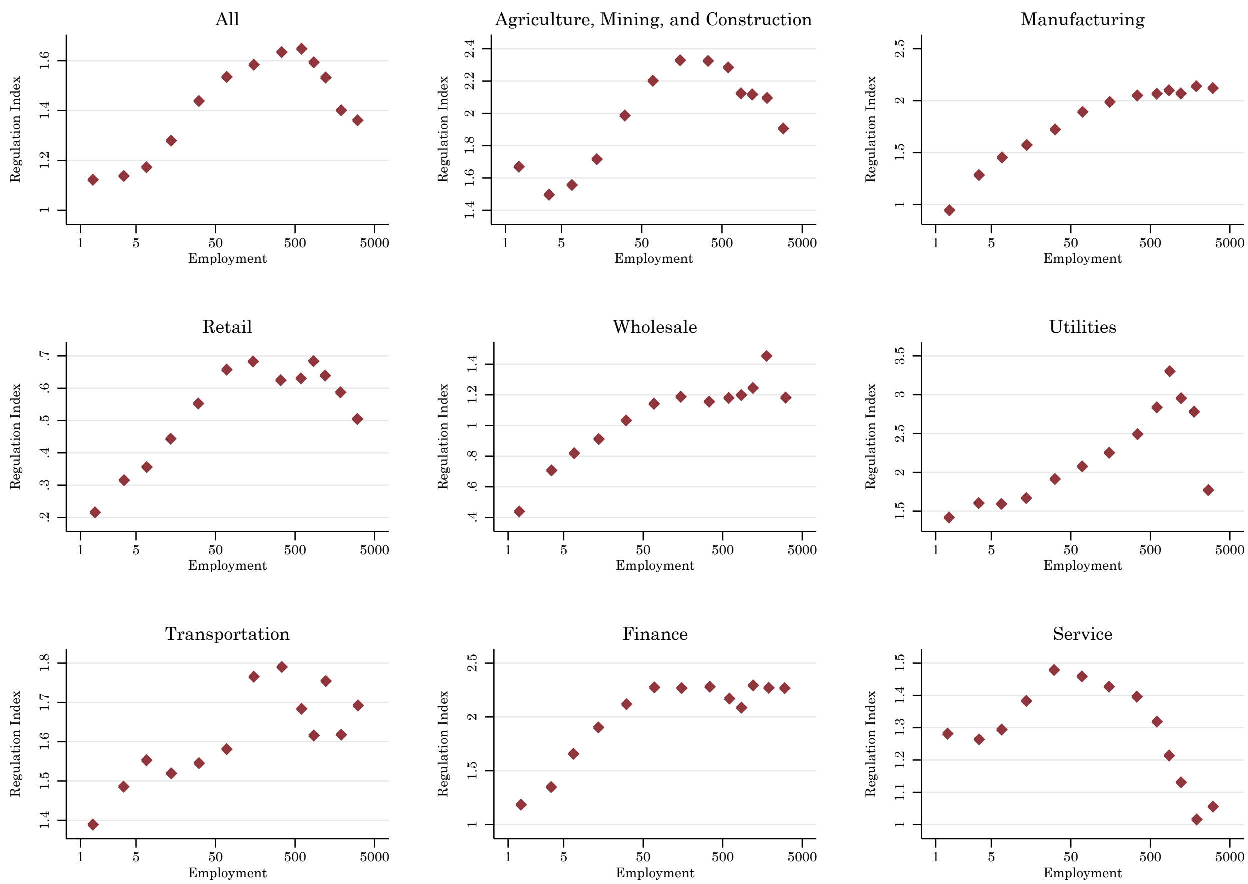


Figure 7: **Within-Occupation Regulatory Compliance Requirement and Firm Size**

This figure plots the demeaned regulatory compliance skill requirement measure for job postings in each firm employment bin, where a job posting's regulatory compliance skill requirement measure is demeaned by subtracting its mean within an SOC 6-digit occupation and year. See Section 4.1.1 for details. The bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

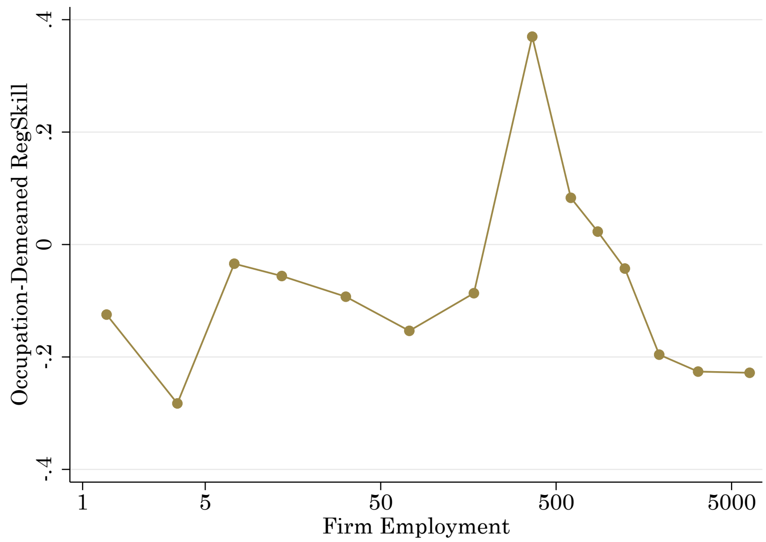


Figure 8: **Total-Input-Based RegIndex and Size**

This figure plots the relation of labor-based RegIndex (baseline), total-input-based RegIndex (including capital) and employment for firms in Panel A and establishments in Panel B. See Section 3.3 for definitions of RegIndex and RegIndex$^{\text{Tot}}$. The dots represent the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

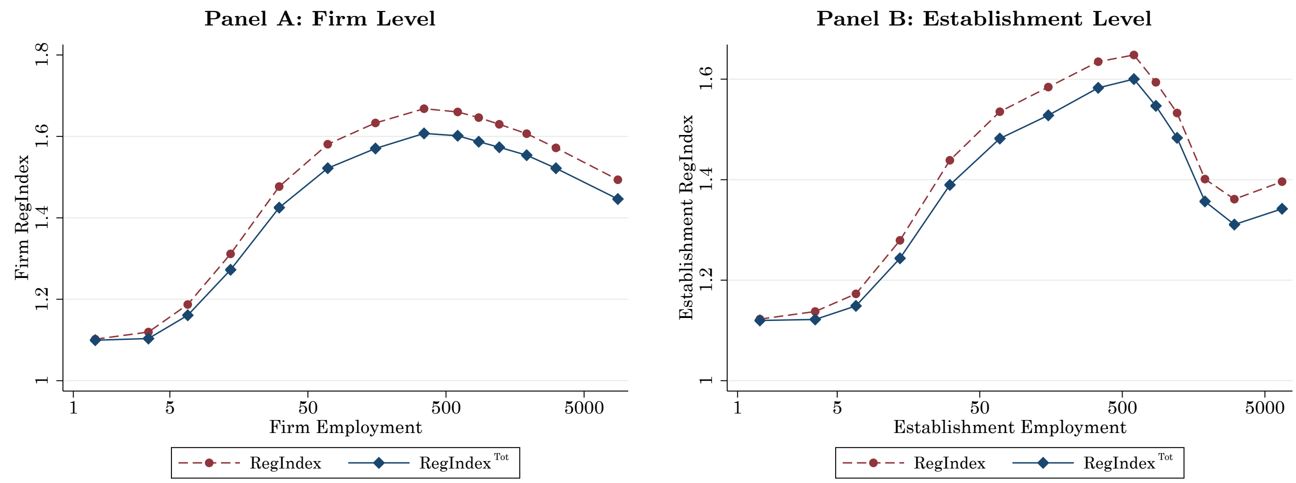


Figure 9: **Outsource-Adjusted RegIndex and Size**

This figure plots the relation of in-house-based RegIndex (baseline), two versions of outsource-adjusted RegIndex and employment for firms in Panels A and C and establishments in Panels B and D. In Panels A and B, we include compliance costs paid to all compliance outsourcing firms, whereas in Panels C and D, we include compliance costs paid to small outsourcing firms with 20 or fewer employees. Each panel reports two versions of outsource-adjusted RegIndex constructed assuming that firms use compliance outsourcing services from local providers within the county and from nationwide providers, respectively. Section 4.1.3 provides the detailed definitions and estimation procedures. See Section 3.3 for the definition of the baseline RegIndex. The dots represent the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

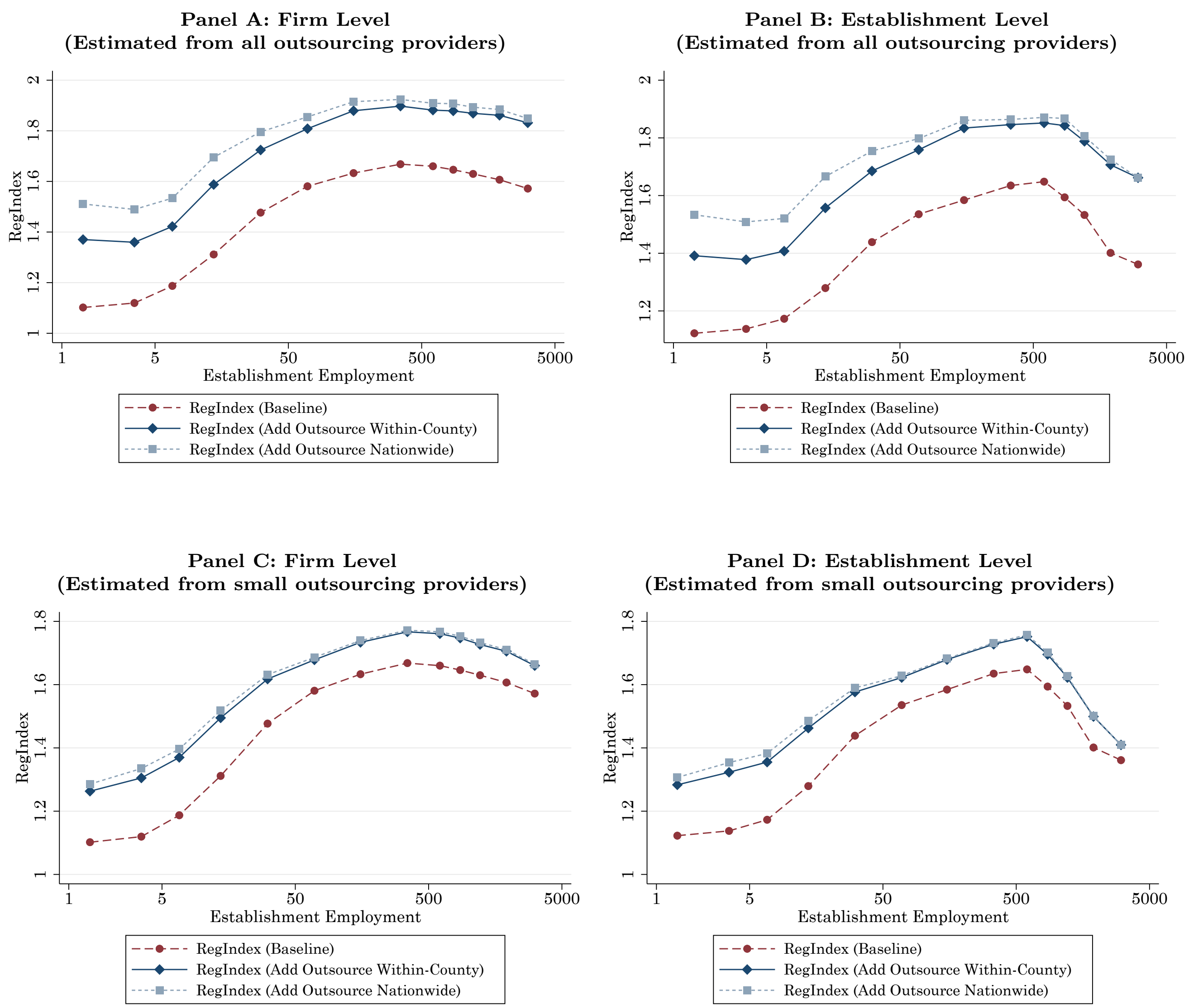


Figure 10: **RegIndex and Size: Multi-State vs. Single State Firms**

This figure plots the relation of RegIndex and employment for firms in Panel A and establishments in Panel B. In each panel, we plot RegIndex for multi-state firms, single-unit firms, and other firms. See more details in Section 4. The dots represent the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

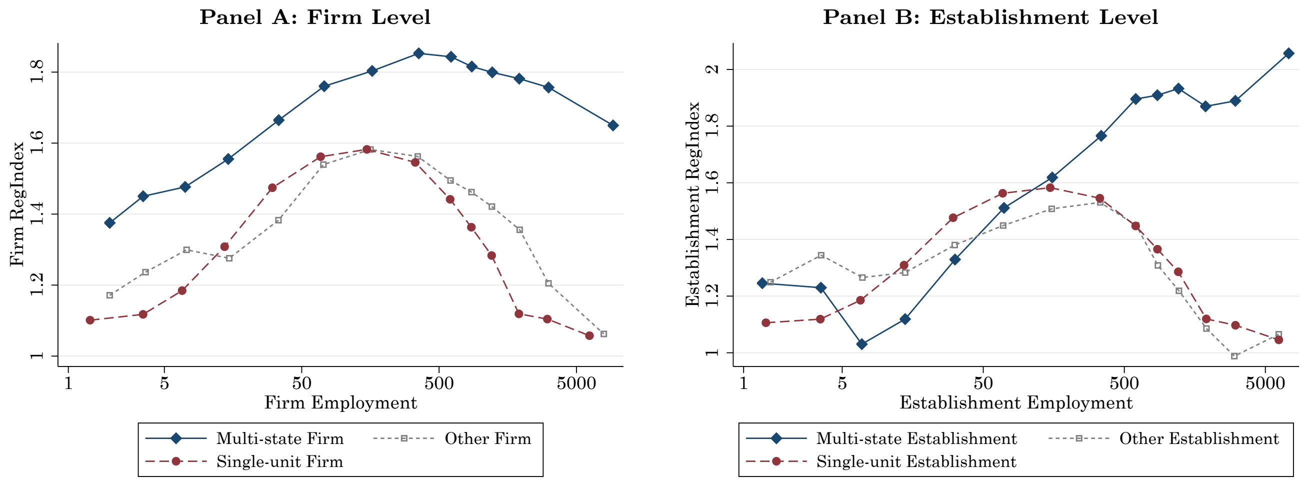
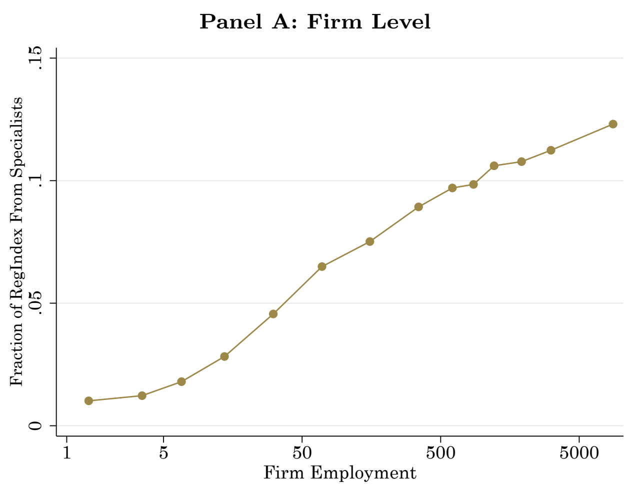


Figure 11: **Regulatory Compliance Specialization and Size**

This figure plots the fraction of RegIndex due to compliance specialists by employment bins for firms in Panel A and establishments in Panel B. Compliance specialists are defined as the top 25 occupations with the highest regulation intensity in Table A.2. We decompose a business's compliance costs into those paid to specialists and those paid to non-specialists. We plot the average fraction from specialists in businesses from each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.



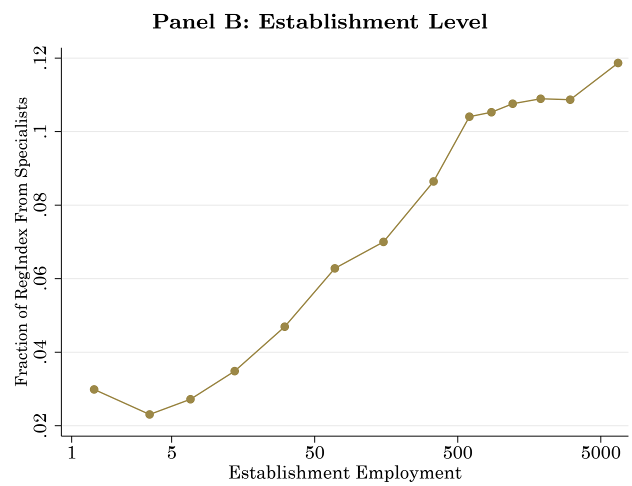


Figure 12: **Decomposing RegIndex: Regulatory Requirements Versus Enforcement**

This figure plots the two components of RegIndex as firms' regulatory compliance due to regulatory requirements and due to regulatory enforcement. See Section 5 for details of the decomposition procedure. We plot the average RegIndex for regulatory requirements and enforcement, respectively, from each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

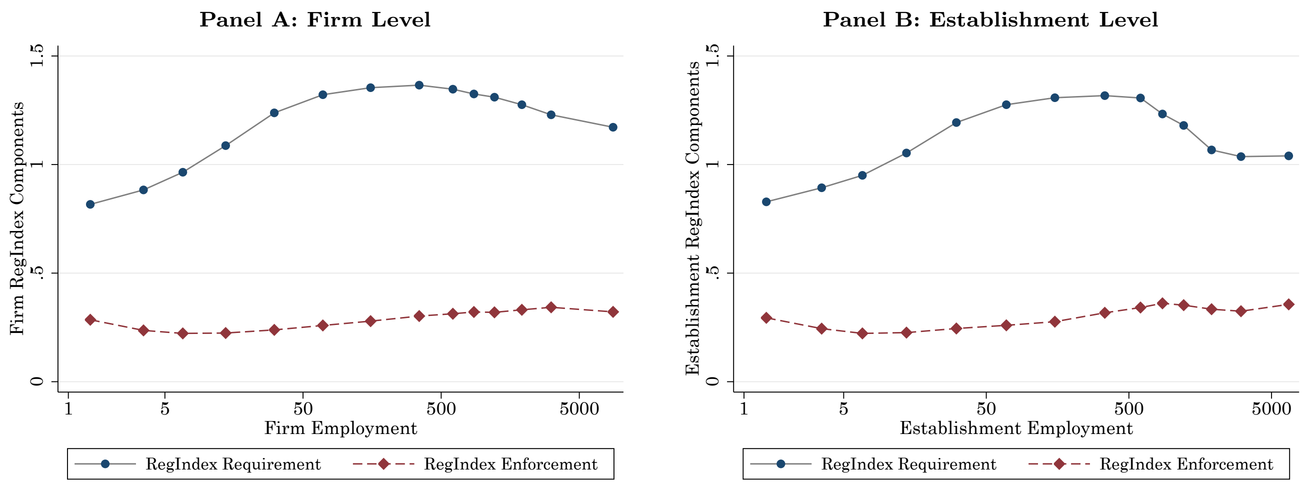


Table 1: **Summary Statistics**

Firms are defined by employer identification number (EIN) following Song et al. (2018). We bundle establishments of an EIN surveyed in year $t-2$ to t as a firm in year t following BLS convention. Industries are defined at NAICS 6-digit level. We aggregate establishments of an industry surveyed in year $t-2$ to t weighted by the establishments' sampling weights to compute industry-level metrics in t. RegIndex is the ratio of labor spending on regulation-related tasks and total labor spending in p.p. Sample period 2002-2014.

Variable	Mean	SD	P0.5	Median	P99.5	Obs.
			Panel A: Firms			
Employment	92.16	617.16	1.00	13.00	2,465.00	3,027,680
Annual Wage ($ mn)	4.07	31.30	0.02	0.46	115.48	3,027,680
RegIndex	1.34	1.88	0	0.86	10.46	3,027,680
			Panel B: Establishments			
Employment	47.79	192.45	1.00	13.00	875.00	3,364,336
Annual Wage ($ mn)	2.09	11.73	0.02	0.44	43.31	3,364,336
RegIndex	1.31	1.90	0	0.80	10.57	3,364,336
			Panel C: Industry			
Employment (1,000)	90.66	285.44	0.01	25.13	2,041.20	15,159
Annual Wage ($ mn)	3,611.91	11,112.21	0.12	1,001.32	67,466.05	15,159
RegIndex	1.66	1.02	0	1.60	5.58	15,159

Table 2: **Validation: Vote Share for Republican Party and RegIndex**

This table reports the results of regressing state vote shares for the Republican Party in the 2016 presidential election, the 2016 House, and the 2017-18 Senate on state 2014 RegIndex. See Figure 4 for the estimation of state 2014 RegIndex.

	Republican Party Vote Share 2016 Presidential Election (1)	Republican Party Vote Share 2016 House Elections (2)	Republican Party Vote Share 2017-18 Senate Elections (3)
State RegIndex	-0.640***	-0.966***	-1.828***
	(0.083)	(0.242)	(0.467)
Constant	1.502***	2.031***	3.408***
	(0.132)	(0.376)	(0.741)
Observations	51	50	50
Adjusted R^2	0.411	0.315	0.119

Table 3: **Validation: RegIndex Before and After IPO**

This table reports the results of regressing a publicly-traded firm's RegIndex on event year indicators during the 3 years before and 3 years after it goes for the initial public offering (IPO) in the U.S from 2002 to 2014. We obtain the firms IPO year from Jay Ritter's website. We treat the year before IPO ($t-1$) as the benchmark year, which is omitted in the regression. The mean RegIndex of IPO firms at $t-1$ is 1.798. In Column (1), we compute a publicly-traded firm's RegIndex by averaging its establishments' RegIndex weighted by its establishments' total labor costs. In Column (2), we demean the firm's RegIndex by subtracting the average RegIndex of its industry in the year, where industry for publicly-traded firms is the NAICS 4-digit obtained from the Compustat database. Section 3.5.4 provides more details. All regressions control for firm and year fixed effects. All standard errors are clustered at the firm level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	RegIndex (1)	RegIndex (Industry Adjusted) (2)
IPO_{t-3}	0.047	0.059
	(0.064)	(0.063)
IPO_{t-2}	-0.017	-0.003
	(0.037)	(0.036)
IPO_{t}	0.045	0.039
	(0.031)	(0.031)
IPO_{t+1}	0.135***	0.108**
	(0.047)	(0.046)
IPO_{t+2}	0.151***	0.111**
	(0.052)	(0.051)
IPO_{t+3}	0.153***	0.099*
	(0.053)	(0.051)
Year FE	Yes	Yes
Firm FE	Yes	Yes
Observations	5,164	5,164
Adjusted R^2	0.701	0.621

Table 4: **Economies of Scale for RegIndex**

Panel A reports the results of regressing firms' RegIndex on their employment and employment squared, where RegIndex is in percentage and employment is at 1,000s. Panel B reports the results at the establishment level. All standard errors are double clustered by year and firm in Panel A and at year and establishment in Panel B. Industry is defined at NAICS 6-digit level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Given a quadratic formula $y = ax^2 + bx + c$, the *max* is computed as $c - \frac{b^2}{4a}$, while the *argmax* is computed as $-\frac{b}{2a}$. The sample period is from 2002 to 2014.

Panel A: Firm-Level					
	(1)	(2)	(3)	(4)	(5)
Emp	2.897***	2.920***	2.008***	1.935***	0.544***
	(0.065)	(0.065)	(0.076)	(0.068)	(0.074)
Emp^2	-2.902***	-2.927***	-1.963***	-1.909***	-0.542***
	(0.068)	(0.068)	(0.073)	(0.064)	(0.069)
max	1.961***	1.965***	1.782***	1.755***	1.517***
	(0.028)	(0.015)	(0.018)	(0.016)	(0.018)
argmax	0.499***	0.499***	0.511***	0.507***	0.501***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.018)
Year FE	-	Yes	-	-	Yes
Year-Ind FE	-	-	Yes	-	-
Year-Ind-State FE	-	-	-	Yes	-
Firm FE	-	-	-	-	Yes
Observations	3,027,680	3,027,680	3,027,241	2,918,296	2,162,080
Adjusted R^2	0.007	0.007	0.378	0.412	0.597
Panel B: Establishment-Level					
	(1)	(2)	(3)	(4)	(5)
Emp	3.607***	3.788***	2.985***	2.902***	0.612***
	(0.265)	(0.217)	(0.148)	(0.134)	(0.125)
Emp^2	-5.255***	-5.510***	-4.452***	-4.397***	-0.992***
	(0.486)	(0.430)	(0.284)	(0.259)	(0.197)
max	1.823***	1.850***	1.724***	1.699***	1.444***
	(0.041)	(0.019)	(0.016)	(0.014)	(0.016)
argmax	0.343***	0.344***	0.335***	0.330***	0.309***
	(0.007)	(0.007)	(0.006)	(0.005)	(0.012)
Year FE	-	Yes	-	-	Yes
Year-Ind FE	-	-	Yes	-	-
Year-Ind-State FE	-	-	-	Yes	-
Establishment FE	-	-	-	-	Yes
Observations	3,362,824	3,362,824	3,362,418	3,255,415	2,194,239
Adjusted R^2	0.005	0.006	0.371	0.408	0.534

Table 5: **Regulatory Compliance Within Occupation**

This table reports the relationship of firms' requirements on regulatory compliance skills in job postings and their size *within an occupation*. The dependent variable is the job posting's regulatory compliance skill requirement constructed as following: We first name match firms in our OEWS sample to firms in the Burning Glass Technologies (BGT) data, which provide 14 million job postings from 2010 to 2014 (the overlapping period between Burning Glass and our OEWS sample) where each job posting describes a set of required skills for the job. We next label each skill as performing regulation-related tasks by applying our methodology in Section 4 to the textual contents for the skill in the BGT data. A job posting's *regulatory compliance skill requirement* is the share of all skill requirements in the job posting that is about performing regulation-related tasks (in percentage). We regress the job posting's regulatory compliance skill requirement on the posting firm's employment (in thousands) and the squared of the employment, while controlling for year-occupation fixed effects. Column (1) reports results for the full sample with all SOC 6-digit occupations. Column (2) requires results for the subsample of occupations that have positive regulation-task intensity (RTI) in our definition in Section 3.3, while Column (3) reports the results for the subsample of occupations with RTI equal to 0. All standard errors are clustered at the firm (EIN) level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Given a quadratic formula $y = ax^2 + bx + c$, the *max* is computed as $c - \frac{b^2}{4a}$, while the *argmax* is computed as $-\frac{b}{2a}$. See Section 4.1.1 for details.

	Dependent Variable: Regulatory Compliance Skill Requirement		
Sample:	All Occ (1)	Occ (RTI>0) (2)	Occ (RTI = 0) (3)
Emp	0.722*** (0.204)	0.536*** (0.168)	0.858*** (0.290)
Emp^2	-0.445*** (0.131)	-0.341*** (0.102)	-0.520*** (0.181)
max	1.558*** (0.092)	2.073*** (0.071)	1.188*** (0.128)
argmax	0.811*** (0.048)	0.787*** (0.048)	0.825*** (0.059)
Year-Occ FE	Yes	Yes	Yes
Observations	14,052,988	5,878,039	8,174,949
Adjusted R^2	0.139	0.138	0.119

Table 6: **Response of RegIndex to Regulatory Requirement and Enforcement Changes**

This table reports the results of regressing 3-year changes in a firm's log RegIndex on the shift-share variables. Panel A shows the results using the full sample, and Panel B shows the results in subsamples divided by firms or establishments' employment size. Equation (7) provides the regression specification. $iv(\Delta\log(p_{it}))$ and $iv(\Delta\log(\tilde{R}_{it}))$ are the shift-share variables for the NAICS 6-digit industry's enforcement changes and regulatory-requirement changes. Section 5 provides details on the construction of the shift-share variables. All regressions include controls of 3-year changes in the log total wage costs of the firm or establishment, 3-year changes in the log wage rate of the regulation-related tasks, and beginning of period log RegIndex to account for mean-reversion of firms' or establishments' RegIndex, and year fixed effects. All variables are standardized to have mean 0 and variance of 1 for ease of interpretation. Standard errors are double clustered by industry and year and presented in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. The sample period is from 2002 to 2011.

Panel A: Full Sample

	Firm-Level			Establishment-Level		
	(1)	(2)	(3)	(4)	(5)	(6)
$iv(\Delta\log(p_{it}))$	0.209**		0.112	0.234***		0.128*
	(0.077)		(0.075)	(0.063)		(0.065)
$iv(\Delta\log(\tilde{R}_{it}))$		0.257***	0.228***		0.269***	0.231***
		(0.038)	(0.054)		(0.036)	(0.052)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	608,500	608,500	608,500	628,733	628,733	628,733
Adjusted R^2	0.322	0.340	0.344	0.323	0.340	0.345

Panel B: Subsamples by Size

	Firm-Level				Establishment-Level			
	1-19	20-399	400-749	≥ 750	1-19	20-399	400-749	≥ 750
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$iv(\Delta\log(p_{it}))$	0.102	0.120	0.099	0.083	0.115*	0.140*	0.168**	0.077
	(0.075)	(0.078)	(0.076)	(0.064)	(0.062)	(0.066)	(0.074)	(0.066)
$iv(\Delta\log(\tilde{R}_{it}))$	0.169**	0.258***	0.289***	0.309***	0.170**	0.271***	0.313***	0.367***
	(0.056)	(0.054)	(0.054)	(0.046)	(0.053)	(0.050)	(0.053)	(0.055)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	189,404	352,779	29,482	36,835	220,464	375,485	19,622	13,162
Adjusted R^2	0.400	0.324	0.243	0.214	0.397	0.321	0.265	0.233

Appendix

"The Cost of Regulatory Compliance in the United States"

Francesco Trebbi   Miao Ben Zhang   Michael Simkovic

- Not for Publication -

A Details of Constructing *RegIndex*

The construction of the regulatory index, RegIndex, starts with identifying which tasks are related to regulatory compliance. We achieve this goal by analyzing the texts of task statements in the O*NET data in two phases: a *keyword matching phase* and an *annotation phase.* In what follows, we detail the procedure for estimating our preferred (conservative) version of the RegIndex, followed by a description of two broader versions of the RegIndex.

Keyword Matching Phase We identify regulation-related tasks by matching the task statements to two different tiers of keywords. Two tiers are necessary in order to balance the rate of false positives and false negatives in the identification of tasks. The first tier of keywords includes the words *regulation*, *regulations*, and *regulatory.* These matches intuitively identify as a "regulation-related" task whose statement explicitly mentions regulation. These matches exhibit a low rate of false positives. Appendix Table A.1 lists ten examples of regulation-related tasks identified by the first-tier keywords.

The first tier of keywords produces, however, an excessive rate of false negatives. Regulation can be described by various alternative words and phrases, and regulatory compliance behavior can be described without directly mentioning *regulation, regulations*, or *regulatory.* For this reason we employ a second tier of keywords for identifying regulation-related tasks. The second tier of keywords includes (A) alternative references to regulation: *law, laws, legislation, legislative, statute, statutes, statutory, ordinance, ordinances, code, codes, standards, public policy, public policies, license, licenses, licensing, permit, permits, certification, certifications*; (B) references to government agencies: *government, governments, governmental, federal, legislature, policy maker, policy makers, governing agency, governing agencies, public agency, public agencies*; and (C) actions of compliance: *compliance, noncompliance.* These matches mitigate false negatives but may introduce false positives.

Annotation Phase In this second phase, we employ two procedures to refine the quality of the keyword-matched regulation-related tasks. First, we manually annotate each matched task statement and exclude further false positives, such as tasks in which the word "code" or "codes" means computer programming codes, tasks in which the word "permit" or "permits" is a verb instead of a noun, tasks which include the word "government", but are unrelated to government regulation, etc. This manual annotation procedure results in a final list of 829 regulation-related tasks out of a total of 19,636 tasks in the O*NET database.[70]

Second, we assign a *regulation-relatedness value* between 0 and 1 to account for the heterogeneity among regulation-related tasks. Specifically, tasks identified by the second-tier keywords are less informative of the tasks' relevance to government regulation (e.g., the task may mention "licenses", but may be unclear about whether the licenses are issued by the government or private entities). We thus assign tasks identified by the first-tier keywords a regulation-relatedness value of 1, tasks identified only by the second-tier keywords a value of 0.75.[71] Moreover, while a task statement is usually just one sentence, the statements may differ in their informativeness of the tasks' relevance to government regulation. Some tasks have only one focus, while others may have multiple focuses. For instance, the following task, "*maintain awareness of advances in medicine, computerized diagnostic and treatment equipment, data processing technology, government regulations, health insurance changes, and financing options*" has six focuses, with only one of them related to government regulations. Thus, we compute the share of regulation-related focuses out of the total number of focuses in each statement and multiply this share by the regulation-relatedness value. The regulation-relatedness value for the 829 regulation-related tasks has a mean of 0.55 and a standard deviation of 0.31.

This second step of assigning a regulation-relatedness value between 0 and 1 to each regulation-related task allows us to measure regulatory compliance costs based on the most informative component of each task statement, which further mitigates false positives in our measure. We corroborate this conservative approach with two broader, alternative approaches to measuring regulatory compliance costs. First, we replace the regulation-relatedness value with 1 for each regulation-related task, which removes the downweighting and treats regulation-related tasks as equally informative about the intensity of regulatory compliance. We name this approach the "medium" approach compared with our main "conservative" approach. Second, we further expand the medium approach by including another 54 tasks that include the keywords *Tax* or *Taxes*. Tax compliance is included in the OIRA's estimated total regulatory compliance

[70]Our regulation-related tasks can be downloaded at www.miaobenzhang.com/Regulation_Related_Tasks.xlsx.

[71]Our findings are robust to alternative levels of downweighting second-tier keywords.

hours. Hence, this "broad" approach helps us compare our regulatory compliance hours with the OIRA's reported regulatory compliance hours. In what follows, we focus on the regulation-relatedness value from the conservative approach to produce our main measure, due to its virtual absence of false positives.

RegIndex Construction Having measured the regulation-relatedness of each task, we next compute the regulation-task intensity (RTI) for each SOC 8-digit occupation by averaging its tasks' regulation-relatedness values weighted by the tasks' importance ratings for that occupation from O*NET.[72] We then aggregate the RTI to the SOC 6-digit level.

Appendix Table A.2 lists the top 25 SOC 6-digit occupations with the highest RTI. For instance, Compliance Officers have the highest RTI of 0.343. We interpret this RTI as indicating that Compliance Officers on average spend 34.3 percent of their work hours on directly performing government regulation-related tasks (e.g., "*evaluate applications, records, or documents to gather information about eligibility or liability issues*" or "*keep informed regarding pending industry changes, trends, or best practices*"). While it is tempting to consider Compliance Officers having RTI of 100 percent, in practice, some tasks of compliance officers may serve firm production without been necessarily linked to regulation. For instance, Environmental Compliance Inspectors (13-1041.01), a subcategory of Compliance Officers (13-1041), perform the task "*maintain and repair materials, worksites, and equipment*", and Regulatory Affairs Specialists (13-1041.07) also perform the task "*develop or track quality metrics*".

Finally, we merge each SOC 6-digit occupation's RTI to the relevant establishments in the OEWS data, which provides each establishment's labor cost (employment times wage rate) for each occupation.[73]

B Details on Shift-Share Variables Construction

The regulatory agencies employed for the construction of the shift-share variables include Department of Labor (DOL), Department of Transportation (DOT), Department of Education (ED), Environmental Protection Agency (EPA), Federal Communications Commission (FCC), Federal Deposit Insurance Cor-

[72]Literature has frequently used O*NET's importance ratings to aggregate measures from the task level to the occupation level (see Acemoglu and Autor, 2011; Acemoglu et al., 2020). We also perform a robustness check by aggregating tasks' RTI to the occupation level using a frequency rating from O*NET which indicates how frequent a task is performed by the occupation. This weighting yields an occupation-level RTI measure that is near identical to our original RTI with a correlation of 99.4%.

[73]Wage rate in the OEWS survey includes "*base rate pay, cost-of-living allowances, guaranteed pay, hazardous-duty pay, incentive pay such as commissions and production bonuses, and tips are included in a wage. Back pay, jury duty pay, overtime pay, severance pay, shift differentials, non-production bonuses, employer costs for supplementary benefits, and tuition reimbursements are excluded.*" See details on the technical notes of the OEWS at https://www.bls.gov/oes/oes_doc_arch.htm.

poration (FDIC), Federal Trade Commission (FTC), Department of Health and Human Services (HHS), Department of Housing and Urban Development (HUD), Nuclear Regulatory Commission (NRC), Securities and Exchange Commission (SEC), and Department of Agriculture (USDA). According to the OIRA reports, regulations from these twelve major agencies require 1.65 billion estimated compliance hours per year on average from 2002 to 2014, accounting for 82% of the total non-treasury compliance hours during that period.[74] Appendix Figures A.13 and A.14 plot the time series for $\Delta \log enf_{kt}$ and $\Delta \log reg_{kt}$ for each main agency and the corresponding levels in our sample.

Our procedure identifies 227 out of 612 occupations of federal employees as "regulation-related." In the Appendix C, we conduct a robustness check by identifying a narrower list of 105 enforcement-focused regulation-related occupations, and we re-estimate our model using this alternative measure of enforcement changes which is potentially a more volatile supply-side variable for regulatory enforcement. Appendix Tables A.14-A.17 show virtually no substantive differences between results using the alternative changes from our main results.

Our IV's exhibit desirable properties that are intuitive. First, Appendix Table A.10 reports the top 3 industries for each regulatory agency. The table shows intuitive profiles of oversight, which supports by and large the intuitive validity of our approach based on keywords. For instance, the Environmental Protection Agency (EPA) reports Waste Management and Remediation Services; Petroleum and Coal Products Manufacturing; and Construction of Buildings as its top industries under oversight. The Securities and Exchange Commission (SEC) reports Securities, Commodity Contracts, and Other Financial Investments and Related Activities; Credit Intermediation and Related Activities; Funds, Trusts, and Other Financial Vehicles as its top industries, and so on. Second, we explore how regulatory and enforcement changes co-move over time. Appendix Figure A.13 traces the time series of the two changes for each main regulatory agency. We observe that there is substantial independent variation in each of the two change series across all regulators, although for some agencies the separation is starker. Dynamics also differ. For instance, SEC sharply accelerates hiring in the aftermath of the financial crisis of 2008, while other agencies, like FCC, do not. Finally, Appendix Table A.11 shows that the overall correlation between the shift-sahre variables is 12 percent at the establishment level.

[74] OIRA report for each year can be downloaded from the Whitehouse website at https://www.whitehouse.gov/omb/information-regulatory-affairs/reports/#ICB.

C Measuring Enforcement Employment

As a robustness check to using our main measure of enforcement changes in Section 5.2.2, which is the based on major regulatory agencies' employment of regulation-related occupations, we construct a refined measure based on major regulatory agencies' employment of enforcement-focused regulation-related occupations. Our data as described in Section 5.2.2 cover detailed information of federal government employees such as their agency, occupation, and fulltime/parttime status from 2002 and 2014. Section 5.2.2 has described our method of identifying regulation-related occupations. We further identify "enforcement" occupations among regulation-related occupations. To do so, we first obtain each federal employee occupation's task description from "Handbook of Occupational Group and Families" at the US OPM website.[75] Then we identify an occupation as enforcement-focused if its task description includes the following keywords: "*enforcement, enforce, enforces, supervision, supervisory, monitor, monitors, oversight, oversee, oversees, sanctions, sanction, penalty, penalties, fine, fines, inspect, inspects, inspection, inspections, investigate, investigates, investigation, investigations, examine, examines, examination, examinations.*" These procedures identify 105 "enforcement" occupations out of a total of 227 regulation-related occupations.

Finally, we apply our definitions of enforcement-related regulatory occupations to the twelve agency's employment, and compute the 3-year log differences for each agency's enforcement employment, which is an alternative measure of $\Delta \log enf_{kt}$ in equation (5). Tables A.14-A.17 present results using enforcement changes, $iv(\Delta\log(p_{it}))$, based on this alternative measure.

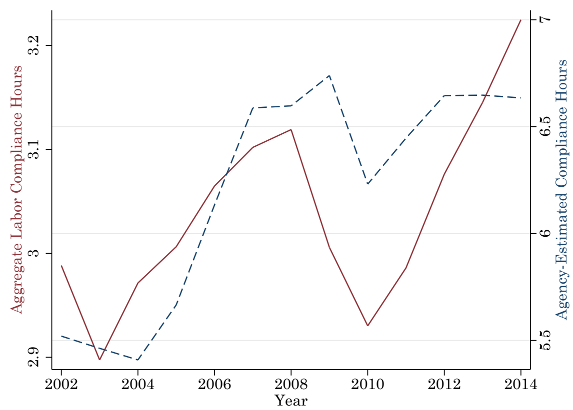
[75]The OPM has its own definitions for government occupations that are different from the SOC system. The handbook for OPM occupation description can be downloaded at https://www.opm.gov/policy-data-oversight/classification-qualifications/classifying-general-schedule-positions/occupationalhandbook.pdf.

D Additional Figures and Tables

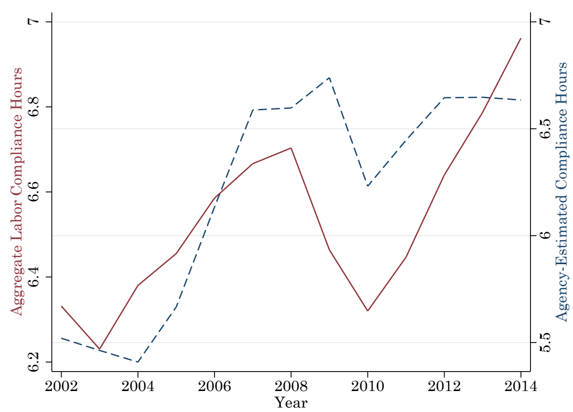
Figure A.1: **Validation of Regulation Index Using Agency-Estimated Compliance Hours—All Versions**

This figure plots the aggregate annual compliance hours (in billions of hours) identified by our most conservative RegIndex measure in Panel A and the hours identified by two broader versions of the RegIndex measure in Panel B (RegIndex without downweighting regulation-related tasks due to focus or keywords in task statements) and Panel C (RegIndex using the Panel B definition plus tax compliance). In each panel, we compare the aggregate compliance hours identified by our measure with the estimated annual compliance hours submitted by various regulatory agencies to OIRA.

Panel A: Compliance Hours Based on Conservative RegIndex (Main Measure)



Panel B: Compliance Hours Based on Medium RegIndex



Panel C: Compliance Hours Based on Broad RegIndex

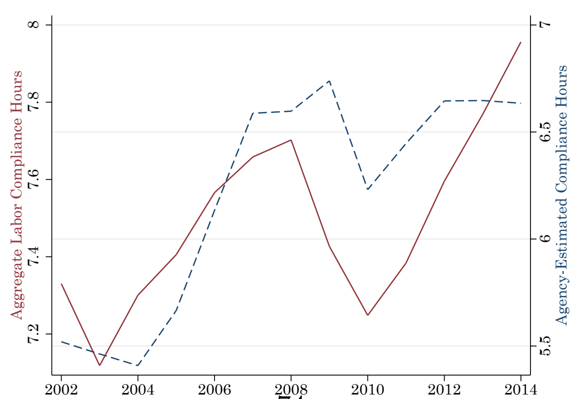


Figure A.2: **Case Studies of Industry Regulatory Policy Changes Using RegData**

This figure plots the response of industries' RegData measure to five industry-level regulatory policy changes. RegData is from RegData Version 3.2. from QuantGov.com and is the natural logarithm of the count of restrictive words in the Code of Federal Regulations governing an industry in the year. Section 3 provides details of the industry regulatory policy changes and discusses the classification of treated and control groups. To ease the comparison, we shift the lines vertically so that they have the same value in the year before the treatment. The value in the year before the treatment is the average of the regulation measures across the treated and control industries in that year. The difference between the two lines after the treatment, minus the difference between the two lines before the treatment reflects the difference-in-differences estimation.

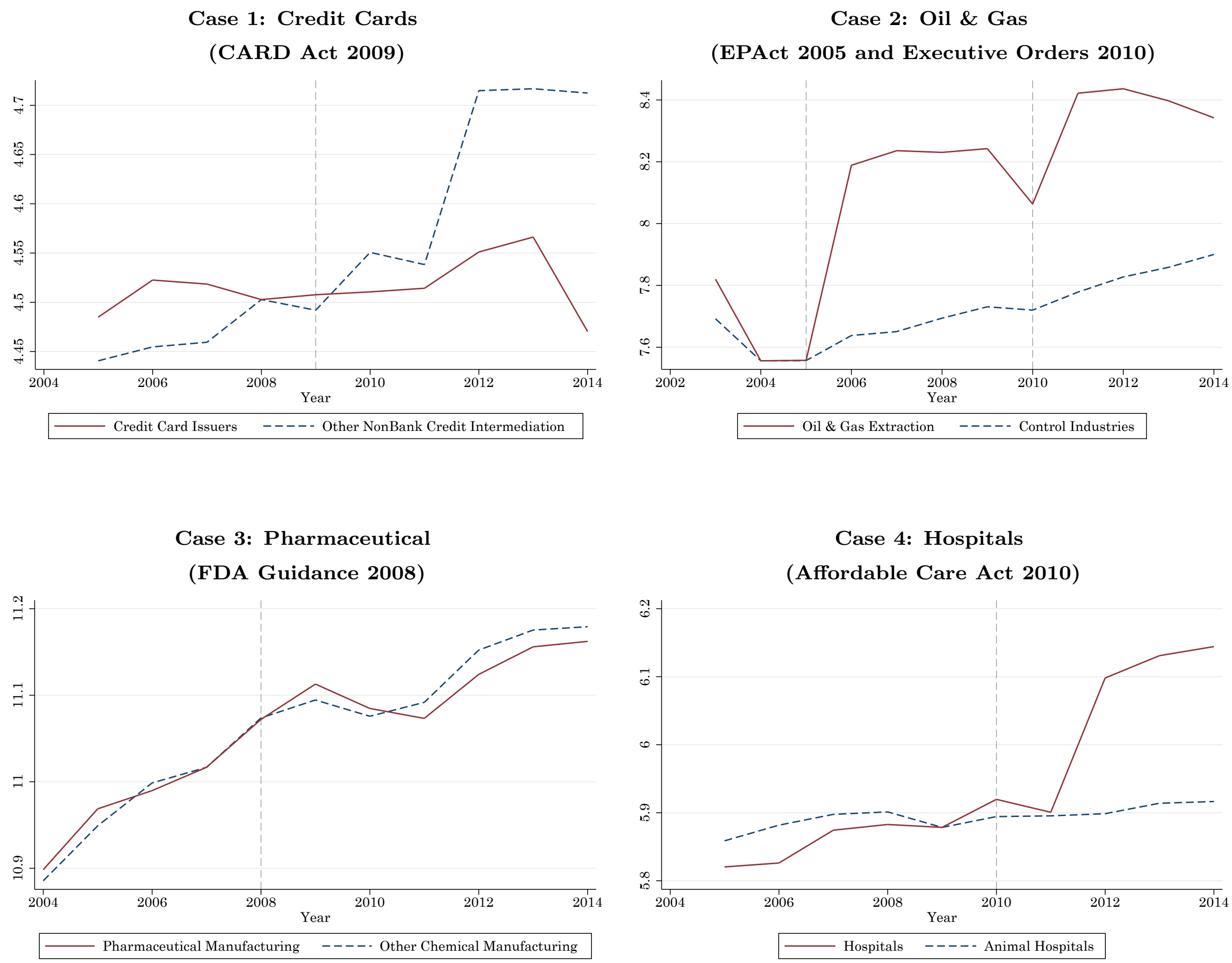


Figure A.3: **Robustness: Industry Regulatory Policy Changes Using Medium Version of RegIndex**

This figure shows a similar result of Figure 3 by plotting the response of the medium version of industries' Regulation Index (RegIndex) to five industry-level regulatory policy changes. RegIndex is the percentage of an industry's annual labor spending on performing the medium version of regulation-related tasks. See Appendix A for the definitions of our RegIndex versions. Section 3 provides details of the industry regulatory policy changes and discusses the classification of treated and control groups. To ease the comparison, we shift the lines vertically so that they have the same value in the year before the treatment. The value in the year before the treatment is the average of the regulation measures across the treated and control industries in that year. The difference between the two lines after the treatment, minus the difference between the two lines before the treatment reflects the difference-in-differences estimation. The shaded areas indicate the 95% confidence interval of the industries' average RegIndex.

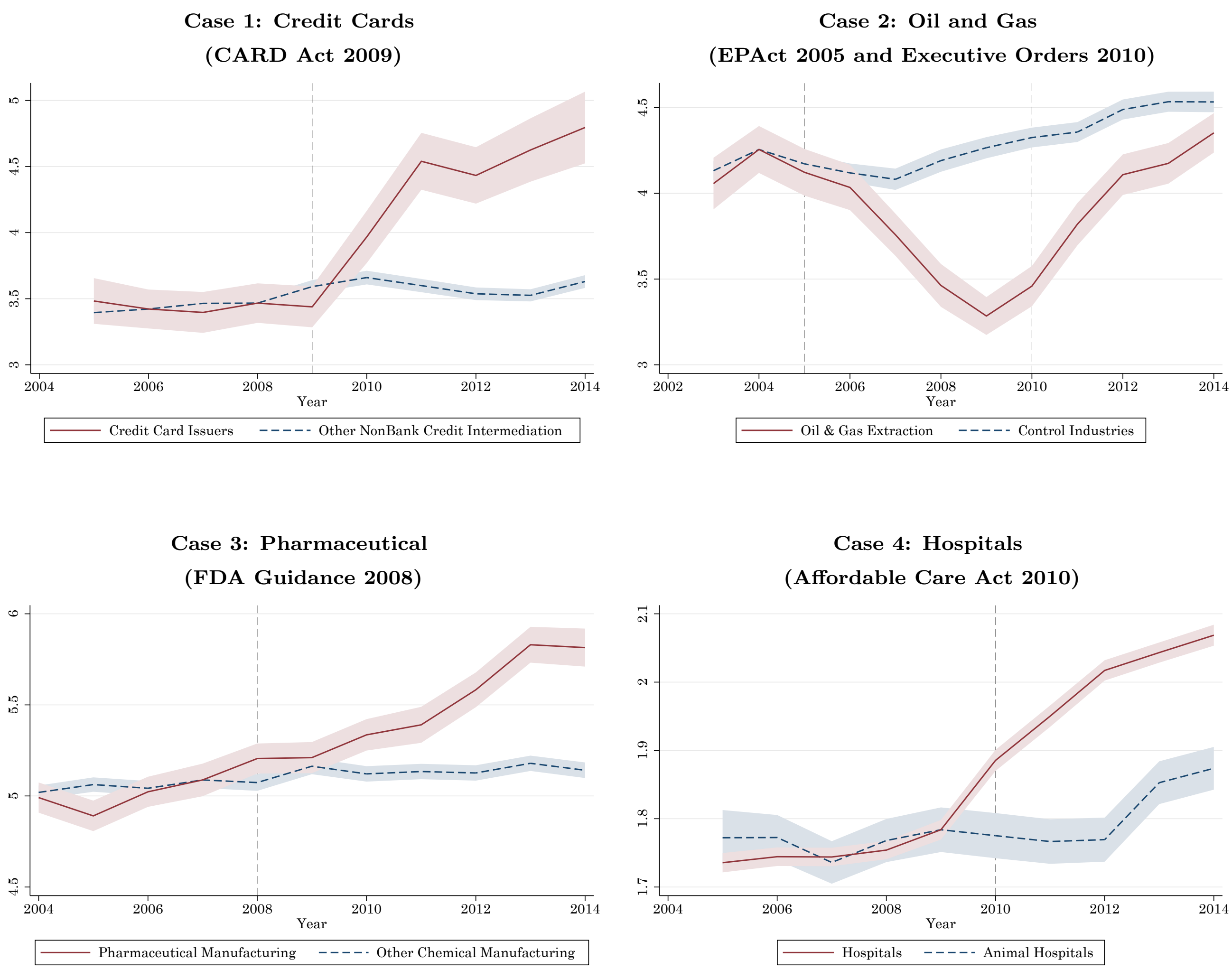


Figure A.4: **Robustness: Industry Regulatory Policy Changes Using Broad Version of RegIndex**

This figure shows a similar result of Figure 3 by plotting the response of the broad version of industries' Regulation Index (RegIndex) to five industry-level regulatory policy changes. RegIndex is the percentage of an industry's annual labor spending on performing the broad version of regulation-related tasks. See Appendix A for the definitions of our RegIndex versions. Section 3 provides details of the industry regulatory policy changes and discusses the classification of treated and control groups. To ease the comparison, we shift the lines vertically so that they have the same value in the year before the treatment. The value in the year before the treatment is the average of the regulation measures across the treated and control industries in that year. The difference between the two lines after the treatment, minus the difference between the two lines before the treatment reflects the difference-in-differences estimation. The shaded areas indicate the 95% confidence interval of the industries' average RegIndex.

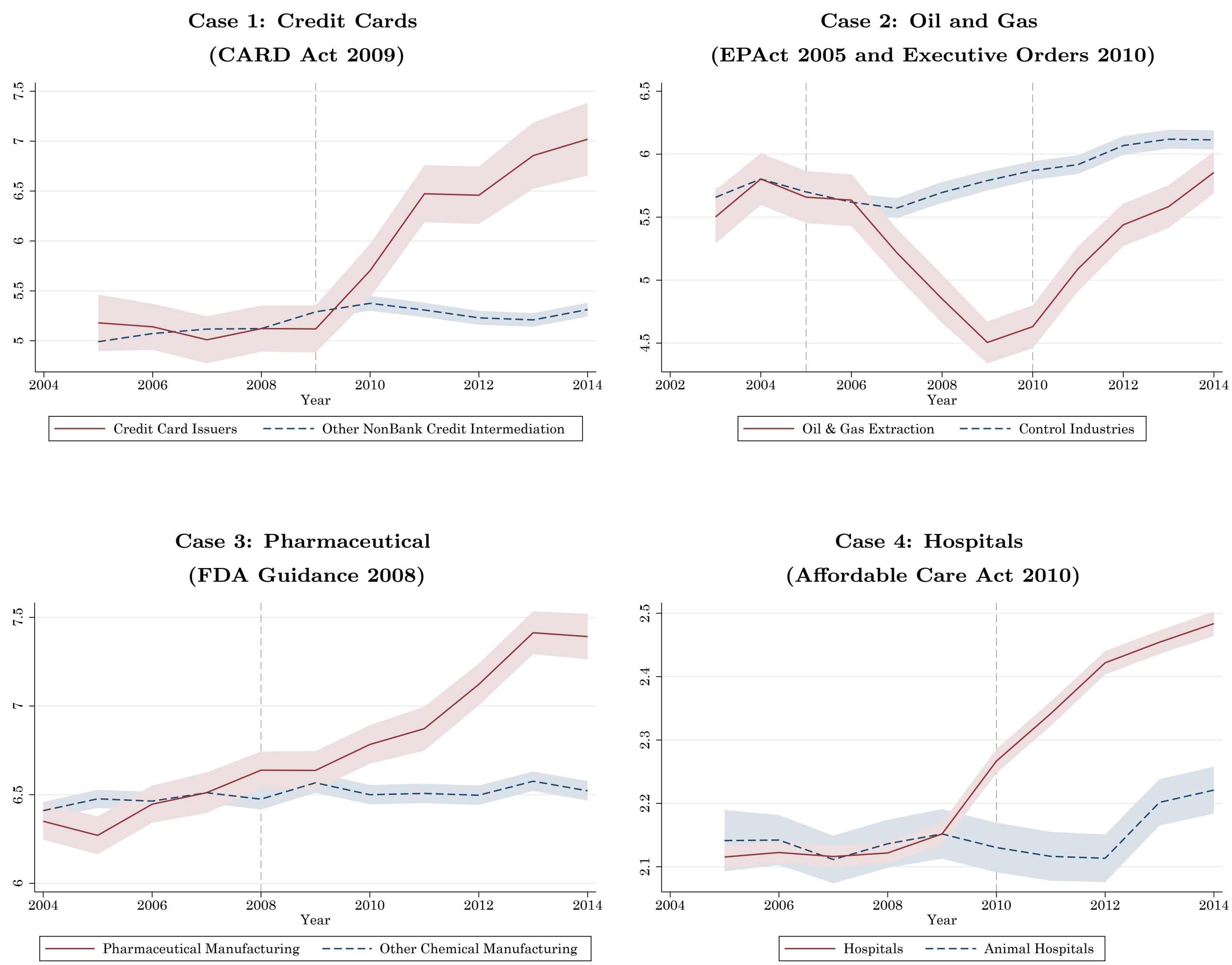


Figure A.5: **RegData Across States**

This figure plots the state-level RegData measure from QuantGov.org.

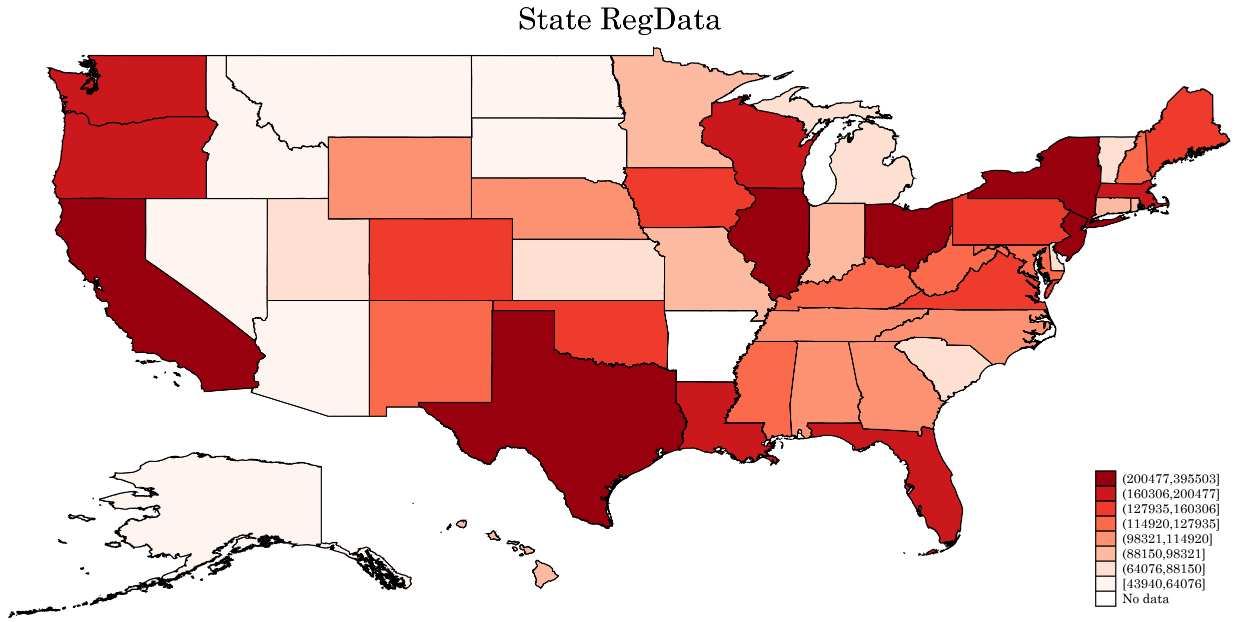


Figure A.6: **Robustness: Other Versions of RegIndex and Size**

This figure shows the robustness check results of Figure 5 by plotting the relation of the medium and broad-versions of RegIndex and employment for firms in Panels A and C and establishments in Panels B and D. See Appendix A for the definitions of our RegIndex versions. The dots represent the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The line represents the LOWESS smoothed fitted curve using the bandwidth of 0.05. The x-axis is on a log scale.

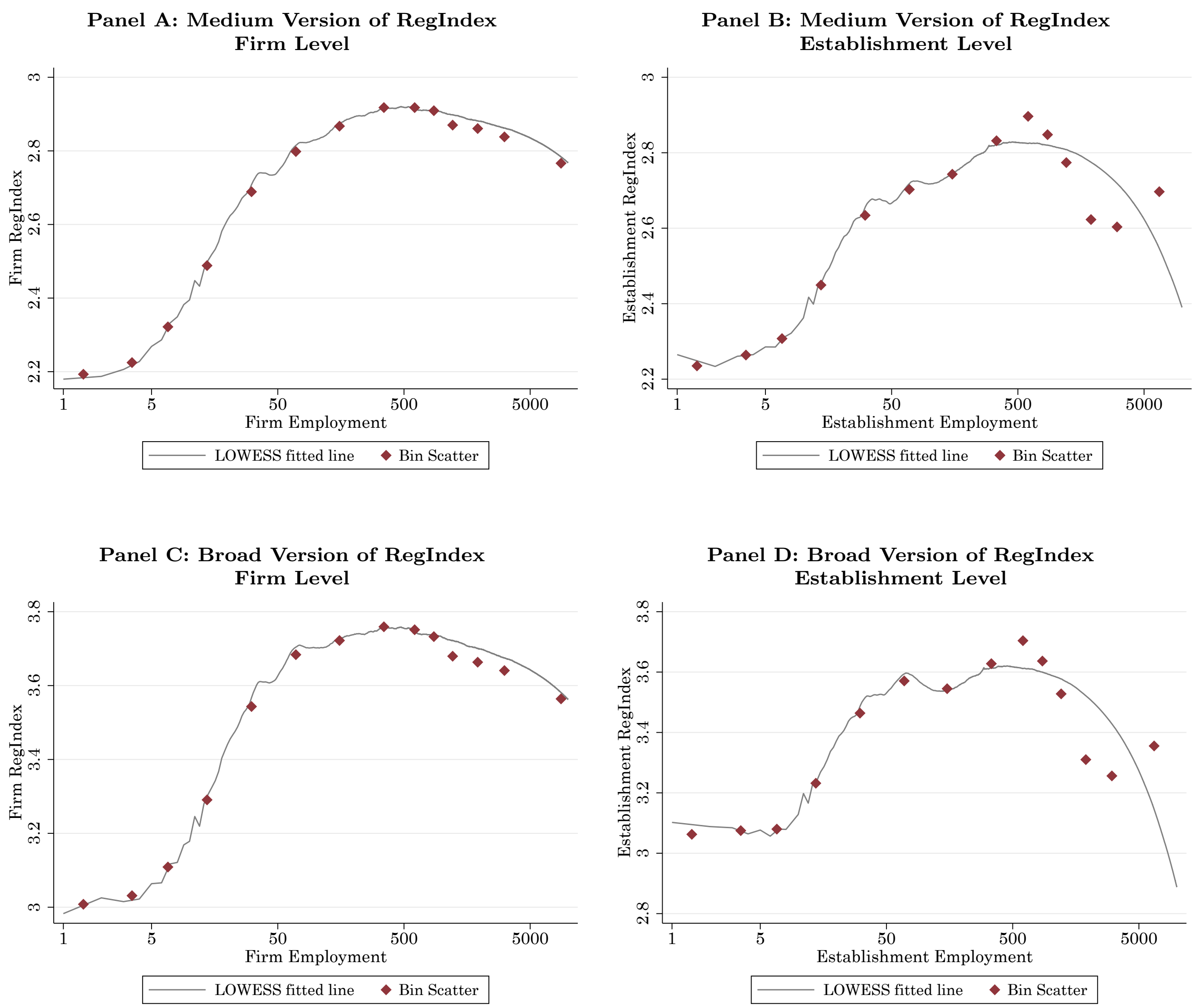


Figure A.7: **RegIndex and Size by Year**

Panel A shows the relation of RegIndex and firm employment in each year from 2002 to 2014. Panel B shows the relation at the establishment level. Each dot represents the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

Panel A: Firm-Level

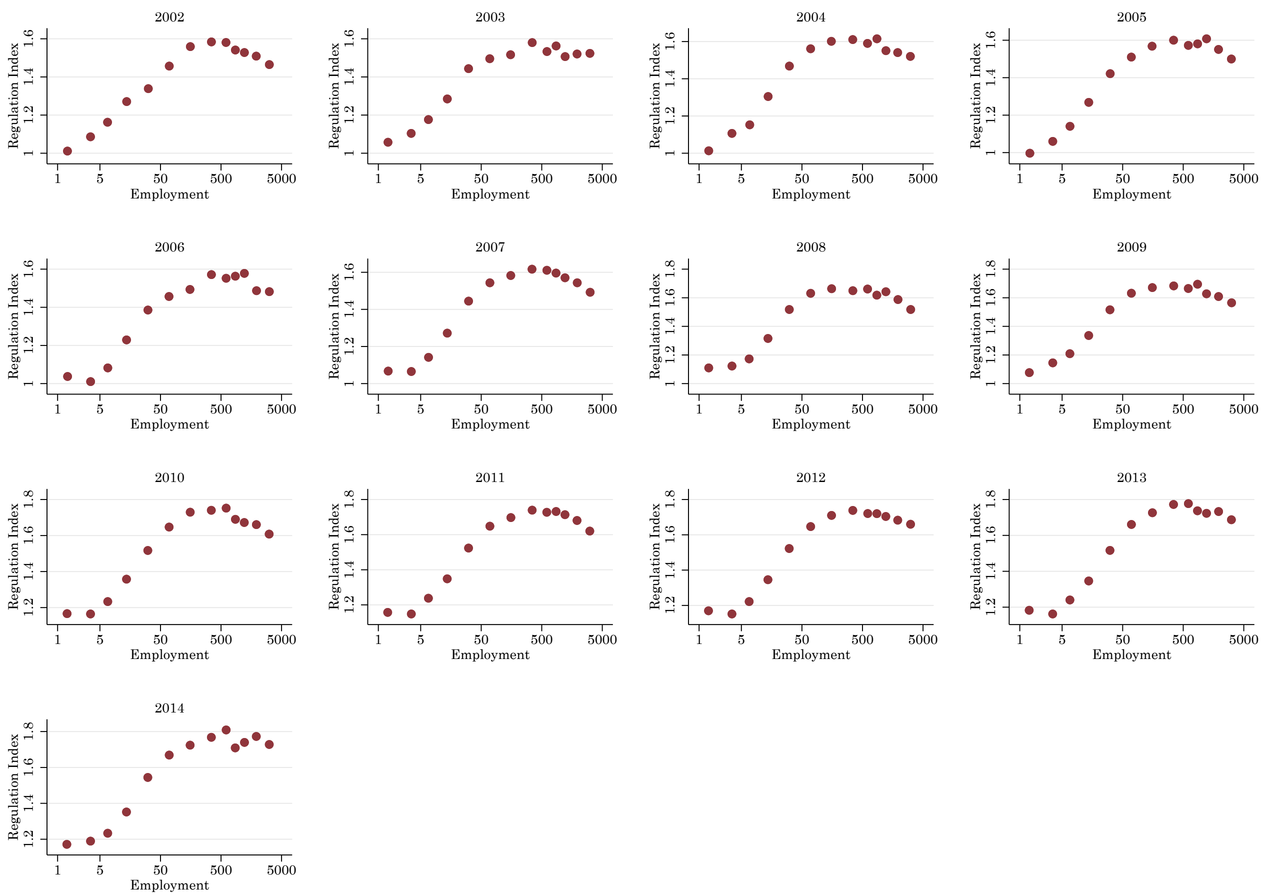


Figure A.7: **RegIndex and Size by Year**—*Continued*

Panel A shows the relation of RegIndex and firm employment in each year from 2002 to 2014. Panel B shows the relation at the establishment level. Each dot represents the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

Panel B: Establishment-Level

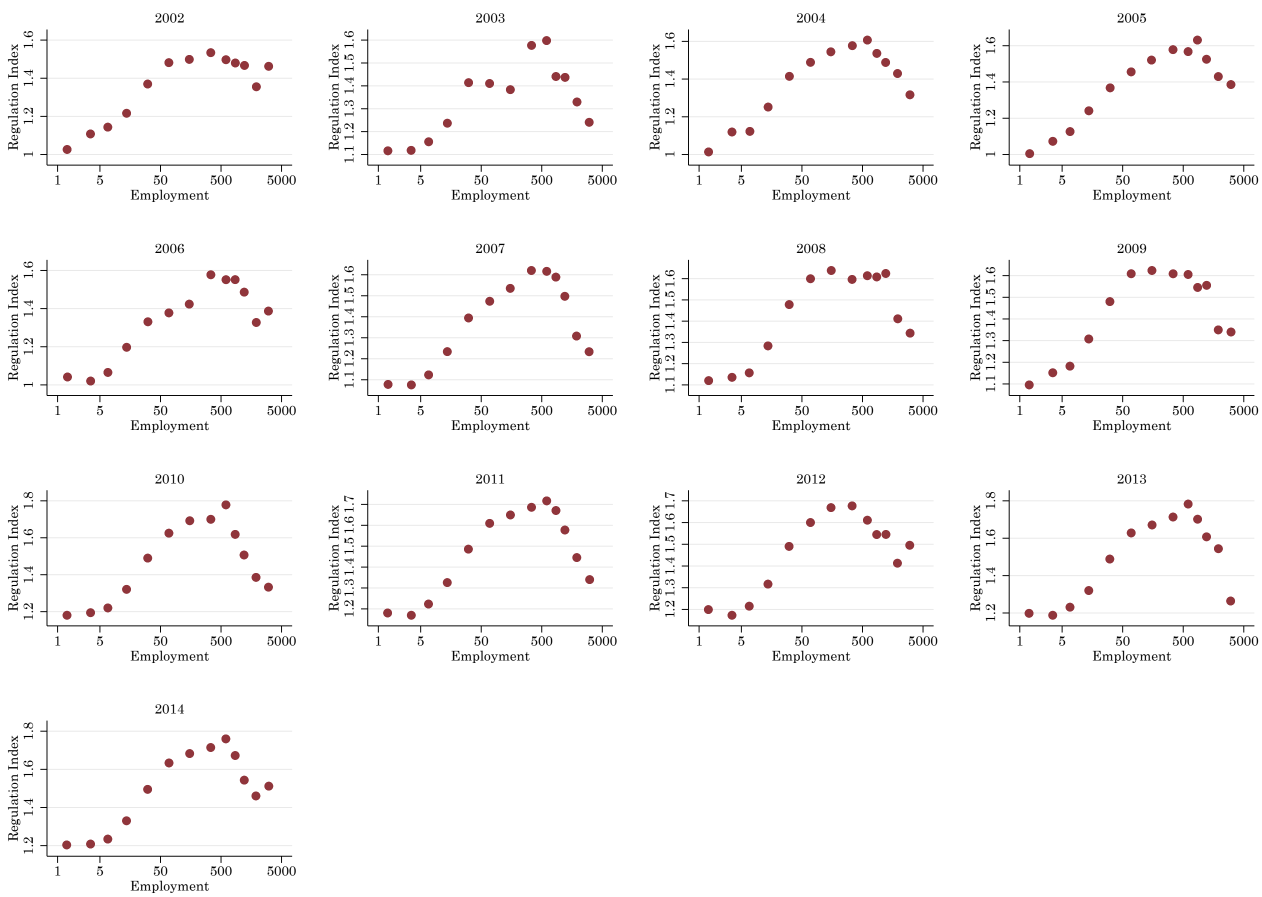


Figure A.8: **Outsource-Adjusted RegIndex Using Financial Industry Report**

This figure plots the baseline RegIndex and the RegIndex adjusted for outsourced compliance costs provided by the Securities Industry Association (2006) report for firms in Panel A and establishments in Panel B from the financial industry (NAICS = 52). The industry report provides the relative outsourced costs and in-house labor costs for regulatory compliance for firms in three size bins. Small, medium, and large size are defined as firms with employees fewer than 100, between 100 and 1000, and more than 1000, respectively. For businesses in each size bin, we compute the outsource-adjusted RegIndex using the outsourced to in-house labor cost ratio. See Section 4.1.3 for more details.

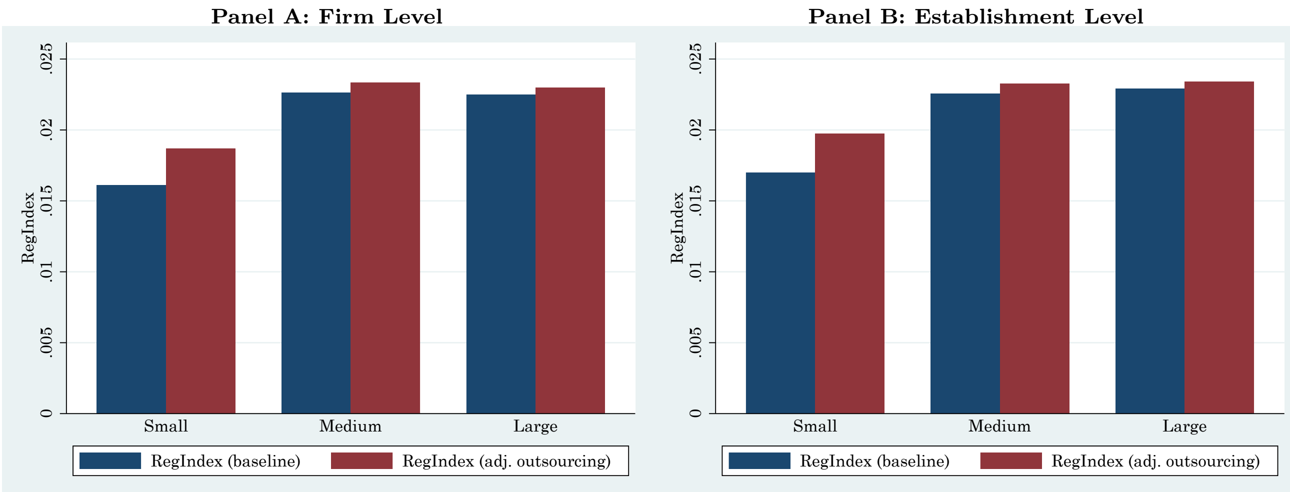
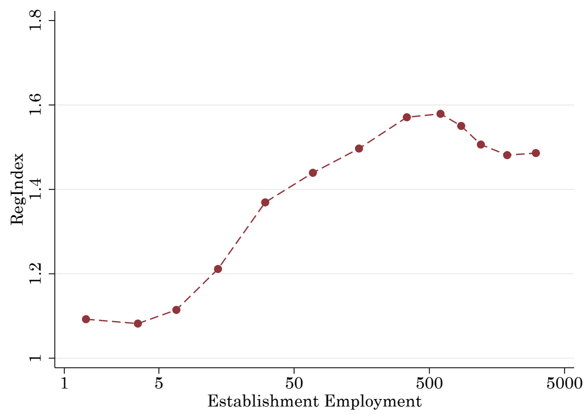


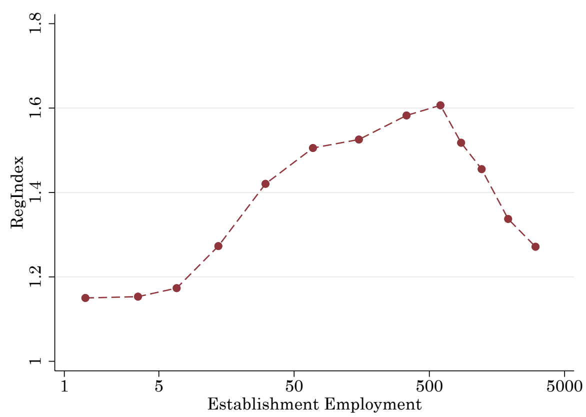
Figure A.9: **First Pass Check: RegIndex and Size by Counties' Outsourcing Abundance**

This figure plots the relation of RegIndex and employment for establishments from three subsamples of counties based on the counties' normalized total outsourced compliance costs. A county's normalized total outsourced compliance costs are the total labor spending on regulation-related tasks from outsourcing firms in the county divided by non-outsourcing firms total labor spending in the county, $OutSoure_{c,t}$ (see Section 4.1.3). We then separate establishments into three subsamples each year based on the terciles of their counties' $OutSoure_{c,t}$. Panels A-C plot the RegIndex by establishment size bins in the low, medium, and high $OutSoure_{c,t}$ counties, respectively. See Section 3.3 for the definition of the RegIndex measure. The dots represent the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

Panel A: Establishments from Low Outsourcing Counties



Panel B: Establishments from Medium Outsourcing Counties



Panel C: Establishments from High Outsourcing Counties

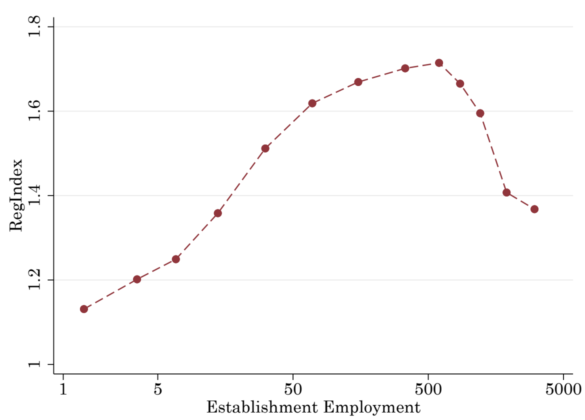


Figure A.10: **Outsourcing Intensity of Regulatory Compliance and Size**

This figure plots the estimated outsourcing intensity for regulatory compliance by establishments' size in Section 4.1.3. The estimates are produced by regressing the county-level outsourcing services abundance, $OutSource_{c,t}$, on establishments' size distribution within the county based on the 14 size bins, $Share_{j,c,t}$. The regression does not impose a constant and controls for the county's total employment, state and urban fixed effects, and year fixed effects. The estimate for the largest size bin is not used in our study and is excluded from this figure due to estimation noise (see details in Section 4.1.3). The maroon line with diamond dots shows the estimates where $OutSource_{c,t}$ is computed based on the county's abundance of small outsourcing service providers defined as those with 20 or fewer employees (see Figure 9). The shaded areas represent the 95% confidence intervals of the estimates. The x-axis is on a log scale.

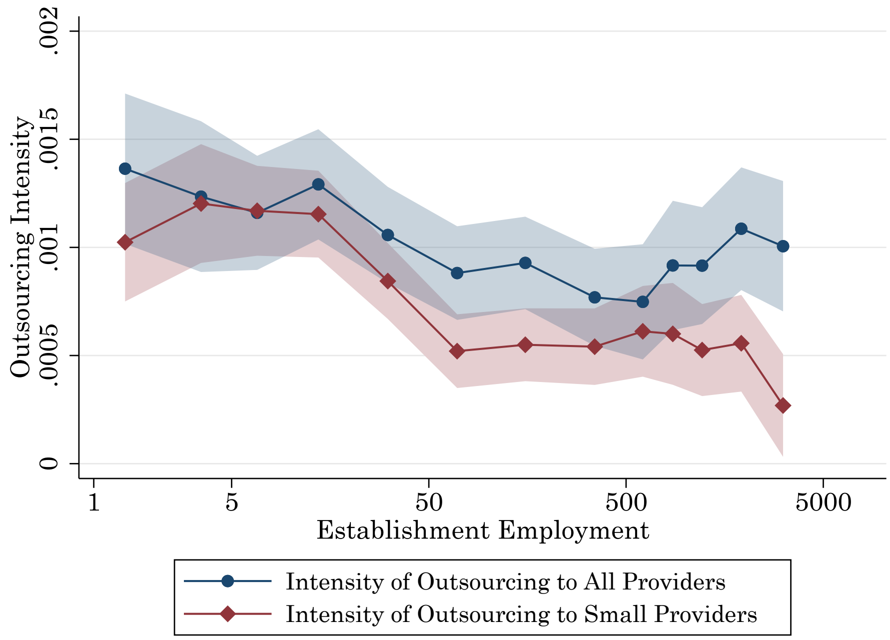


Figure A.11: **Robustness: Outsource-Adjusted RegIndex and Size (Alternative Cutoffs for Defining Small Outsourcing Providers)**

This figure plots the robustness check results for Figure 9 regarding the relation of in-house-based RegIndex (baseline), two versions of outsource-adjusted RegIndex and employment for firms in Panels A and C and establishments in Panels B and D. Panels A and B show to results of adding compliance costs paid alternatively-defined small outsourcing firms with 10 or fewer employees (instead of 20 or fewer in Figure 9), whereas panels C and D show the results for small outsourcing firms defined as by 50 or fewer employees. Each panel reports two versions of outsource-adjusted RegIndex constructed assuming that firms use compliance outsourcing services from local providers within the county and from nationwide providers, respectively. Section 4.1.3 provides the detailed definitions and estimation procedures. See Section 3.3 for the definition of the baseline RegIndex. The dots represent the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

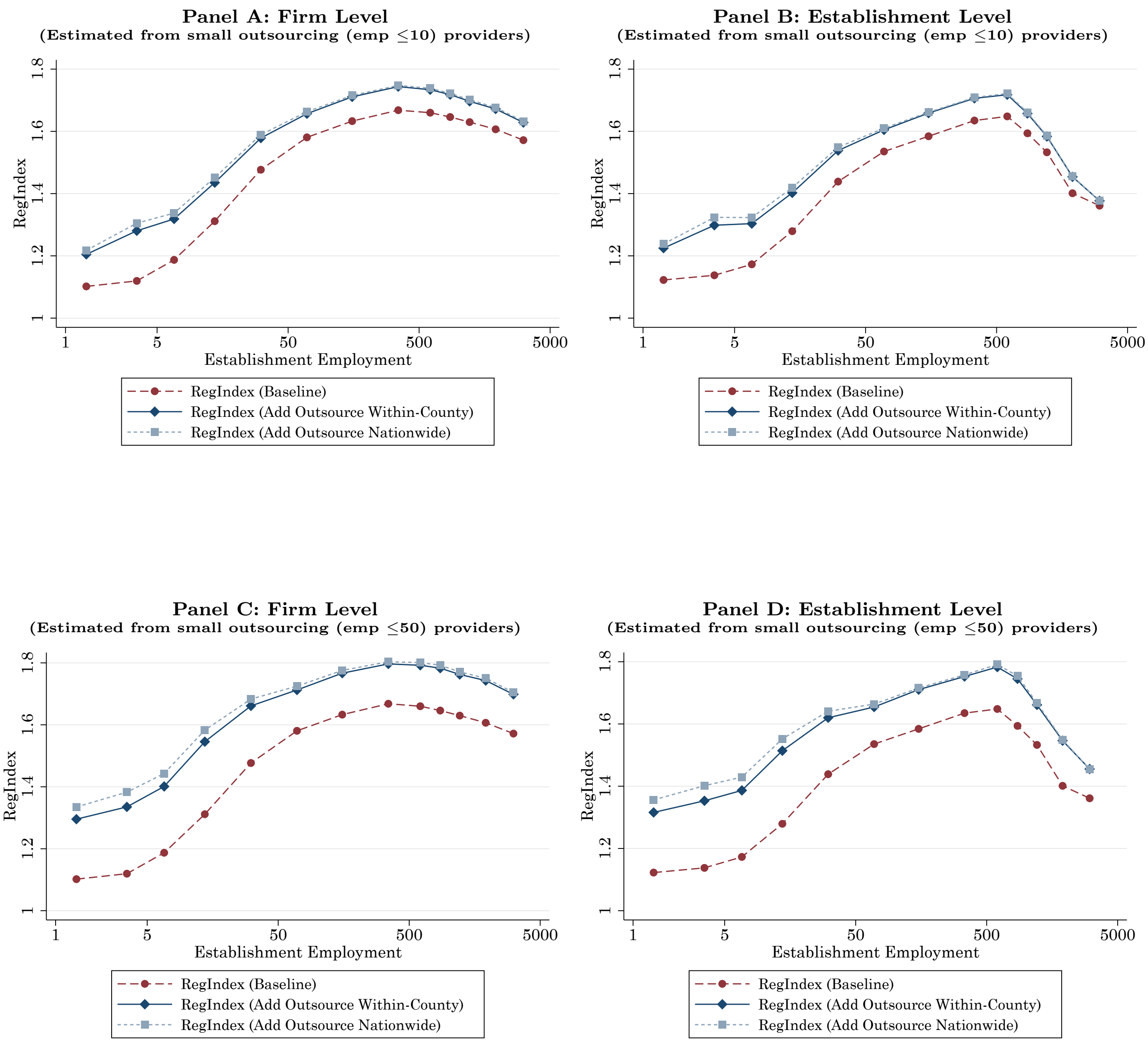


Figure A.12: **Robustness: Outsource-Adjusted RegIndex and Size (Medium and Broad Versions of RegIndex)**

This figure plots the robustness check results for Figure 9 regarding the relation of in-house-based Medium and Broad version RegIndex (baseline), two versions of outsource-adjusted RegIndex and employment for firms in Panels A and C and establishments in Panels B and D. See Section 3.3 for the definition of the versions of the RegIndex measure. In all panels, we include compliance costs paid to small outsourcing firms with 20 or fewer employees. In Panels A and B, we rerun the co-location estimation using the medium version of RegIndex throughout, while in Panels C and D, we rerun the estimation using the broad version of RegIndex throughout. Each panel reports two versions of outsource-adjusted RegIndex constructed assuming that firms use compliance outsourcing services from local providers within the county and from nationwide providers, respectively. Section 4.1.3 provides the detailed definitions and estimation procedures. The dots represent the average RegIndex in each employment bin, where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale.

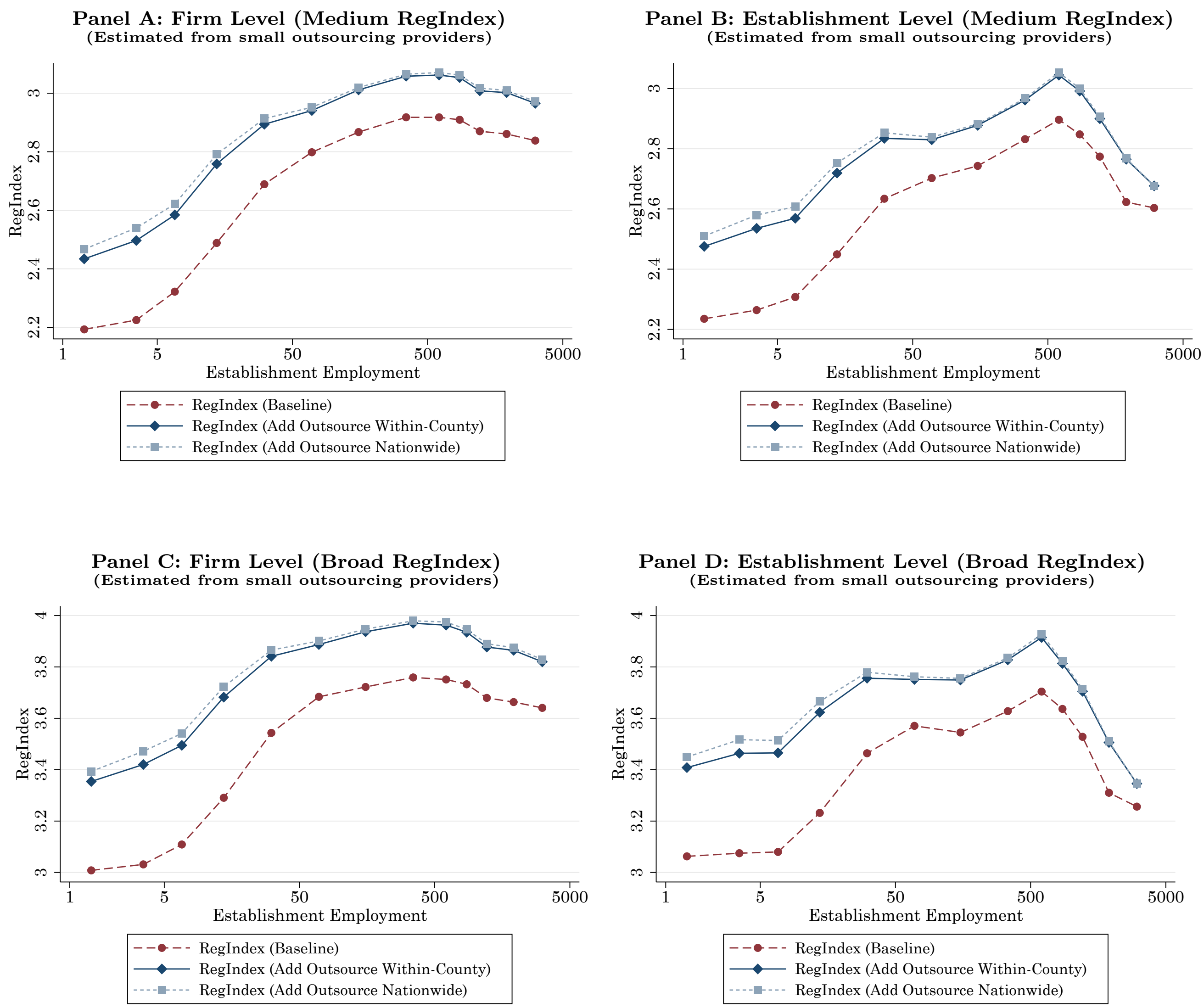


Figure A.13: **Changes in Regulatory Requirements and Enforcement Measures by Agency**

This figure plots the changes in an agency's regulatory enforcement and regulation requirements. Enforcement changes (*ENF*) are measured by the 3-year changes in the natural logarithm of the agency's regulation-related employment from t to $t+3$. Regulation-requirement changes (*REG*) are measured by the 3-year changes in the natural logarithm of the agency's estimated compliance hours of its regulations excluding adjustments from t to $t+3$. See Section 5 for more details.

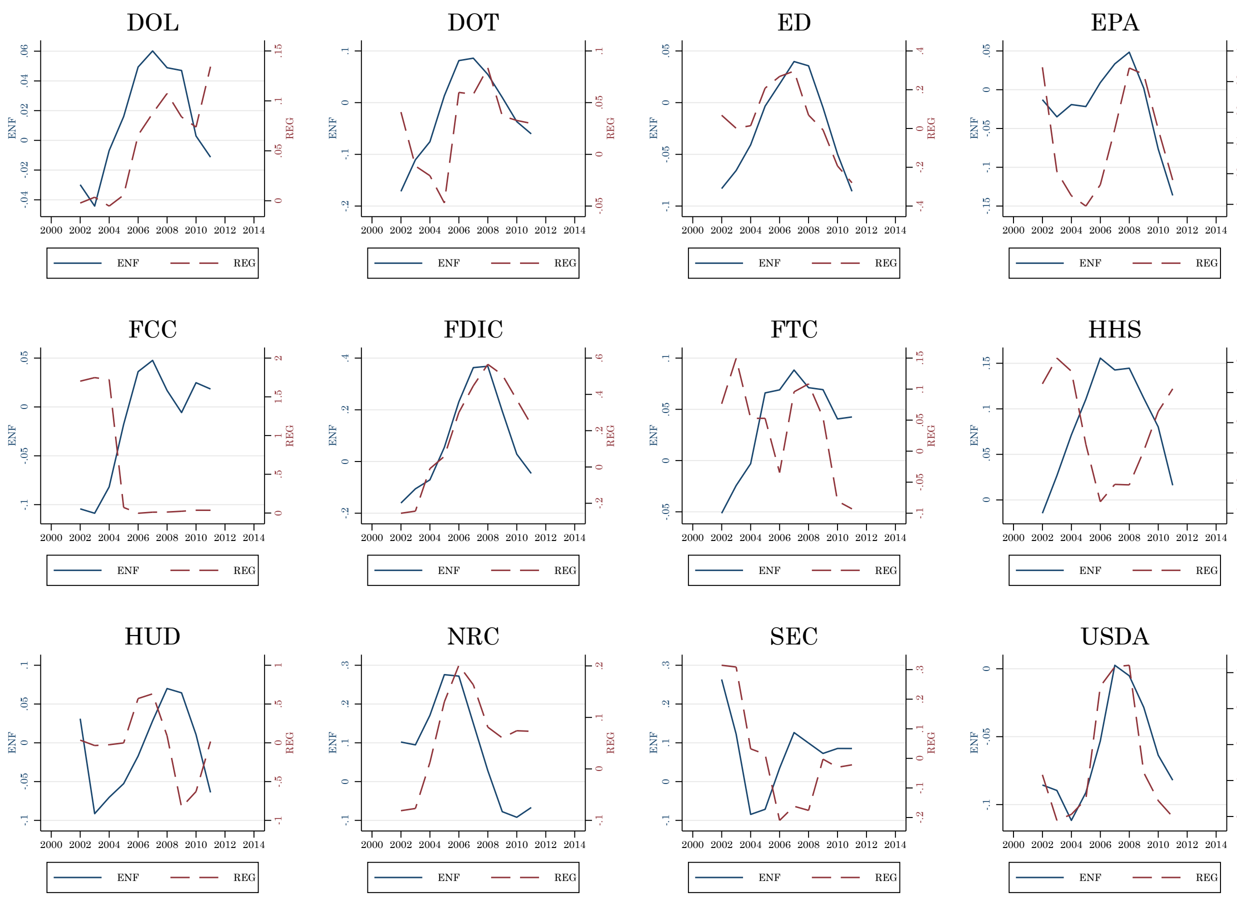


Figure A.14: **Levels of Regulatory Requirements vs. Enforcement Measures by Agency**

This figure plots each major agency's regulation-related employment, which is used to construct enforcement changes, and the estimated compliance hours (in millions) of the agency's regulations excluding adjustments, which are used to construct regulation-requirement changes. See Section 5 for more details.

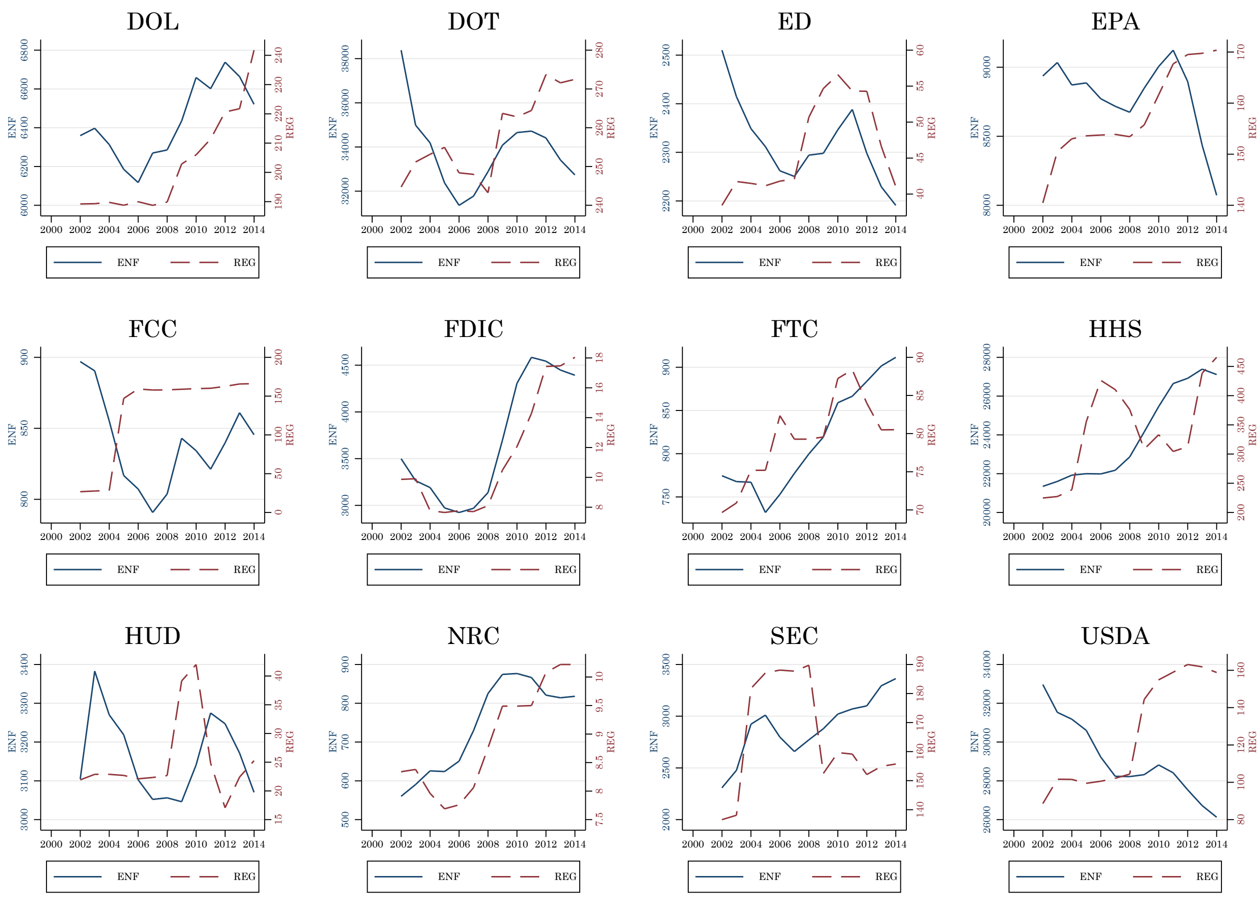


Figure A.15: **Changes in Firm Size Distribution in the U.S.**

This figure illustrates the relationship between RegIndex and changes in businesses' employment size distribution in the U.S. from 2002 to 2014. In Panel A, the navy line shows the change in firms' size distribution from 2002 to 2014 for each size bin. The maroon dots, which correspond to the right y-axis, represent the average RegIndex in each employment bin of firms (EIN), where the bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. The x-axis is on a log scale. In Panel B, we show a similar figure for establishments. See Section 5.3 for more details.

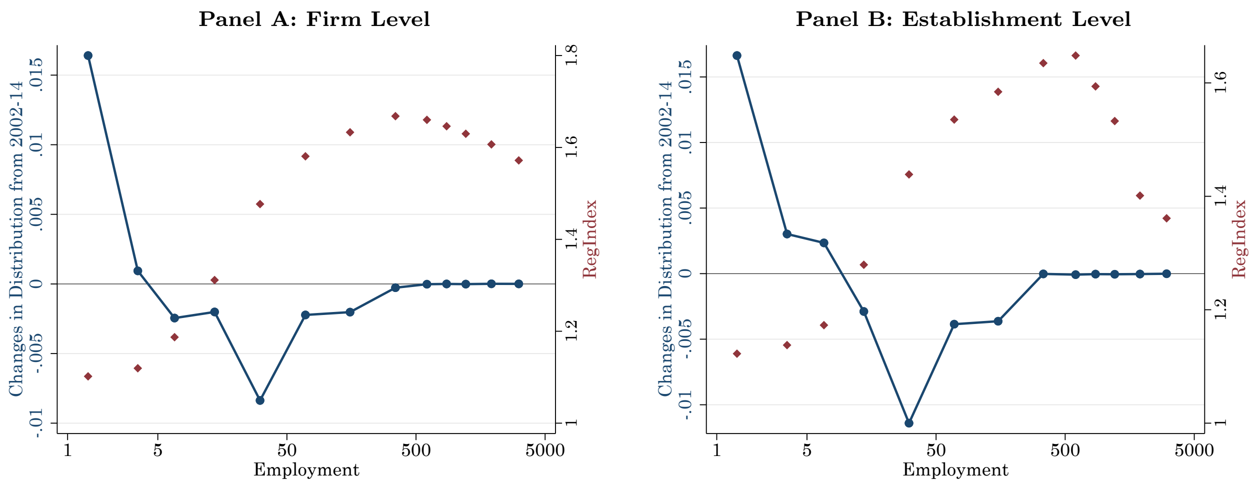


Table A.1: **Examples of Regulation-related Tasks**

This table lists 10 regulation-related tasks from the O*NET database. See Section 3 for our definition of regulation-related tasks. *Import.* is the importance rating of the task for the occupation ranging from 1 to 5 provided by the O*NET database.

Occupation and Task	Import.
Construction Managers	
Inspect or review projects to monitor compliance with building and safety codes or other regulations.	3.91
Agricultural Inspectors	
Inspect agricultural commodities or related operations, as well as fish or logging operations, for compliance with laws and regulations governing health, quality, and safety.	4.59
Construction and Building Inspectors	
Evaluate project details to ensure adherence to environmental regulations.	4.12
Financial Examiners	
Establish guidelines for procedures and policies that comply with new and revised regulations and direct their implementation.	3.69
Industrial Engineering Technologists and Technicians	
Monitor environmental management systems for compliance with environmental policies, programs, or regulations.	2.67
Occupational Health and Safety Specialists	
Inspect or evaluate workplace environments, equipment, or practices to ensure compliance with safety standards and government regulations.	4.21
Urban and Regional Planners	
Determine the effects of regulatory limitations on land use projects.	4.00
Aircraft Mechanics and Service Technicians	
Conduct routine and special inspections as required by regulations.	4.49
Food Service Managers	
Monitor compliance with health and fire regulations regarding food preparation and serving, and building maintenance in lodging and dining facilities.	4.45
Compensation and Benefits Managers	
Fulfill all reporting requirements of all relevant government rules and regulations, including the Employee Retirement Income Security Act (ERISA).	4.35

Table A.2: **Top 25 Occupations with the Highest Regulation-Task Intensity**

This table reports the top 25 SOC 6-digit occupations with the highest regulation-task intensity (RTI). See Section 3 for the construction of RTI.

SOC	Occupation Title	RTI
13-1041	Compliance Officers	0.343
47-4011	Construction and Building Inspectors	0.340
45-2011	Agricultural Inspectors	0.278
17-3026	Industrial Engineering Technicians	0.262
13-2061	Financial Examiners	0.256
19-3051	Urban and Regional Planners	0.229
33-2021	Fire Inspectors and Investigators	0.223
23-1011	Lawyers	0.204
17-2081	Environmental Engineers	0.189
19-2041	Environmental Scientists and Specialists, Including Health	0.183
19-3011	Economists	0.180
19-1012	Food Scientists and Technologists	0.176
43-4031	Court, Municipal, and License Clerks	0.156
33-1021	First-Line Supervisors of Fire Fighting and Prevention Workers	0.154
17-2111	Health and Safety Engineers, Ex. Mining Safety Engineers & Inspectors	0.152
53-6051	Transportation Inspectors	0.140
35-1011	Chefs and Head Cooks	0.134
19-3094	Political Scientists	0.132
13-2082	Tax Preparers	0.130
29-9012	Occupational Health and Safety Technicians	0.129
33-3051	Police and Sheriff's Patrol Officers	0.121
33-9091	Crossing Guards	0.119
11-9151	Social and Community Service Managers	0.119
11-9021	Construction Managers	0.117
33-3041	Parking Enforcement Workers	0.117

Table A.3: **Top 25 Industries with the Highest Regulation Index**

This table reports the top 25 NAICS 3-digit industries with the highest regulation-index (RegIndex). See Section 3 for the construction of RegIndex.

NAICS	Industry Title	RegIndex (%)
485	Transit and Ground Passenger Transportation	3.930
525	Funds, Trusts, and Other Financial Vehicles	3.359
325	Chemical Manufacturing	3.274
324	Petroleum and Coal Products Manufacturing	2.992
551	Management of Companies and Enterprises	2.882
523	Securities, Commodity Contracts, and Other Financial Related Activities	2.734
221	Utilities	2.733
211	Oil and Gas Extraction	2.705
483	Water Transportation	2.628
236	Construction of Buildings	2.624
334	Computer and Electronic Product Manufacturing	2.621
486	Pipeline Transportation	2.594
813	Religious, Grantmaking, Civic, Professional, and Similar Organizations	2.565
238	Specialty Trade Contractors	2.511
336	Transportation Equipment Manufacturing	2.470
533	Lessors of Nonfinancial Intangible Assets (except Copyrighted Works)	2.460
237	Heavy and Civil Engineering Construction	2.452
531	Real Estate	2.430
522	Credit Intermediation and Related Activities	2.414
482	Rail Transportation	2.230
327	Nonmetallic Mineral Product Manufacturing	2.180
212	Mining (except Oil and Gas)	2.164
313	Textile Mills	2.153
326	Plastics and Rubber Products Manufacturing	2.139
562	Waste Management and Remediation Services	2.114

Table A.4: **Validation: Case Studies of Industry Regulatory Changes**

This table reports the response of establishments' RegIndex to major industry regulatory changes in five case studies. Section 3 provides the details of each case study. *Treated* is a dummy variable that equals 1 if the industry is treated by the regulation and 0 if not. *Post* is a dummy variable that equals 1 if the year is after the law enactment year and 0 if before. We exclude the law enactment year. All standard errors are double clustered by year and NAICS 6-digit industry. Each observation is weighted by a product of the establishment's weight assigned by the OEWS survey and the establishment's total annual wage payment. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Treated Industry: Case:	Credit Cards CARDAct 2009 (1)	Oil and Gas EPAct 2005 (2)	Oil and Gas Executive Order 2010 (3)	Pharmaceutical FDA Guidance 2008 (4)	Hospitals ACA 2010 (5)
Treated × Post	0.423***	-0.521***	0.569***	0.272**	0.062***
	(0.116)	(0.090)	(0.153)	(0.110)	(0.006)
Treated	0.257	1.413**	-1.839***	-0.245	0.606***
	(0.195)	(0.419)	(0.330)	(0.186)	(0.004)
Year FE	Yes	Yes	Yes	Yes	Yes
Establishment FE	Yes	Yes	Yes	Yes	Yes
Observations	10,082	3,140	5,877	13,005	25,043
Adjusted R^2	0.364	0.441	0.267	0.327	0.398

Table A.5: **Robustness—Validation: State Vote Share for Republican Party and RegIndex**

This table reports the robustness check of Table 2 by controlling for states' 2017 RegData measure, and states' number of establishments in 2017. See Figure 4 for the estimation of states' 2014 RegIndex. State RegData is from QuantGov.org. States' number of establishment is from the Census Statistics of U.S. Businesses (SUSB).

	State Vote Share for Republican Party in 2016 Presidential Election		State Vote Share for Republican Party in 2016 House Delegation Elections		State Vote Share for Republican Party in 2017-18 Senate Elections	
	(1)	(2)	(3)	(4)	(5)	(6)
State RegIndex	-0.538***	-0.539***	-0.958***	-0.957***	-1.835***	-1.817***
	(0.109)	(0.109)	(0.253)	(0.254)	(0.468)	(0.478)
State RegData	-0.561***	-0.512**	-0.547***	-0.607**	-2.109***	-2.978**
	(0.102)	(0.198)	(0.137)	(0.271)	(0.597)	(1.133)
#Establishments		-0.026		0.032		0.457
		(0.082)		(0.106)		(0.472)
Constant	1.414***	1.413***	2.089***	2.090***	3.693***	3.708***
	(0.174)	(0.175)	(0.390)	(0.395)	(0.748)	(0.776)
Observations	49	49	49	49	49	49
Adjusted R^2	0.339	0.325	0.365	0.352	0.218	0.214

Table A.6: **Robustness: RegIndex Before and After IPO**

This table reports the robustness of Table 3 by showing all three versions of RegIndex before and after the firm go for IPO from $t-3$ to $t+3$ where year $t-1$ is omitted as the benchmark year. The mean RegIndex for the three conservative, medium, and broad versions of IPO firms at $t-1$ are 1.798, 3.309 and 4.410, respectively. See Appendix A for the definitions of our RegIndex versions. All standard errors are clustered by firm. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Version:	Conservative		Medium		Broad	
	RegIndex (1)	Adj.RegIndex (2)	RegIndex (3)	Adj.RegIndex (4)	RegIndex (5)	Adj.RegIndex (6)
IPO_{t-3}	0.047	0.059	0.075	0.083	0.139	0.118
	(0.064)	(0.063)	(0.099)	(0.097)	(0.130)	(0.129)
IPO_{t-2}	-0.017	-0.003	-0.010	0.001	0.017	0.013
	(0.037)	(0.036)	(0.063)	(0.061)	(0.085)	(0.082)
IPO_{t}	0.045	0.039	0.046	0.040	0.055	0.059
	(0.031)	(0.031)	(0.047)	(0.047)	(0.062)	(0.062)
IPO_{t+1}	0.135***	0.108**	0.222***	0.181***	0.272***	0.243***
	(0.047)	(0.046)	(0.073)	(0.070)	(0.093)	(0.089)
IPO_{t+2}	0.151***	0.111**	0.222***	0.179**	0.283***	0.264**
	(0.052)	(0.051)	(0.081)	(0.078)	(0.107)	(0.104)
IPO_{t+3}	0.153***	0.099*	0.204**	0.149*	0.279**	0.255**
	(0.053)	(0.051)	(0.082)	(0.080)	(0.112)	(0.110)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,164	5,164	5,164	5,164	5,164	5,164
Adjusted R^2	0.701	0.621	0.699	0.611	0.717	0.629

Table A.7: **Correlation Between RegIndex and Other Firm-Level Regulation Measures**

This table reports the pairwise correlations between our RegIndex measures and two popular measures of publicly-traded firms' regulation intensity. See Appendix A for the definitions of our RegIndex versions. The measure from Kalmenovitz (2023) is the firm-level RegIn based on firms' exposure to the estimated total compliance hours for the paperwork regulations. The measure from Armstrong et al. (2025) is the firm-level total government agency exposure based on firms mentions of government agencies in their 10-K filings. We thank the authors for making their measures publicly available. We follow Armstrong et al. (2025) and select publicly-traded firms to be incorporated in the U.S. and have positive revenue in the year from 2002 to 2014. See Section 3.5.4 for more discussions.

	RegIndex (Conservative)	RegIndex (Medium)	RegIndex (Broad)
Kalmenovitz (2023)	0.165***	0.182***	0.173***
Armstrong et al. (2025)	0.240***	0.281***	0.251***

Table A.8: **Examples of Requirement-Sensitive and Enforcement-Sensitive Tasks**

This table reports examples of regulation requirement-sensitive tasks in Panel A and enforcement-sensitive tasks in Panel B for regulation-related tasks. See details about the categorization methodology in Section 5.2.

Task ID	Task	Occupation
	Panel A: Regulation Requirement-Sensitive Tasks	
10840	Research and keep informed of pertinent information and developments in areas such as EPA laws and regulations.	Environmental Compliance Inspectors
2522	Keep informed about changes in tax and deduction laws that apply to the payroll process.	Payroll and Timekeeping Clerks
3371	Consult with or serve as a technical liaison between business, industry, government, and union officials.	Job Analysts
3774	Interpret laws, rulings and regulations for individuals and businesses.	Lawyers
18058	Maintain current knowledge base of existing and emerging regulations, standards, or guidance documents.	Regulatory Affairs Specialists
9849	Learn and follow safety regulations.	Operating Engineers and Other Construction Equipment Operators
	Panel B: Regulation Enforcement-Sensitive Tasks	
11067	Inspect or evaluate workplace environments, equipment, or practices to ensure compliance with safety standards and government regulations.	Occupational Health and Safety Specialists
7239	Confer with legal authorities to ensure that renting and advertising practices are not discriminatory and that properties comply with state and federal regulations.	Property, Real Estate, and Community Association Managers
14570	Inspect vehicles or equipment to ensure compliance with rules, standards, or regulations.	Transportation Vehicle, Equipment and Systems Inspectors, Except Aviation
7230	Maintain contact with insurance carriers, fire and police departments, and other agencies to ensure protection and compliance with codes and regulations.	Property, Real Estate, and Community Association Managers
10825	Prepare reports required by state and federal laws.	Aquacultural Managers
12820	Explain and enforce safety regulations and policies.	Nursery and Greenhouse Managers

Table A.9: **50 Keywords for Each Regulatory Agency**

This table lists the 50 keywords we used to identify each regulatory agency. To obtain these keywords, we first extract all relevant volumes from the Code of Federal Regulations to each of the 12 agencies in Section 5. Then, we compute the term-frequency ratio for each word as the count of the word in the agency i's relevant CFR volumes over the count of that word in all 12 agencies' relevant CFR volume. This table lists the top 50 keywords with the highest term-frequency ratio for each agency.

Rank	DOL	DOT	ED	EPA	FCC	FDIC
1	labors	transported	education	environmental	communications	deposit
2	workers	traveling	educational	pollution	telecommunication	depositor
3	unemployment	transit	school	epa	transmitting	depositors
4	employer	cargo	schools	environment	telecommunications	depository
5	workforce	freight	learners	conservation	transmit	deposits
6	worker	trains	academic	epas	broadcasting	fdic
7	wages	traffic	teachers	ecological	channels	fdics
8	jobs	taxiing	diploma	emissions	radiocommunication	bank
9	employers	buses	teacher	pollutants	broadcasts	banks
10	workplaces	cargocarrying	student	eco	telephony	banking
11	employees	intercity	colleges	contamination	broadcast	insured
12	bargaining	passengers	instructional	renewable	reception	savings
13	wage	bus	literacy	recycling	channel	investments
14	job	vehicle	graduate	pollutant	conversation	dividend
15	employee	highways	students	ecosystem	signals	fdi
16	miners	driving	parents	contaminated	multichannel	fdicsupervised
17	workmens	railroads	teachout	endangerment	telegraphy	forex
18	machinery	train	vocational	ecology	transmissions	paycheck
19	occupation	taxi	graduation	chlorinated	networks	eximbank
20	subsistence	locomotive	curricula	preventable	telephones	fsi
21	welders	cargoes	tuition	greenhouse	radiotelephony	pd
22	unemployed	railroad	postsecondary	preventative	broadcaster	unfunded
23	farmworkers	cruising	undergraduate	pesticide	radio	surcharge
24	wageloss	cars	institution	pesticides	transmitters	ssfa
25	demanding	roadside	baccalaureate	chemicals	wireless	loans
26	workweek	trips	parental	cleaner	broadband	mortgagebacked
27	workday	haul	elementary	hazardous	signalling	portfolio
28	jobrelated	roadway	extracurricular	compliance	fcc	collateral
29	workrelated	passengercarrying	faculty	ordinance	modulation	unsecured
30	workdays	taxiway	achievement	contaminants	transmitter	portfolios
31	workings	towing	semester	ecosystems	interference	institutionaffiliated
32	surplus	commuter	schoolwide	habitats	cochannel	securitization
33	cutting	baggage	mathematics	containment	transceiver	loantovalue
34	employmentrelated	car	coursework	cleaners	telecommand	lending
35	recruitment	ferry	bachelors	remediation	cable	securitizations
36	welder	passenger	campus	ozone	television	brokered
37	occupations	flight	preschool	aeration	broadcastingsatellite	fdia
38	worksites	itinerary	children	wetlands	bandwidths	qfc
39	layoff	luggage	cognitive	warming	audio	dif
40	apprenticeship	congestion	enrolled	decontaminated	rf	securitized
41	contracture	fares	talent	recycled	bandwidth	safekeeping
42	erecting	highway	geography	wastewaters	decoders	gaap
43	shafting	carriage	doctoral	permits	fccs	fiduciary
44	contractorissued	routes	racial	petroleum	voip	liquidity
45	farmworker	route	childs	wastewater	messages	creditworthiness
46	jobsite	drivertrainees	scholar	biocides	stations	institution
47	economical	riding	accrediting	landfills	handsets	statelicensed
48	men	movement	athletic	wildlife	interconnected	assets
49	sickness	aboard	peer	antidegradation	interconnection	lei
50	clothing	airline	disabilities	epadc	antenna	pledged

Table A.9: **50 Keywords for Each Regulatory Agency**—*Continued*

This table lists the 50 keywords we used to identify each regulatory agency. To obtain these keywords, we first extract all relevant volumes from the Code of Federal Regulations to each of the 12 agencies in Section 5. Then, we compute the term-frequency ratio for each word as the count of the word in the agency i's relevant CFR volumes over the count of that word in all 12 agencies' relevant CFR volume. This table lists the top 50 keywords with the highest term-frequency ratio for each agency.

Rank	FTC	HHS	HUD	NRC	SEC	USDA
1	seller	health	housing	nuclear	securities	livestock
2	sellers	hospitals	dwellings	reactorrelated	brokers	grazing
3	buyers	medicine	residential	reactor	brokerdealers	tomatoes
4	marketer	healthrelated	redevelopment	reactors	securitiesxexxd	growers
5	advertised	doctors	neighborhoods	radioactive	broker	seedlings
6	gp	physician	homes	fission	investor	potatoes
7	franchised	hospital	cities	uranium	brokerdealer	seeds
8	marketers	hospitalspecific	apartment	neutron	investors	organically
9	valued	physicians	rents	plutonium	trader	grower
10	merchandise	ambulance	neighborhood	isotope	shareholders	potato
11	solicitations	inpatients	tenancy	irradiation	shareholder	berries
12	acquisitions	patients	renting	atomic	currencies	germination
13	clothes	clinicians	dwelling	tritium	accountant	varieties
14	opt	manpower	homeowners	radiation	prospectus	variety
15	franchise	hospitalization	rent	radioactivity	accountants	weed
16	pearl	nurse	condominium	irradiator	prospectuses	peanuts
17	wholesalers	hipaa	bedrooms	radionuclides	stockholder	apples
18	camera	diseases	reside	deuterium	offerings	seedling
19	apparel	clinics	households	isotopes	syndicate	pear
20	warrantor	hospices	homebuyers	radionuclide	advisers	grapes
21	advertiser	professions	tenants	irradiated	depositor	pears
22	diamond	inpatient	homelessness	strontium	prudential	seed
23	paypercall	doctor	mortgages	irradiators	adviser	almonds
24	advertisement	medically	residents	nrc	securityholder	cotton
25	octane	nursing	homeownership	radiological	edgar	leaf
26	deception	clinical	mortgagees	fissile	underwriter	upland
27	fur	clinic	amenities	securityrelated	offering	usda
28	rvalue	hospitalbased	homeowner	strategic	interdealer	pork
29	optout	profession	homeless	thorium	dividends	cottonseed
30	franchisee	aides	poverty	nrcs	intermediarys	rot
31	furs	telemedicine	homebuyer	technetium	futures	goat
32	advertisers	medicare	rental	unrestricted	securitybased	ripe
33	franchisees	hhs	shelter	doenrc	depositary	flesh
34	abc	clinician	shelters	licenses	registrants	clover
35	ftc	fdas	mortgaged	license	dealers	seedless
36	telemarketing	hmos	developments	fsar	promoter	insects
37	unfair	medicaid	landlord	commissionapproved	nms	onions
38	wool	practitioners	mortgage	gamma	counterpartys	roots
39	biomassbased	patient	architect	snm	counterparties	tobacco
40	recyclable	stewardship	household	rulemakings	reliance	cherries
41	consumers	care	incomes	safeguards	fasb	raisins
42	imitation	biomedical	restructuring	physicist	dealer	stems
43	conveys	drugs	builder	licensee	repurchase	apple
44	telemarketer	hospice	modernization	licensees	penny	insect
45	textile	caregivers	occupancy	byproduct	diversified	olives
46	furnisher	shortage	tenant	engineered	soliciting	aggregating
47	franchisor	servings	buildingcomplex	enrichment	sx	kernel
48	freezer	disease	vacant	coc	crs	spready
49	rayon	interventions	occupy	repository	intercompany	lamb
50	conditioners	surgeons	vacancies	nb	ob	dirty

Table A.10: **Top 3 Industries for Each Regulatory Agency**

This table reports each regulatory agency's top 3 most exposed industries, where industry is defined at the NAICS 3-digit level. r_k is the ratio of the industry's labor spending on agency k's regulation-related tasks and total labor spending. RegIndex is the ratio of the industry's labor spending on all regulation-related tasks and total labor spending. See Section 5 for details.

Agency	Rank	NAICS3	Title	r_k/RegIndex
USDA	1	722	Food Services and Drinking Places	0.0707
USDA	2	311	Food Manufacturing	0.0634
USDA	3	115	Support Activities for Agriculture and Forestry	0.0621
DOT	1	492	Couriers and Messengers	0.2298
DOT	2	485	Transit and Ground Passenger Transportation	0.2196
DOT	3	482	Rail Transportation	0.1721
EPA	1	562	Waste Management and Remediation Services	0.1775
EPA	2	324	Petroleum and Coal Products Manufacturing	0.1736
EPA	3	236	Construction of Buildings	0.1651
FCC	1	512	Motion Picture and Sound Recording Industries	0.1220
FCC	2	492	Couriers and Messengers	0.1064
FCC	3	515	Broadcasting (except Internet)	0.1054
FDIC	1	523	Securities, Commodity Contracts, and Other Financial Activities	0.1738
FDIC	2	522	Credit Intermediation and Related Activities	0.1726
FDIC	3	525	Funds, Trusts, and Other Financial Vehicles	0.1490
HHS	1	446	Health and Personal Care Stores	0.2442
HHS	2	621	Ambulatory Health Care Services	0.2400
HHS	3	622	Hospitals	0.2166
HUD	1	531	Real Estate	0.1808
HUD	2	236	Construction of Buildings	0.1562
HUD	3	238	Specialty Trade Contractors	0.1361
FTC	1	313	Textile Mills	0.1213
FTC	2	315	Apparel Manufacturing	0.1205
FTC	3	314	Textile Product Mills	0.1176
NRC	1	221	Utilities	0.1041
NRC	2	325	Chemical Manufacturing	0.1006
NRC	3	562	Waste Management and Remediation Services	0.0928
ED	1	624	Social Assistance	0.1172
ED	2	485	Transit and Ground Passenger Transportation	0.1107
ED	3	713	Amusement, Gambling, and Recreation Industries	0.1050
DOL	1	113	Forestry and Logging	0.1487
DOL	2	813	Religious, Grant-making, Civic, Professional, Similar Organizations	0.1480
DOL	3	448	Clothing and Clothing Accessories Stores	0.1477
SEC	1	523	Securities, Commodity Contracts, and Other Financial Activities	0.1963
SEC	2	522	Credit Intermediation and Related Activities	0.1786
SEC	3	525	Funds, Trusts, and Other Financial Vehicles	0.1638

Table A.11: **Correlation of Shift-Share Variables**

This table reports the Pearson correlation between the shift-share variable for enforcement changes, $iv(\Delta\log(p_{it}))$, and the shift-share variable for regulatory-requirement changes, $iv(\Delta\log(\tilde{R}_{it}))$, in full sample and in each NAICS 1-digit sector. The full firm-level sample includes 608,500 observations that have 3-year changes in log RegIndex. The full establishment-level sample includes 628,733 observations that have 3-year changes in log RegIndex. Section 5 provides more details on the construction of the shift-share variables.

	Firm-Level Sample	Establishment-Level Sample
All Sectors	0.104	0.116
Agriculture, Mining, and Construction	0.159	0.157
Manufacturing	0.007	-0.084
Retail	-0.183	-0.219
Wholesale	-0.127	-0.131
Utilities	0.024	-0.049
Transportation	-0.092	-0.139
Finance	0.167	0.139
Service	0.077	0.123

Table A.12: **Regulatory Requirements vs. Enforcement for Firms in Each Sector**

This table reports the results of regressing 3-year changes in a firm's log RegIndex on shift-share variables in Table 6 in each NAICS 1-digit sector. Equation (7) provides the regression specification. See Table 6 for variable definitions. All regressions control for $\Delta\log(\text{Wage})$, $\Delta\log(w^r)$, log(RegIndex), and year fixed effects. All variables are standardized to have mean 0 and variance of 1 for ease of interpretation. Standard errors are double clustered by industry and year and presented in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. The sample period is from 2002 to 2014.

	All Sizes			Subsample by Firm Size			
				1-19	20-399	400-749	$\geq$ 750
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Agriculture, Mining, and Construction			
$iv(\Delta\log(p_{it}))$	0.190**		0.106	0.103	0.106	0.068	0.213**
	(0.078)		(0.078)	(0.081)	(0.075)	(0.103)	(0.082)
$iv(\Delta\log(\tilde{R}_{it}))$		0.269***	0.243***	0.196**	0.276***	0.279**	0.207*
		(0.044)	(0.059)	(0.061)	(0.059)	(0.094)	(0.103)
				Manufacturing			
$iv(\Delta\log(p_{it}))$	0.349***		0.183*	0.120	0.190*	0.243**	0.234*
	(0.087)		(0.087)	(0.068)	(0.094)	(0.104)	(0.103)
$iv(\Delta\log(\tilde{R}_{it}))$		0.331***	0.285***	0.241***	0.307***	0.323***	0.411***
		(0.057)	(0.047)	(0.044)	(0.052)	(0.048)	(0.043)
				Retail			
$iv(\Delta\log(p_{it}))$	0.132***		0.097**	0.088*	0.127**	0.062	0.033
	(0.040)		(0.033)	(0.039)	(0.044)	(0.047)	(0.038)
$iv(\Delta\log(\tilde{R}_{it}))$		0.134***	0.120***	0.122***	0.128***	0.099**	0.103***
		(0.028)	(0.025)	(0.022)	(0.032)	(0.034)	(0.022)
				Wholesale			
$iv(\Delta\log(p_{it}))$	0.214***		0.125**	0.076*	0.164**	0.096	0.200
	(0.057)		(0.042)	(0.034)	(0.059)	(0.103)	(0.110)
$iv(\Delta\log(\tilde{R}_{it}))$		0.194***	0.169***	0.125***	0.194***	0.285***	0.272***
		(0.038)	(0.032)	(0.035)	(0.039)	(0.074)	(0.070)

Table A.12: **Regulatory Requirements vs. Enforcement for Firms by Sector** –*Continued*

This table reports the results of regressing 3-year changes in a firm's log RegIndex on shift-share variables in Table 6 in each NAICS 1-digit sector. Equation (7) provides the regression specification. See Table 6 for variable definitions. All regressions control for Δlog(Wage), $\Delta\log(w^r)$, log(RegIndex), and year fixed effects. All variables are standardized to have mean 0 and variance of 1 for ease of interpretation. Standard errors are double clustered by industry and year and presented in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. The sample period is from 2002 to 2014.

	All Sizes			Subsample by Firm Size			
				1-19	20-399	400-749	$\geq$ 750
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Utilities							
$iv(\Delta\log(p_{it}))$	0.435***		0.196*	0.123*	0.210	-0.053	0.140
	(0.127)		(0.103)	(0.055)	(0.167)	(0.207)	(0.116)
$iv(\Delta\log(\tilde{R}_{it}))$		0.443***	0.373**	0.544***	0.254	0.539***	0.378***
		(0.130)	(0.120)	(0.120)	(0.201)	(0.044)	(0.046)
Transportation							
$iv(\Delta\log(p_{it}))$	0.167		0.100	0.059	0.105	0.129*	0.184**
	(0.101)		(0.067)	(0.062)	(0.071)	(0.058)	(0.071)
$iv(\Delta\log(\tilde{R}_{it}))$		0.374***	0.360***	0.310***	0.367***	0.448***	0.514***
		(0.037)	(0.030)	(0.037)	(0.028)	(0.065)	(0.058)
Finance							
$iv(\Delta\log(p_{it}))$	0.380***		0.190**	0.191*	0.192**	0.211	0.088
	(0.053)		(0.075)	(0.093)	(0.080)	(0.143)	(0.108)
$iv(\Delta\log(\tilde{R}_{it}))$		0.304***	0.239**	0.206*	0.275***	0.222***	0.336**
		(0.067)	(0.079)	(0.093)	(0.067)	(0.061)	(0.112)
Service							
$iv(\Delta\log(p_{it}))$	0.196**		0.097	0.101	0.096	0.055	0.034
	(0.078)		(0.082)	(0.081)	(0.081)	(0.066)	(0.062)
$iv(\Delta\log(\tilde{R}_{it}))$		0.251***	0.222***	0.169**	0.258***	0.269***	0.279***
		(0.040)	(0.058)	(0.058)	(0.057)	(0.045)	(0.047)

Table A.13: **Regulatory Requirements vs. Enforcement for Establishments by Sector**

This table reports the results of regressing 3-year changes in an establishment's log RegIndex on shift-share variables in Table 6 in each NAICS 1-digit sector. Equation (7) provides the regression specification. See Table 6 for variable definitions. All regressions control for $\Delta\log(\text{Wage})$, $\Delta\log(w^r)$, log(RegIndex), and year fixed effects. All variables are standardized to have mean 0 and variance of 1 for ease of interpretation. Standard errors are double clustered by industry and year and presented in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. The sample period is from 2002 to 2014.

	All Sizes			Subsample by Establishment Size			
				1-19	20-399	400-749	$\geq$ 750
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Agriculture, Mining, and Construction				
$iv(\Delta\log(p_{it}))$	0.207**		0.118	0.127*	0.112	0.170	0.037
	(0.065)		(0.069)	(0.068)	(0.068)	(0.095)	(0.123)
$iv(\Delta\log(\tilde{R}_{it}))$		0.275***	0.243***	0.188***	0.283***	0.241*	0.162
		(0.041)	(0.054)	(0.055)	(0.053)	(0.119)	(0.112)
			Manufacturing				
$iv(\Delta\log(p_{it}))$	0.388***		0.229**	0.154**	0.242***	0.304**	0.349***
	(0.073)		(0.072)	(0.066)	(0.072)	(0.096)	(0.103)
$iv(\Delta\log(\tilde{R}_{it}))$		0.342***	0.288***	0.238***	0.312***	0.331***	0.460***
		(0.054)	(0.028)	(0.028)	(0.030)	(0.036)	(0.039)
			Retail				
$iv(\Delta\log(p_{it}))$	0.195***		0.154***	0.116**	0.188***	0.193**	0.008
	(0.054)		(0.040)	(0.041)	(0.044)	(0.071)	(0.138)
$iv(\Delta\log(\tilde{R}_{it}))$		0.189***	0.171***	0.122***	0.202***	0.287***	0.413***
		(0.049)	(0.045)	(0.031)	(0.057)	(0.072)	(0.104)
			Wholesale				
$iv(\Delta\log(p_{it}))$	0.252***		0.164***	0.112**	0.205***	0.464***	0.504***
	(0.052)		(0.039)	(0.042)	(0.048)	(0.085)	(0.120)
$iv(\Delta\log(\tilde{R}_{it}))$		0.191***	0.155***	0.128***	0.182***	0.230**	0.089**
		(0.035)	(0.025)	(0.031)	(0.027)	(0.082)	(0.032)

Table A.13: **Regulatory Requirements vs. Enforcement for Establishments by Sector**–*Continued*

This table reports the results of regressing 3-year changes in an establishment's log RegIndex on shift-share variables in Table 6 in each NAICS 1-digit sector. Equation (7) provides the regression specification. See Table 6 for variable definitions. All regressions control for Δlog(Wage), $\Delta\log(w^r)$, log(RegIndex), and year fixed effects. All variables are standardized to have mean 0 and variance of 1 for ease of interpretation. Standard errors are double clustered by industry and year and presented in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. The sample period is from 2002 to 2014.

	All Sizes			Subsample by Establishment Size			
				1-19	20-399	400-749	$\geq$ 750
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Utilities			
$iv(\Delta\log(p_{it}))$	0.362***		0.150	0.233	0.047	0.161	0.078
	(0.075)		(0.086)	(0.141)	(0.095)	(0.272)	(0.286)
$iv(\Delta\log(\tilde{R}_{it}))$		0.373***	0.319***	0.313*	0.293***	0.058	0.293
		(0.081)	(0.075)	(0.165)	(0.068)	(0.232)	(0.223)
				Transportation			
$iv(\Delta\log(p_{it}))$	0.192*		0.134*	0.086*	0.143*	0.193*	0.284
	(0.086)		(0.062)	(0.045)	(0.068)	(0.086)	(0.167)
$iv(\Delta\log(\tilde{R}_{it}))$		0.366***	0.350***	0.294***	0.373***	0.393***	0.526***
		(0.033)	(0.023)	(0.034)	(0.023)	(0.064)	(0.066)
				Finance			
$iv(\Delta\log(p_{it}))$	0.377***		0.196**	0.199**	0.207***	0.161***	-0.073
	(0.039)		(0.062)	(0.081)	(0.063)	(0.045)	(0.132)
$iv(\Delta\log(\tilde{R}_{it}))$		0.294***	0.219**	0.216**	0.234**	0.298***	0.285***
		(0.063)	(0.080)	(0.089)	(0.078)	(0.067)	(0.075)
				Service			
$iv(\Delta\log(p_{it}))$	0.232***		0.121	0.117	0.127	0.129**	0.064
	(0.062)		(0.073)	(0.071)	(0.071)	(0.052)	(0.060)
$iv(\Delta\log(\tilde{R}_{it}))$		0.260***	0.218***	0.167**	0.259***	0.264***	0.285***
		(0.040)	(0.058)	(0.057)	(0.056)	(0.063)	(0.074)

Table A.14: **Robustness—Regulatory Requirements vs. Enforcement: Using Alternative Enforcement Measures**

This table reports the robustness check of Panel A of Table 6 by reconstructing the shift-share variable for enforcement changes using only enforcement-related regulatory occupations in each agency. Equation (7) provides the regression specification. See Appendix C and Table 6 for details.

	Firm-Level			Establishment-Level		
	(1)	(2)	(3)	(4)	(5)	(6)
$iv(\Delta\log(p_{it}))$	0.223**		0.115	0.237***		0.125*
	(0.070)		(0.078)	(0.061)		(0.067)
$iv(\Delta\log(\tilde{R}_{it}))$		0.257***	0.224***		0.269***	0.231***
		(0.038)	(0.056)		(0.036)	(0.052)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	608,500	608,500	608,500	628,733	628,733	628,733
Adjusted R^2	0.323	0.340	0.344	0.323	0.340	0.344

Table A.15: **Robustness—Regulatory Requirements vs. Enforcement in Subsamples by Size: Using Alternative Enforcement Measures**

This table reports the robustness check of Panel B of Table 6 by reconstructing the shift-share variable for enforcement changes using only enforcement-related regulatory occupations in each agency. Equation (7) provides the regression specification. See Appendix C and Table 6 for details.

	Firm-Level				Establishment-Level			
	1-19	20-399	400-749	$\geq$ 750	1-19	20-399	400-749	$\geq$ 750
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$iv(\Delta\log(p_{it}))$	0.103	0.124	0.105	0.086	0.112	0.136*	0.169*	0.090
	(0.079)	(0.081)	(0.077)	(0.063)	(0.065)	(0.069)	(0.077)	(0.073)
$iv(\Delta\log(\tilde{R}_{it}))$	0.166**	0.253***	0.287***	0.309***	0.170**	0.271***	0.313***	0.365***
	(0.059)	(0.056)	(0.055)	(0.046)	(0.054)	(0.051)	(0.053)	(0.057)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	189,404	352,779	29,482	36,835	220,464	375,485	19,622	13,162
Adjusted R^2	0.400	0.324	0.243	0.214	0.397	0.320	0.265	0.233

Table A.16: **Robustness—Regulatory Requirements vs. Enforcement for Firms by Sector: : Using Alternative Enforcement Measures**

This table reports the robustness check of Table A.12 by reconstructing the shift-share variable for enforcement changes using only enforcement-related regulatory occupations in each agency. Equation (7) provides the regression specification. See Appendix C and Table A.12 for details.

	All Sizes			Subsample by Firm Size			
				1-19	20-399	400-749	$\geq$ 750
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Agriculture, Mining, and Construction							
$iv(\Delta\log(p_{it}))$	0.195**		0.096	0.093	0.095	0.056	0.206**
	(0.069)		(0.080)	(0.084)	(0.077)	(0.107)	(0.089)
$iv(\Delta\log(\tilde{R}_{it}))$		0.269***	0.241***	0.195**	0.273***	0.278**	0.198*
		(0.044)	(0.061)	(0.064)	(0.061)	(0.099)	(0.106)
Manufacturing							
$iv(\Delta\log(p_{it}))$	0.360***		0.195**	0.122*	0.200**	0.260**	0.251**
	(0.074)		(0.079)	(0.064)	(0.085)	(0.094)	(0.100)
$iv(\Delta\log(\tilde{R}_{it}))$		0.331***	0.281***	0.239***	0.303***	0.320***	0.412***
		(0.057)	(0.043)	(0.042)	(0.048)	(0.044)	(0.041)
Retail							
$iv(\Delta\log(p_{it}))$	0.127***		0.096**	0.094**	0.125**	0.063	0.028
	(0.038)		(0.035)	(0.041)	(0.046)	(0.051)	(0.040)
$iv(\Delta\log(\tilde{R}_{it}))$		0.134***	0.122***	0.122***	0.130***	0.102**	0.105***
		(0.028)	(0.025)	(0.020)	(0.032)	(0.033)	(0.022)
Wholesale							
$iv(\Delta\log(p_{it}))$	0.215***		0.125**	0.073*	0.165**	0.094	0.202*
	(0.052)		(0.040)	(0.034)	(0.058)	(0.097)	(0.109)
$iv(\Delta\log(\tilde{R}_{it}))$		0.194***	0.168***	0.125***	0.193***	0.288***	0.275***
		(0.038)	(0.032)	(0.035)	(0.040)	(0.073)	(0.070)
Utilities							
$iv(\Delta\log(p_{it}))$	0.538***		0.307**	0.204*	0.336*	-0.028	0.159
	(0.119)		(0.100)	(0.100)	(0.153)	(0.264)	(0.135)
$iv(\Delta\log(\tilde{R}_{it}))$		0.443***	0.335***	0.501***	0.211	0.529***	0.376***
		(0.130)	(0.090)	(0.081)	(0.167)	(0.054)	(0.045)
Transportation							
$iv(\Delta\log(p_{it}))$	0.192*		0.107	0.068	0.111	0.142**	0.198**
	(0.086)		(0.065)	(0.060)	(0.069)	(0.058)	(0.073)
$iv(\Delta\log(\tilde{R}_{it}))$		0.374***	0.354***	0.305***	0.362***	0.441***	0.503***
		(0.037)	(0.031)	(0.038)	(0.028)	(0.065)	(0.059)
Finance							
$iv(\Delta\log(p_{it}))$	0.371***		0.167**	0.164	0.169*	0.152	0.063
	(0.058)		(0.073)	(0.090)	(0.079)	(0.141)	(0.110)
$iv(\Delta\log(\tilde{R}_{it}))$		0.304***	0.248**	0.215**	0.284***	0.246***	0.347**
		(0.067)	(0.079)	(0.094)	(0.067)	(0.060)	(0.107)
Service							
$iv(\Delta\log(p_{it}))$	0.215**		0.107	0.106	0.110	0.069	0.044
	(0.070)		(0.085)	(0.084)	(0.085)	(0.068)	(0.065)
$iv(\Delta\log(\tilde{R}_{it}))$		0.251***	0.217***	0.164**	0.251***	0.265***	0.276***
		(0.040)	(0.060)	(0.061)	(0.060)	(0.047)	(0.047)

Table A.17: **Robustness—Regulatory Requirements vs. Enforcement for Establishments by Sector: Using Alternative Enforcement Measures**

This table reports the robustness check of Table A.13 by reconstructing the shift-share variable for enforcement changes using only enforcement-related regulatory occupations in each agency. Equation (7) provides the regression specification. See Appendix C and Table A.13 for details.

	All Sizes			Subsample by Firm Size			
				1-19	20-399	400-749	$\geq$ 750
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Agriculture, Mining, and Construction							
$iv(\Delta\log(p_{it}))$	0.194**		0.097	0.105	0.091	0.146	0.033
	(0.065)		(0.069)	(0.071)	(0.067)	(0.096)	(0.128)
$iv(\Delta\log(\tilde{R}_{it}))$		0.275***	0.247***	0.193***	0.286***	0.241*	0.161
		(0.041)	(0.054)	(0.056)	(0.052)	(0.119)	(0.115)
Manufacturing							
$iv(\Delta\log(p_{it}))$	0.389***		0.233**	0.154**	0.246***	0.311**	0.359***
	(0.072)		(0.073)	(0.066)	(0.071)	(0.100)	(0.108)
$iv(\Delta\log(\tilde{R}_{it}))$		0.342***	0.288***	0.239***	0.313***	0.330***	0.459***
		(0.054)	(0.027)	(0.027)	(0.029)	(0.037)	(0.041)
Retail							
$iv(\Delta\log(p_{it}))$	0.165*		0.149***	0.111**	0.185***	0.179*	0.023
	(0.073)		(0.045)	(0.046)	(0.048)	(0.080)	(0.138)
$iv(\Delta\log(\tilde{R}_{it}))$		0.189***	0.183***	0.132***	0.215***	0.296***	0.411***
		(0.049)	(0.044)	(0.030)	(0.057)	(0.076)	(0.102)
Wholesale							
$iv(\Delta\log(p_{it}))$	0.237***		0.156***	0.107**	0.192***	0.509***	0.435*
	(0.062)		(0.040)	(0.039)	(0.052)	(0.065)	(0.202)
$iv(\Delta\log(\tilde{R}_{it}))$		0.191***	0.161***	0.133***	0.189***	0.223**	0.093***
		(0.035)	(0.025)	(0.029)	(0.029)	(0.082)	(0.028)
Utilities							
$iv(\Delta\log(p_{it}))$	0.420***		0.219**	0.293*	0.101	0.269	0.067
	(0.090)		(0.096)	(0.133)	(0.109)	(0.259)	(0.312)
$iv(\Delta\log(\tilde{R}_{it}))$		0.373***	0.297***	0.291*	0.276***	0.009	0.298
		(0.081)	(0.056)	(0.139)	(0.052)	(0.222)	(0.232)
Transportation							
$iv(\Delta\log(p_{it}))$	0.231**		0.146**	0.100*	0.154*	0.202**	0.294
	(0.074)		(0.064)	(0.047)	(0.070)	(0.084)	(0.164)
$iv(\Delta\log(\tilde{R}_{it}))$		0.366***	0.341***	0.286***	0.363***	0.384***	0.508***
		(0.033)	(0.023)	(0.034)	(0.023)	(0.060)	(0.060)
Finance							
$iv(\Delta\log(p_{it}))$	0.352***		0.162**	0.163**	0.167**	0.101	-0.095
	(0.049)		(0.052)	(0.070)	(0.052)	(0.067)	(0.147)
$iv(\Delta\log(\tilde{R}_{it}))$		0.294***	0.234**	0.232**	0.252***	0.327***	0.290***
		(0.063)	(0.075)	(0.086)	(0.071)	(0.081)	(0.076)
Service							
$iv(\Delta\log(p_{it}))$	0.239***		0.124	0.116	0.133*	0.134**	0.075
	(0.058)		(0.074)	(0.072)	(0.072)	(0.054)	(0.066)
$iv(\Delta\log(\tilde{R}_{it}))$		0.260***	0.217***	0.167**	0.257***	0.263***	0.282***
		(0.040)	(0.058)	(0.057)	(0.056)	(0.065)	(0.075)

Table A.18: **Examples of Threshold-Imposed Regulations in the U.S.**

This table lists examples of U.S. regulations that have thresholds and the scope of their affected industries.

Regulation	**Size Threshold**	**What Threshold Triggers**	**Affected Industry**
Affordable Care Act (Federal)	50+ employees	Must offer affordable health coverage or face penalties and report offers to IRS.	All Industries
Equal Employment Opportunity Commission EEO-1 Report (Federal)	100+ employees (50+ for federal contractors)	Must file annual workforce demographics reports to EEOC/DOL.	All Industries
OSHA Injury & Illness Recordkeeping (Federal)	11+ employees	Must maintain OSHA injury/illness logs; Some (250+ employees or 20-249 employees in high-hazard industries) must e-submit data.	Most Industries (ex. OSHA partially-exempted NAICS)
EPA Toxic Release Inventory Reporting under EPCRA §313 (Federal)	10+ employees	Must disclose chemical releases, waste management, and pollution prevention activities.	Manufacturing, Mining, Electric Utilities, and Hazardous Waste
OSHA Field Sanitation under 29 CFR 1928.110 (Federal)	11+ employees (hand-labor)	Must provide drinking water, toilets, and hand washing in fields; maintain facilities; allow use during work.	Agriculture
Migrant and Seasonal Agricultural Worker Protection Act—Small Business Exemption (Federal)	exempted if use <500 man-days in any quarter of prior year	Exempted from MSPA's wage, housing, transport rules.	Crop & Animal Production & Support Activities for Agriculture
FDA Food Safety Modernization Act (Federal)	500+ employees	Must comply with all Preventive Controls requirements by the earliest deadline without exemptions or extensions.	Manufacturing, Wholesale, Retail (food-related)
Systemically Important Financial Institution / Enhanced Prudential Standards (Federal)	\$50B+ assets (\$100B/\$250B after 2018)	Must comply with enhanced prudential standards on capital, liquidity, resolution plans, etc.	Finance & Insurance
Investment Adviser Registration—SEC vs. State (Federal + State)	\$100M+ assets under management (AUM) for SEC; \$25-\$100M AUM for State	Must register with SEC (or state) as investment adviser if their assets under management passes the respectively threshold	Other Financial Investment Activities
Sarbanes-Oxley Act (Federal)	\$75M+ by public float with additional thresholds by public float and revenues	Must comply with SOX 404(a) and 404(b) on accurate reporting, internal controls, audits, and fraud prevention.	All Industries (Among Publicly-Traded Firms)
Americans with Disabilities Act (Federal)	15+ employees ($\geq$20 weeks)	Must provide equal employment opportunities and reasonable accommodation to qualified individuals with disabilities.	All Industries
National Labor Relations Board (NLRB) jurisdictional standards (Federal)	\$500k+ gross annual volume for retail \$50k+ for non-retail	Must comply with the National Labor Relations Act enforced by NLRB.	All Private Industries
Illinois Secure Choice Savings Program (State)	5+ employees (in operation for $\geq$ 2 yrs with no existing retirement plan)	Must auto-enroll workers in state IRA unless plan offered.	All Industries in Illinois
California Family Rights Act (State)	50+ employees (5+ after 2021)	Must accommodate employees with up to 12 weeks unpaid, job-protected leave for qualifying reasons.	All Industries in California

Table A.19: **Within-Sector Changes in Firm Size Distribution and RegIndex**

This table reports the association of RegIndex and changes in business size distribution within NAICS 1-digit sectors from 2002 to 2014. Panel A reports the results on changes in firm size distribution, and Panel B reports the results on changes in establishment size distribution. In Panel B, the dependent variable is the changes in the share of establishments (by number) from each employment size bin within each NAICS 1-digit sector from 2002 to 2014. We compute each size bin's share of establishments within each sector as the number of establishments from the size bin divided by total number of establishments within the sector and year, where establishments are weighted by the BLS's sampling weights. In Panel A, we similarly construct the shares of firms (EINs) from each size bin within each sector and year. See Section 5.3 for more details. The size bins are [1, 2], [3, 4], [5, 9], [10, 19], [20, 49], [50, 99], [100, 249], [250, 499], [500, 749], [750, 999], [1000, 1499], [1500, 2499], [2500, 3999], and above 4000. *RegIndex* is the average RegIndex of the sector-size bin in our sample, and Δ*RegIndex* is the change in the sector-size bin's RegIndex from 2002 to 2014. Columns labeled by *Small* and *Large* represent the subsamples based on the size bins above and below 250 employees, respectively. Standard errors are clustered at the sector level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Changes in Firm Size Distribution within Sectors (2002-2014)						
	All	Small	Large	All	Small	Large
	(1)	(2)	(3)	(4)	(5)	(6)
RegIndex	-0.0167**	-0.0407***	-0.0002			
	(0.0053)	(0.0059)	(0.0009)			
ΔRegIndex				-0.0135**	-0.0269***	0.0003
				(0.0040)	(0.0074)	(0.0003)
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	112	56	56	112	56	56
Adjusted R^2	0.077	0.225	0.023	0.021	0.038	0.030
Panel B: Changes in Establishment Size Distribution within Sectors (2002-2014)						
	All	Small	Large	All	Small	Large
	(1)	(2)	(3)	(4)	(5)	(6)
RegIndex	-0.0161***	-0.0514***	0.0004			
	(0.0044)	(0.0109)	(0.0003)			
ΔRegIndex				-0.0135***	-0.0490***	-0.0002
				(0.0016)	(0.0048)	(0.0001)
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	112	56	56	112	56	56
Adjusted R^2	0.074	0.322	0.198	0.059	0.343	0.192