The Cost of Regulatory Compliance in the United States

Francesco Trebbi   Miao Ben Zhang   Michael Simkovic*

July 2023

Abstract

One of the key questions in the study of regulation is whether the costs of regulatory compliance fall homogeneously on all businesses or whether certain firms, for instance small ones, are especially penalized. We quantify firms' compliance costs in terms of their labor spending to adhere to government rules. Using comprehensive establishment-level occupational microdata and occupation-specific task information, we recover the proportion of a firm's wage bill attributable to employees engaged in regulatory compliance. On average for 2002-14, regulatory costs account for 1.34% to 3.33% of a firm's wage bill, totaling up in 2014 to $239 billion, and to $289 billion when adding capital equipment costs. Our findings reveal an inverted-U relation between firms' regulatory compliance costs and their scale of employment, indicating that firms with approximately 500 employees face compliance costs that are about 40 percent higher as a share of total wages compared to small or large firms. Finally, we develop an instrumental variable methodology to disentangle the influence of regulatory requirements and enforcement in 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 Matt Backus, Matilde Bombardini, Julieta Caunedo, Ernesto Dal Bó, Steven Davis, Klaus Desmet, Fred Finan, Bård Harstad, Alessandro Lizzeri, John Matsusaka, Tom Palfrey, Amit Seru, Reed Walker, Jessie Wang, and seminar participants at the NBER Economic Analysis of Regulation Conference, Tepper-LAEF Conference, PSL Public Governance Seminar, Berkeley Political Economy Research Lab, 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 Michael Soloy at the BLS for their excellent assistance with the confidential data. 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 falling entry, 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 free entry and 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-level regulatory shocks. 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 presence of size-dependent regulatory constraints (Restuccia and Rogerson, 2008; Guner et al., 2008; García-Santana and Pijoan-Mas, 2011; Hovenkamp and Morton, 2019) and their extent is unknown.

While these questions are central for the ongoing policy debate on regulation, measuring 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 each other (Agarwal et al., 2014). 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

[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.

[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).

or frequency of inspection (Stiglitz, 2009; Kang and Silveira, 2021), which may result in heterogeneous on-equilibrium compliance effort. Fifth, small and large firms face different technological constraints and reach different efficiency levels in complying with regulation, for example due to economies of scale or fixed costs.

A first contribution of this paper – and a necessary step for studying the relation between regulation and firm size – is to systematically measure the costs of regulation for establishments and firms. We focus on labor costs for regulatory compliance first, which by many accounts are a central component of compliance.[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 2002-14.

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

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. Given the specific focus of our approach, these figures appear sensible.[4] 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.[5] When aggregating RegIndex to the national level, we observe that regulatory

[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.*" See https://www.jpmorganchase.com/content/dam/jpmc/jpmorgan-chase-and-co/investor-relations/documents/2013AR-Chairman-CEO-letter.pdf. 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.

[4]For example, in 2003 Milton Friedman estimated an upper bound of "*10% or so of income that residents and businesses spend in response to government mandates and regulation.*" The Wall Street Journal "What Every American Wants" January 15th, 2003, by Milton Friedman.

[5]We use the most conservative version of RegIndex that we construct 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

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 regulatory and deregulatory shocks, 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.

After discussing labor costs, 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 used 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). 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. On the one hand, regulatory costs increasing with size imply 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).[6] On the other hand, regulatory costs decreasing with scale favor 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

excludes taxes, but does not downweight regulatory related tasks based on the presence of other task objectives or keyworks. These are available from the authors upon request.

[6]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).

fixed costs,[7] or it may even derive from regulatory capture and special deals for large players.[8]

Our main finding in Section 4 is an inverted-U relation between RegIndex and firm size. In particular, firms under 500 employees experience increasing regulatory compliance costs as 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. We observe this relation after controlling for industry fixed effects, state fixed effects, and even firm fixed effects. 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 size. Specifically with respect to within-occupation variation, 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 economic to consolidate their compliance efforts in the hands of dedicated employees to save costs. We then propose and implement a shift-share instrumental variable 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. For the construction of our IV's we exploit the fact that each firm has a different exposure to the various federal agency changes in regulatory requirements and in en-

[7] 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); Calomiris et al. (2020) find larger corporations being at an advantage (while focusing only on public companies).

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

forcement/inspection effort due to sector, production technology, location, and organization. We find evidence that the fact that small business 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 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. capital structures like reinforced concrete walls, pumping or draining infrastructure for mine water, etc.) and foregone investment opportunities and profits due to regulatory risk.[9] 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. 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. Another important limitation of our study is that we do not include in our measure costs of compliance that are borne by firms through outsourcing (e.g. external legal, compliance, and accounting services). The aforementioned surveys also 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. Based on the data, this would not appear to be a damning omission.

This paper's methodology is a partial, but relevant step in the direction of a more reliable, granular assessment of the economic costs of regulation borne by producers. In particular, our measure has several important features. First, our approach has the advantage of being a firm-specific micro measure, which is crucial to capture heterogeneity in regulatory burden, as the same regulation may affect different firms differently, even within the same sector. Second, our results focus on actual compliance costs paid by firms and not on statements, expectations, or projections by either firm management or government agencies. Third, our measure looks at on-equilibrium effects of regulation, incorporating the endogenous responses by firms to enforcement, to regulatory ambiguity, etc. Fourth, our work covers both very large and very small firms – which is an extremely important feature in the context of assessing the effects of government rules on small businesses. Fifth, our approach captures regulations from all relevant sources, including federal, state, local,

[9] 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).

and industry privately-enforced regulations. Sixth, our approach is operational and reproducible by government agencies in their ongoing assessments and validation of rules. For example, it can be adopted as a straightforward empirical measure for ex post compliance costs, complementing OIRA's ex ante projections.

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] For instance, 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 (2021) employs text data from Form 83-I filed to OIRA to indicate the expectations of regulators about the cost of compliance with each regulation.[11] Davis (2017) and Calomiris et al. (2020) employ an original linguistic approach to measuring regulatory exposure for large publicly traded companies. Davis (2017) focuses on Part 1A of forms 10-K to gauge publicly traded firms' exposure to regulatory and policy risk.[12] Calomiris et al. (2020) focuses on the transcripts of corporate earnings calls made by publicly traded corporations and shows that its measure of increasing regulation is predictive of sales and assets growth, leverage, and other measures of firm performance. 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.[13]

More to its central questions, the paper speaks to the literature on size dependence and regulation, and heterogeneous policies more broadly. The paper has bearing for the question of whether entry

[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. Bailey and Thomas (2017) study wages of high-skilled labor, particularly, accountants and lawyers, from publicly-available OEWS data to examine industry regulatory shocks. Finally, 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.

[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 the estimated dollar costs of compliance. Relative to Al-Ubaydli and McLaughlin (2017), the method in Kalmenovitz (2021) has the major benefit of aggregating the expectations of experts. It is, however, an aggregate measure of regulatory burden as expected by regulators – potentially different from the rules' actual incidence, and not ex post verified. It is also a measure that cannot be directly traced to establishments.

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

[13] Indeed, Calomiris et al. (2020) underscore compliance risk as the most relevant channel behind their finding, rather than physical compliance costs.

and firm dynamism may be hurt by regulation (Alesina et al., 2018; Goldschlag and Tabarrok, 2018) to the extent that the excess compliance costs paid by medium firms may be an obstacle to reach efficient scale of production. Interestingly, and in line with the findings in this paper, using a different measure of costs, Singla (2023) shows that increases in regulatory costs contribute to small firms getting smaller and larger firms getting bigger. The potentially distortive effects of non-neutral policies on productivity and innovation are presented in a series of studies, too long to exhaustively summarize here, that look at specific types of tax or regulatory distortions (Pashigian, 1984; Guner et al., 2008; Garicano et al., 2016; Aghion et al., 2021) or more general quantitative models (Parente and Prescott, 2002; Restuccia and Rogerson, 2008). Moving forward, the findings in this paper may provide empirical moments for quantitative models to match. Further, while assessing the productivity consequences of the regulatory compliance requirements that we identify is beyond the scope of this paper, future research should address this issue explicitly at the establishment level.

Legal scholarship also relates regulatory and compliance costs to firm size in areas such as corporate governance and securities regulation, anti-trust, and federalism. First, large and mid-sized firms are typically organized as Delaware Corporations, while smaller firms favor other business forms (Daines, 2001; Eldar and Magnolfi, 2020). Delaware law imposes on corporate directors and officers a duty to monitor the corporation and to implement appropriate compliance systems (Arlen, 2008; Langevoort, 2017). Second, whereas the largest firms are typically publicly traded, smaller private firms are not subject to federal securities regulation (Ewens et al., 2021). In addition, some aspects of securities regulation phase in as firm size increases, for example some of the requirements under the Sarbanes Oxley Act. Third, anti-trust enforcement historically targeted larger firms (Hovenkamp and Morton, 2019). And finally, the division of legal authority between federal, state, and local governments has led to a complex web of regulations affecting firms as they branch across jurisdictions. Efforts to mitigate resulting costs include uniform commercial laws, contractual choice of law provisions, pre-emption, and constitutional protections for interstate commerce (Ribstein, 1992; Schwartz and Scott, 1994; Nelson, 2000; Redish and Nugent, 1987).

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; 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.[14] 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.[15] 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.[16]

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

[14] 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.

[15] 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.

[16] 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.

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 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 $\tilde{R}_{it} = \tilde{R}(L_{it}) = \rho L^{\alpha}_{it}$ and $p_{it} = \pi L^{\beta}_{it}$, where α, β, ρ, π are policy parameters governing regulatory requirements and enforcement effort targeting establishment i. Simple static profit maximization implies

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

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^r_{it}.

We define the index of regulatory compliance costs as:

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

that is, the share of the total wage bill $W_{it} = w^r_{it}R_{it} + w_{it}L_{it}$ 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 Online 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. Online 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.[17]

A firm's RegIndex is similarly obtained based on occupational labor costs aggregated from those of 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.[18]

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.[19]

[17] 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 $\text{RegIndex}_{i,o,t}$.

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

[19] 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

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.[20] 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, 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 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

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%.

[20]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.

income taxes amounted to $353 billion in the same year.[21]

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.[22]

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.[23] Panel A shows that our estimates based on the conservative RegIndex account for about half of the hours estimated by federal agencies through OIRA. Using

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

[22] 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.

[23] Compliance hours for individual income taxes account for about 40 percent of the Department of Treasury's total 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/.

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.

3.5.2 Establishment RegIndex and Industry Regulatory Shocks

As a second validation exercise, this subsection illustrates the response of RegIndex to four major industry-level regulatory shocks. 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 shocks, 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).[24] 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).

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

[24]The Card Act of 2009 was enacted after research suggested that credit card pricing was less than perfectly competitive.

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 Online Appendix Figure A.3 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.[25]

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 prior to the shocks.[26] 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).[27]

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 Online Appendix Figure A.3 that the supply-side RegData measure exhibits an increase in regulation of oil and gas

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

[26]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 shocks. 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.

[27]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.

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.[28]

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 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 Online Appendix Figure A.3. 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

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

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 Online Appendix Figure A.3 that the RegData measure can also identify increased regulatory restrictions for hospitals from animal hospitals.

Finally, Online Appendix Table A.4 reports the statistical significance of the graphical evidence of the cases in Figure 3.

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,[29] 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 Online Appendix Figure A.2). 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.

[29]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.

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 elections. Table 2 shows that state-specific RegIndex is significantly and negatively related to the state Republican vote share in all three elections.[30]

4 Returns to Scale in Regulatory Compliance

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, $\frac{\partial \text{RegIndex}_{it}}{\partial L_{it}}$.[31] 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.[32] 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 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 $\frac{\partial \tilde{R}_{it}}{\partial L_{it}} \gtreqless 0$.[33] 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

[30] In the Online 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.

[31] 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 $L_{it} + R_{it}$.

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

[33] 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.

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.[34]

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 non-parametric 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 uncover the percentage of labor costs for regulatory compliance for 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 3, 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.

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

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

expect from the vast differences in sector specific regulations.

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 occupations, we can explicitly control for Year×Occupation fixed effects. 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.

Column (1) of Table 4 provides clear evidence that firms' demand for regulatory compliance tasks also exhibits an inverted-U relationship with firm size, 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' actual percentage 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, $\text{RegIndex}^{tot}_{i,t}$, is employed in replicating Figure 5.

Figure 7 shows how labor-based $\text{RegIndex}_{i,t}$ and total-input-based $\text{RegIndex}^{tot}_{i,t}$ compare in terms of relationship with firm and establishment size. We observe again a distinctive inverted-U shape for $\text{RegIndex}^{tot}_{i,t}$, 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 $\text{RegIndex}^{tot}_{i,t}$ 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 $\text{RegIndex}^{tot}_{i,t}$ 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.[35]

4.1.3 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 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.[36]

Figure 8 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

[35]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.

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

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

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. We then engage in a more complex exercise, 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.

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 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 9, 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 applied onto firms, $\tilde{R}_{it}$, and enforcement effort, E_{it}, which in turn affects the likelihood of inspection p_{it}. These competing drivers, which may also be size-dependent, introduce a challenge in interpreting the findings in Section 4. For example, a small firm may exhibit low RegIndex because the agency rules are designed to be lighter on smaller firms (proxied by lower $\tilde{R}_{it}$ for smaller firms). Alternatively, this may happen because small firms are more likely to fly under the radar of regulators (proxied by lower p_{it} for smaller firms) and therefore in equilibrium they choose to comply less to rules that otherwise would apply equally to all firms. These considerations are relevant to decomposing the upward sloping element of the non-monotonic shape of the relationship between compliance costs and firm size found in Section 4.

This section presents a method to analyze the impact of the two main factors on driving the RegIndex of firms of a given size. In order to identify the differential role of regulatory requirements and of enforcement in driving regulatory compliance, we add additional information originating from the supply side, i.e. from the regulatory agencies who design and enforce the rules. In particular, we exploit the fact that some shifts in RegIndex track changes in the official text of regulations, while others track changes in regulatory agencies' enforcement and personnel budgets. We first describe and implement two shift-share regulatory supply shocks help identify regulation-related tasks' sensitivities to regulatory requirements and to enforcement. Next, we decompose each firm's regulatory compliance costs into a component paid for requirement-sensitive tasks and a component paid for enforcement-sensitive tasks. For instance, one would expect that the task of "*learn and follow safety regulations*" to be more sensitive to new regulatory requirement, while the task to "*enforce safety regulations and policies*" to be more sensitive to regulatory enforcement. With this additional information at hand, we construct two components of RegIndex: The share of a firm's labor spending on requirement-sensitive tasks, $\text{RegIndex}^{req}_{i,t}$ and the share of a firm's labor spending on enforcement-sensitive tasks, $\text{RegIndex}^{enf}_{i,t}$. This is constructed so that $\text{RegIndex}^{req}_{i,t} + \text{RegIndex}^{enf}_{i,t} = \text{RegIndex}_{i,t}$.

5.2.1 Framework for Estimating Task Sensitivity to Regulatory Requirement and Enforcement Shocks

For each regulatory agency k at time t, suppose one can measure both its regulatory requirement, reg_{kt}, and the extent, for given reg_{kt}, of the enforcement and supervision effort, enf_{kt}. Then, we can construct measures of *shocks* to reg_{kt} and enf_{kt} for each agency k at time t, which we discuss in Section 5.2.2 below.

We posit that different industries have different *exposure* to regulations falling under agency k's oversight. Suppose one can measure industry j's RegIndex originating from regulatory agency k, r_{jkt}. Then, one can, in turn, create two shift-share instrumental variables tracing changes in regulatory requirements and in enforcement pertinent to firm i from industry j:

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

where $\Delta\log reg_{kt}$ and $\Delta\log enf_{kt}$ are the agency-specific requirement shocks and enforcement shocks from year t to $t+3$, and r_{jkt} is industry j's exposure to agency k's regulations in year t. ν_{it} and υ_{it} are firm-specific i.i.d. shocks.

Our key assumption is that certain regulation-related tasks are more relevant to complying with regulatory requirement while other regulation-related tasks are more relevant to complying with enforcement. Firms respond to regulatory requirement shocks by spending more on *requirement-sensitive tasks* and respond to enforcement shocks by spending more on *enforcement-sensitive tasks*. We can thus categorize each regulation-related task o as *requirement-sensitive* or *enforcement-sensitive* by estimating the following regression for each regulation-related 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 log change in spending on task o from year t to $t+3$[37], and β_o and γ_o measure task o's sensitivity to regulatory requirement shocks and enforcement shocks, respectively. We also include firm level controls X_{it}, including log changes in firms' total labor costs and log changes in the average wage rate of firms' regulation-related tasks, which are motivated by our framework in Section 3.2,[38] firms' initial RegIndex level to account firm heterogeneity, and also year fixed effects to account for macro conditions.

We run the above regression in equation (6) for each regulation-related task o. We categorize a task as follows: We standardize the two IVs in equation (6) and run a test of coefficients for the two standardized IVs, β_o and γ_o. If $\beta_o > \gamma_o$ and is statistically significant at five percent level, we regard this task o as *requirement-sensitive*. If $\beta_o < \gamma_o$ and is significant at five percent level, we regard this task as *enforcement-sensitive*. If the two coefficients are not statistically differentiable at five percent significance level, we regard task o as *mixed*, half requirement-sensitive and half enforcement-sensitive.[39] With this procedure, 49% of regulation-related tasks are labeled as *requirement-sensitive*, 16% as *enforcement-sensitive,* and 35% as *mixed*. Once each task is categorized,

[37]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 shocks). 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) also the two IVs because establishments are covered by the OEWS sample in every three years.

[38]In particular, plugging equation (1) into equation (2) and take log changes, we have $\Delta\log(\text{RegIndex}_{it}) = \frac{1}{2}\Delta\log(\tilde{R}_{it}) + \frac{1}{2}\Delta\log(p_{it}) + \frac{1}{2}\Delta\log(w^r_{it}) - \Delta\log W_{it}$.

[39]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.

we can then adopt the exact definition in equation (2) to create a requirement-specific $\text{RegIndex}_{i,t}^{req}$ by aggregating only requirement-sensitive tasks, and an enforcement-specific $\text{RegIndex}_{i,t}^{enf}$ by aggregating only enforcement-sensitive tasks.

5.2.2 Constructing the Instruments for Regulatory Requirements and Enforcement Shocks

We now come back to describe our empirical components of the two instruments in equation (5). Each regulatory agency's regulatory requirement shocks, $\Delta \log reg_{kt}$, enforcement shocks, $\Delta \log enf_{kt}$,and each NAICS 6-digit industry's exposure to each regulatory agency, r_{jkt}. All details for the construction are available in the Online Appendix B.

We measure regulatory requirements shocks of each major agency in the U.S., $\Delta \log reg_{kt}$, 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. We use 3-year log differences in reported compliance hours due to changes in enactment and retirement of regulations to measure regulation-requirement shocks of the agency to firms.[40]

We measure enforcement shocks for each major agency k, $\Delta \log enf_{kt}$, using panel data of U.S. federal government employees from 2002 to 2014. 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.[41] Variables crucial for our study include the employee's agency, occupation, and full-time/part-time employment status.[42] 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 shocks of the agency to firms. Finally, we select twelve agencies that have both regulation-requirement shocks and enforcement shocks from 2002 to 2014. These twelve agencies account for 81 percent of non-Treasury regulatory compliance hours for an average year from 2002 to 2014.

Our shift-share IV's in equation (5) also require measuring each industry's exposure to the twelve regulatory agencies' shocks, r_{jkt}. To do so, we extract the top 50 identifying keywords for each

[40] 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.

[41] 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.

[42] 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.

regulatory agency using natural language processing of the CFR text.[43] We compute a Google BERT similarity between each regulation-related task and an agency's keywords, and we standardize the 12 similarities for each task to sum up to 1. These standardized similarities capture a regulation-related task's exposure to the twelve agencies. Note that each NAICS 6-digit industry j's RegIndex is simply the sum of the industry's percentage labor spending on each regulation-related tasks. Multiplying the industry's percentage labor spending on each regulation-related task and each task's exposure to agency k, we obtain industry j's percentage labor spending towards compliance with agency k's regulations in year t, i.e., r_{jkt} in equation (5). With all elements in equation (5) now available, one can construct $iv(\Delta\log(p_{it}))$ and $iv(\Delta\log(\tilde{R}_{it}))$.

5.2.3 Results of Decomposing RegIndex

Figure 10 plots the two components of RegIndex, RegIndex^{req} and RegIndex^{enf} for each size bin. 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. For both the firm and establishment graphs, we observe that the inverted-U relationship between RegIndex and size (in Figure 5) is primarily driven by regulatory requirements, i.e. RegIndex^{req}, rather than enforcement, RegIndex^{enf}.

In particular, while firms in the middle-sized bin 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 is flat. Moreover, RegIndex^{req} accounts for about 100% of the difference in RegIndex between mid-sized firms and the largest size group. The results are similar at the establishment level. In sum, Figure 10 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.

We next corroborate the above decomposition of the *level* of RegIndex with further analyze *changes* in firm RegIndex and the two shocks. Specifically, we run regression of log changes in firm or establishment RegIndex on the regulatory requirement instrument and on the enforcement

[43] We obtain the keywords in three steps: First, we select all the CFR volumes related to the 12 agencies and to their sub-agencies, where the CFR texts are available at https://www.govinfo.gov/app/collection/cfr/2021/. Second, we count the term frequency of each word in the overall selected CFR text ($count_{all}$) and the 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 BERT. Our final keywords are the top 50 keywords that have the highest BERT similarity score. Online Appendix Table A.6 lists the 50 keywords for each of the 12 agencies.

instrument similar to equation (6) as follows:

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

where all variables are described in equation (6) earlier.

Panel A of Table 5 reports our estimation of equation (7) for the full sample of firms and establishments. 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 ($\Delta\log(\text{RegIndex}_{it})$), respectively, to enforcement shocks, $iv(\Delta\log(p_{it}))$. In both columns, we report 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' share of regulatory costs ($\Delta\log(\text{RegIndex}_{it})$), respectively, to regulation requirements shocks, $iv(\Delta\log(\tilde{R}_{it}))$. The coefficient for $iv(\Delta\log(\tilde{R}_{it}))$ is higher 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). Finally, Columns (3) and (6) include both shocks in the estimation of equation (7) to compare the marginal effects of changes in regulatory requirements, $iv(\Delta\log(\tilde{R}_{it}))$, from that of changes in enforcement effort, $iv(\Delta\log(p_{it}))$. This horse-race specification reveals a stronger role for regulatory requirements than for enforcement effort in driving changes in RegIndex. The effect of regulatory requirements is estimated at 0.23 and statistically significant, while that of enforcement effort is at 0.11 and statistically insignificant for firms (and similarly for establishments).[44]

To see how these instrumental variable estimates may be informative about the mechanisms behind the relation between the level of RegIndex and business size, let us 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 5 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 employees or lower. 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 less stringent regulatory requirements for small business, such as for example in the cases of small community banks or for labor laws applying to firms below certain employment levels, shocks to the requirements affecting medium and

[44] Appendix Tables A.5 and A.6 display variations of the above results across broad sectors of firms and establishments, respectively.

large firms have only indirect effects on smaller entities, because small business are largely exempt. This in turn explains the smaller coefficient on $iv(\Delta\log(\tilde{R}_{it}))$. Finally, notice that as regulatory tiering is a feature of the design of regulatory requirements, it is important to hold enforcement of regulation fixed to cleanly isolate this mechanism in the tests of Table 5. This is why controlling for $iv(\Delta\log(p_{it}))$, which may also be size-dependent in principle, is a crucial step in isolating regulatory tiering.

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

6 Conclusion

This paper presents a new approach to estimating the compliance costs of regulation in the United States. Based on 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.

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 shape 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 shift-share instrumental variable methods 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 may extend the use of our methodology to other high-income countries, where similar microdata 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.

Arlen, J., 2008. The story of allis-chalmers, caremark, and stone: Directors' evolving duty to monitor. NYU Law and Economics Research Paper .

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) .

Bailey, J.B., Thomas, D.W., 2017. Regulating away competition: The effect of regulation on entrepreneurship and employment. Journal of Regulatory Economics 52, 237–254. doi:10.1007/s11149-017-9343-9.

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.

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.

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 .

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.

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

Daines, R., 2001. Does delaware law improve firm value? Journal of Financial Economics 62, 525–558.

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

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

Eldar, O., Magnolfi, L., 2020. Regulatory competition and the market for corporate law. American Economic Journal: Microeconomics 12, 60–98.

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., 2021. Regulatory costs of being public: Evidence from bunching estimation. Technical Report. National Bureau of Economic Research.

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.

Goldschlag, N., Tabarrok, A., 2018. Is regulation to blame for the decline in american entrepreneurship? Economic Policy 33, 5–44.

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 .

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

Kalmenovitz, J., 2021. Regulatory intensity and firm-specific exposure. mimeograph Drexel University .

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.

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

Langevoort, D.C., 2017. Caremark and compliance: A twenty-year lookback. Temp. L. Rev. 90, 727.

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.

Nelson, C., 2000. Preemption. Virginia Law Review , 225–305.

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.

Pashigian, B.P., 1984. The effect of environmental regulation on optimal plant size and factor shares. The Journal of Law and Economics 27, 1–28.

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.

Redish, M.H., Nugent, S.V., 1987. Dormant commerce clause and the constitutional balance of federalism, the. Duke LJ , 569.

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

Ribstein, L.E., 1992. Choosing law by contract. J. Corp. L. 18, 245.

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.

Schwartz, A., Scott, R.E., 1994. Political economy of private legislatures. U. Pa. L. Rev. 143, 595.

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

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.

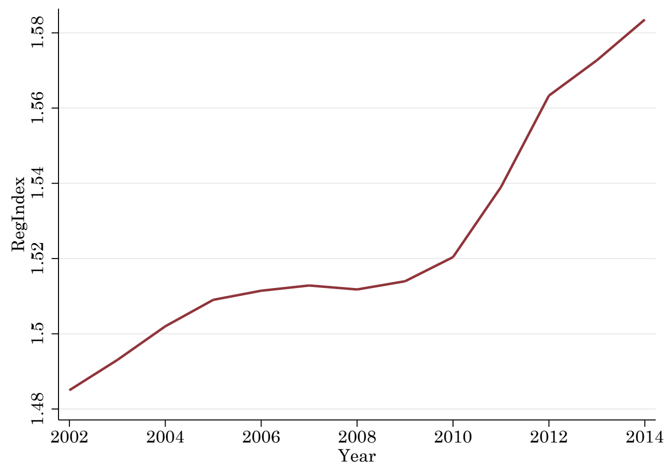


Figure 2: **Validation of Regulation Index Using Agency-Estimated Compliance Hours**

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). Online Appendix Figure A.1 plots additional graphs using the hours identified by two broader versions of the RegIndex measure: Medium RegIndex and Broad RegIndex.

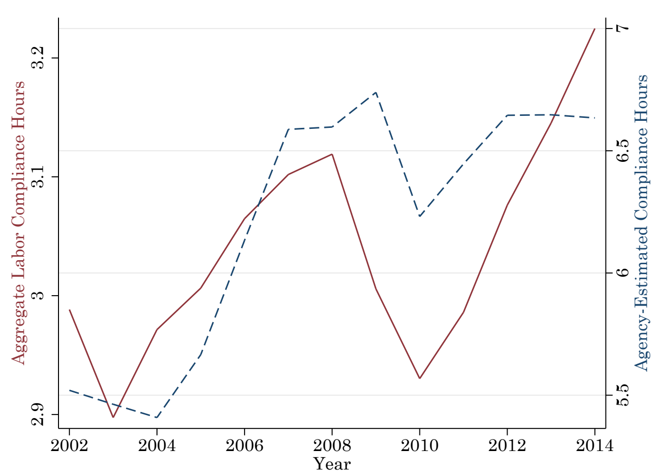


Figure 3: **Validation of Regulation Index Using Industry Regulatory Shocks**

This figure plots the response of industries' Regulation Index (RegIndex) to five industry-level regulatory shocks. RegIndex is the percentage of an industry's annual labor spending on performing regulation-related tasks. Section 3 provides details of the industry shocks 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.

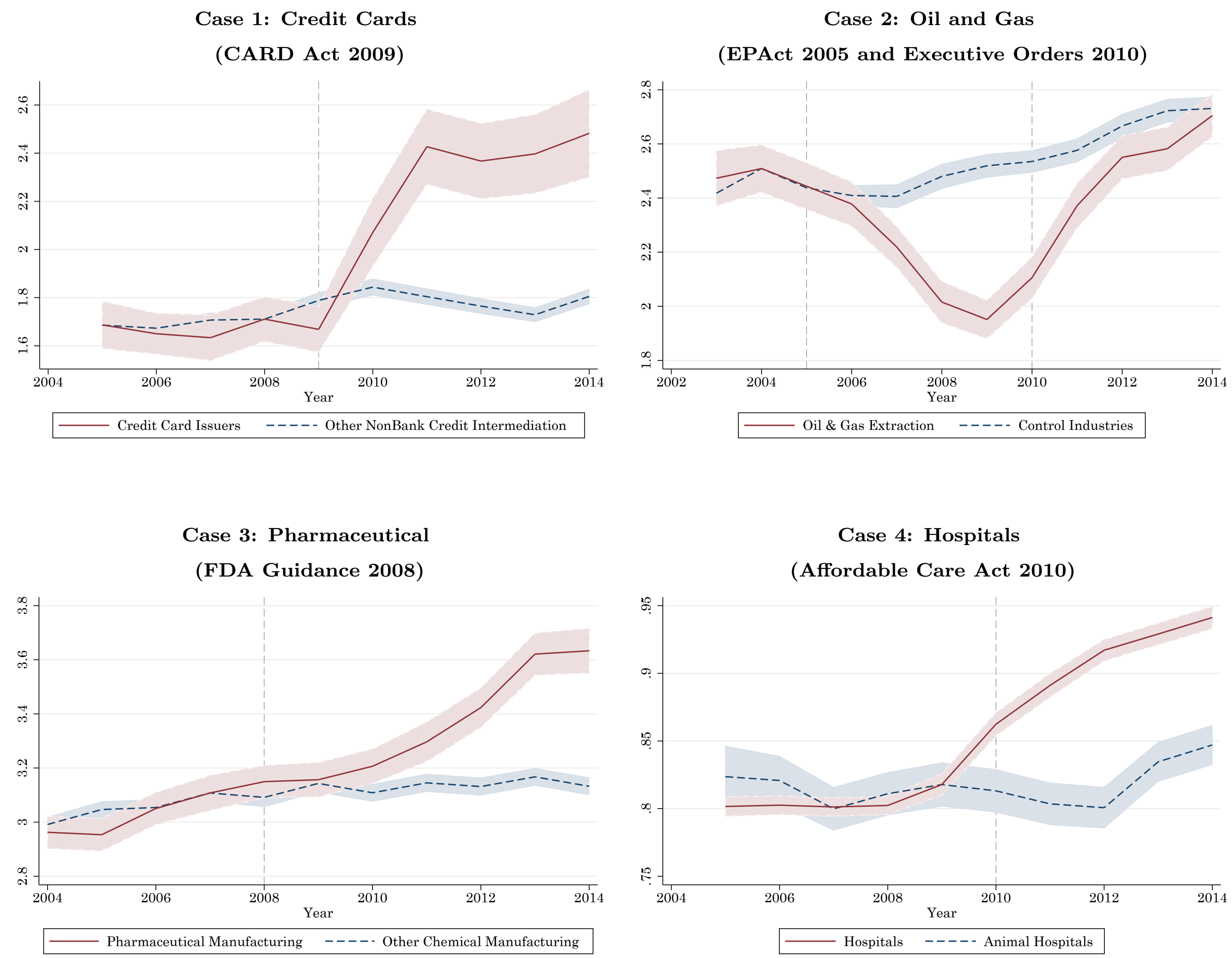


Figure 4: **Regulation Index Across States**

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.

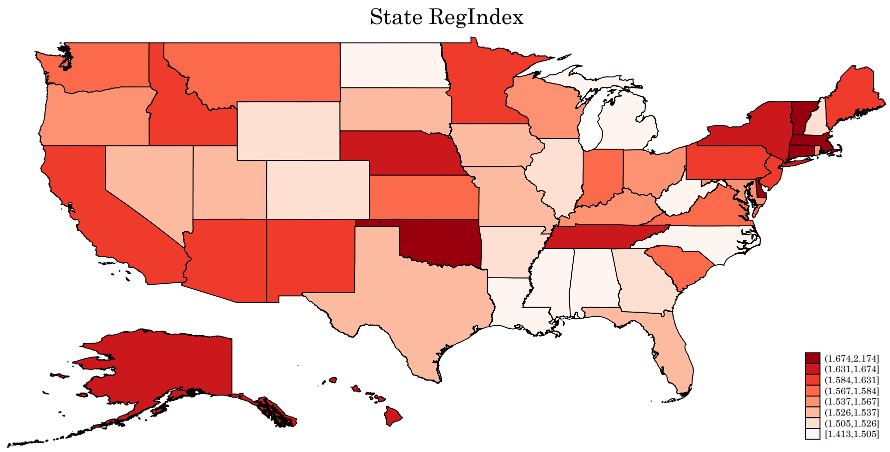


Figure 5: **RegIndex and Size**

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 in log scales.

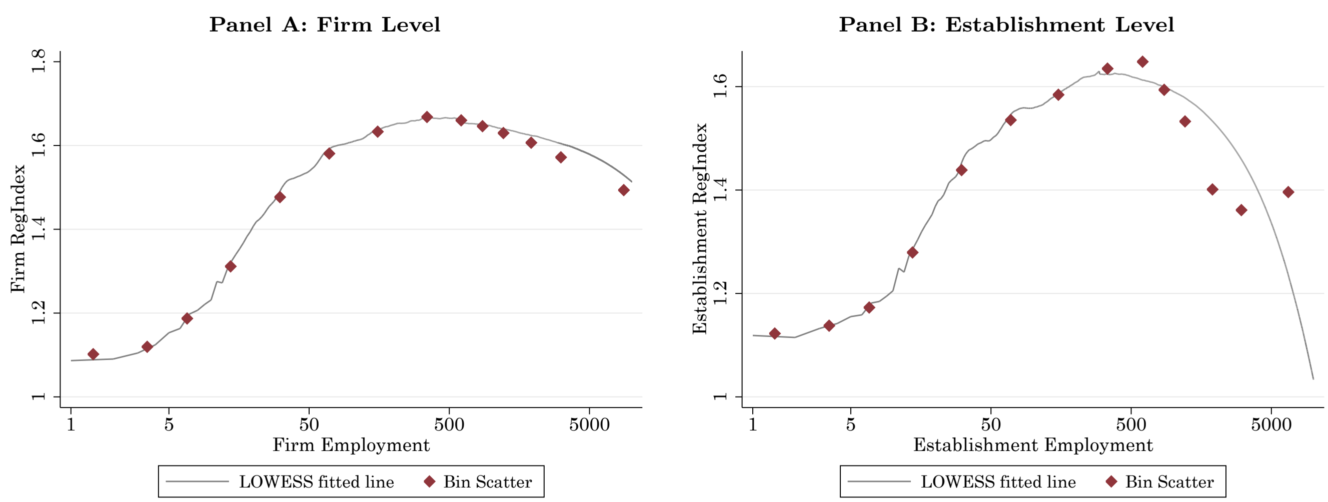


Figure 6: **RegIndex and Size by Sector**

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 in log scales.

Panel A: Firm-Level

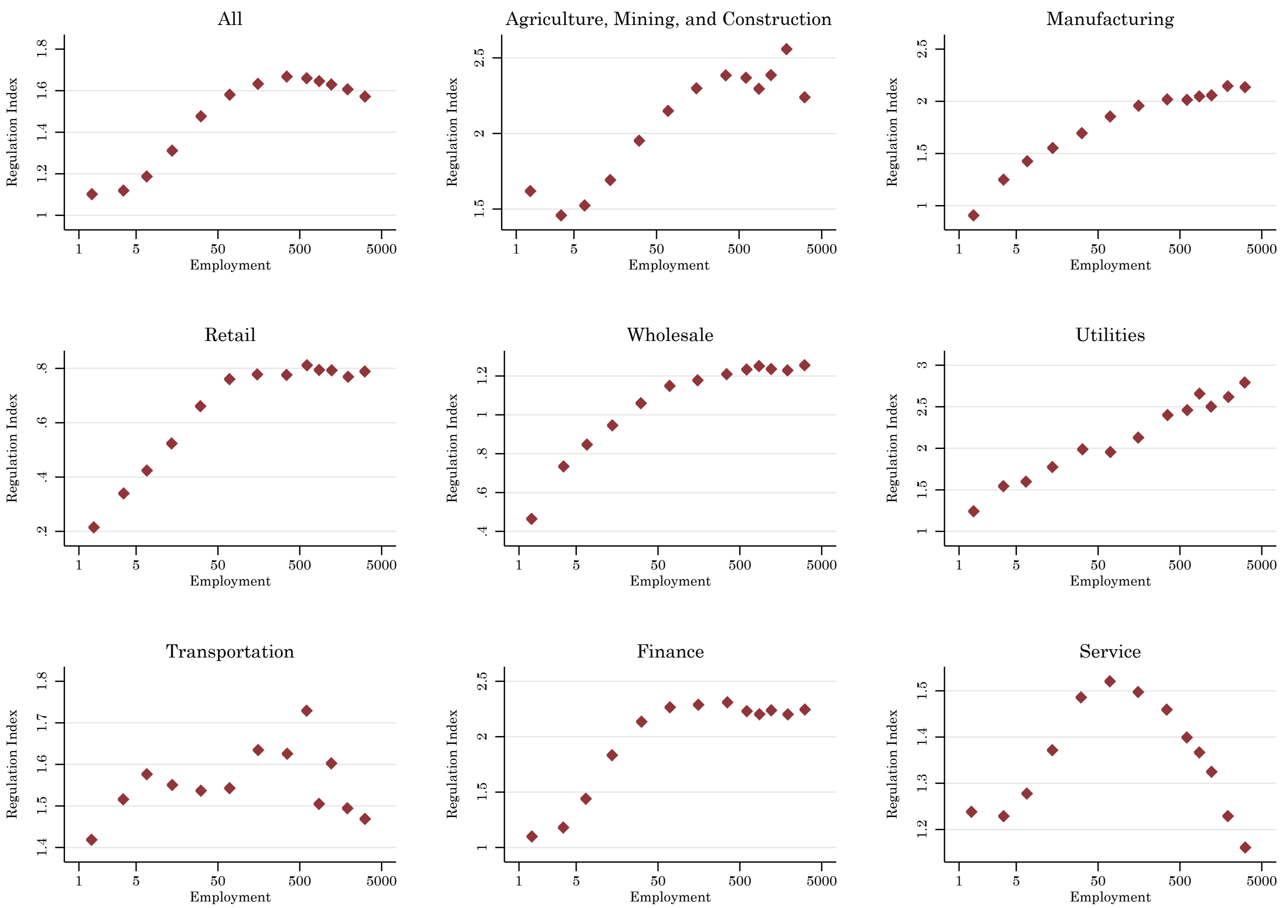


Figure 8: **RegIndex and Size by Sector**—*Continued*

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 in log scales.

Panel B: Establishment-Level

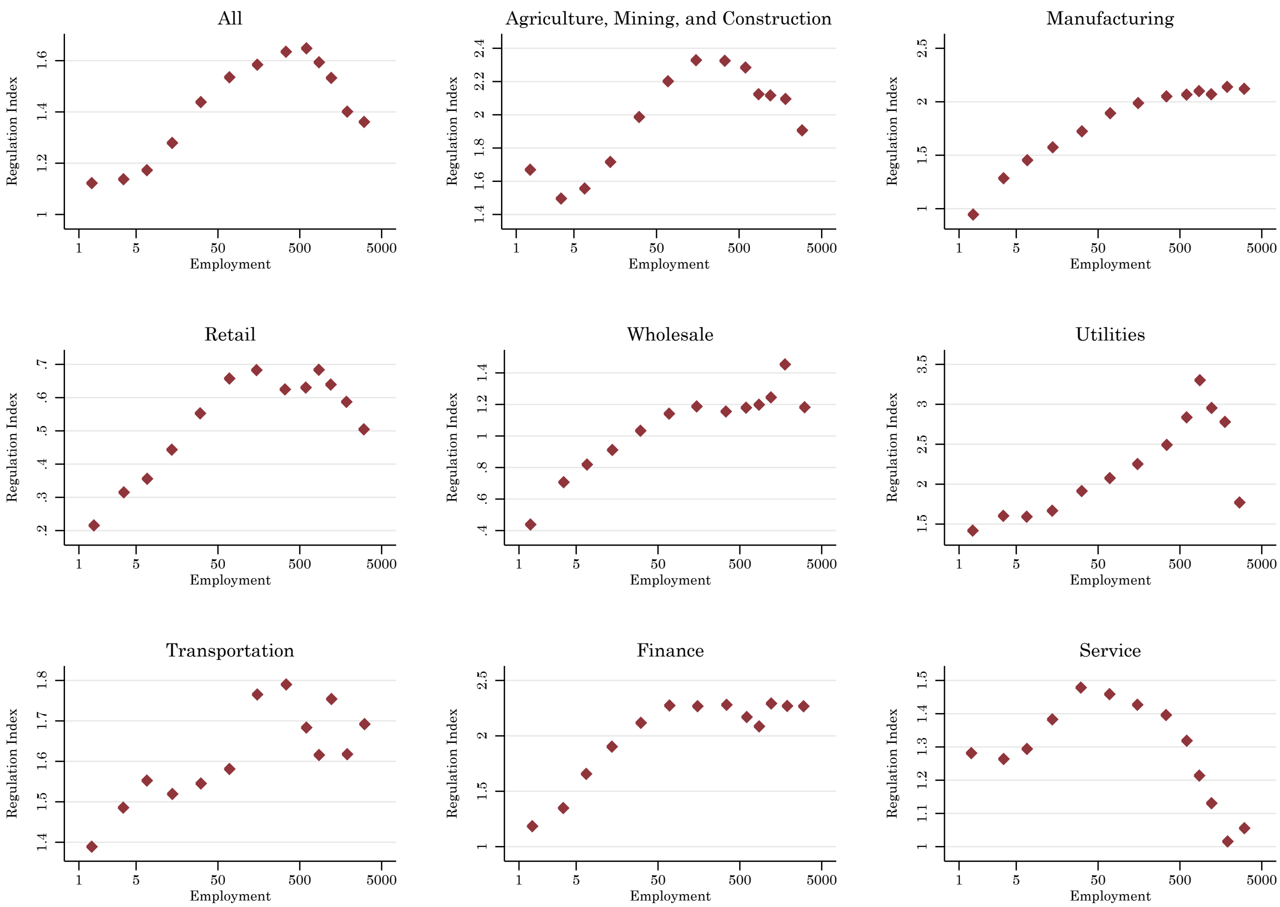


Figure 7: **Total-Input-Based RegIndex and Size**

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^{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 in log scales.

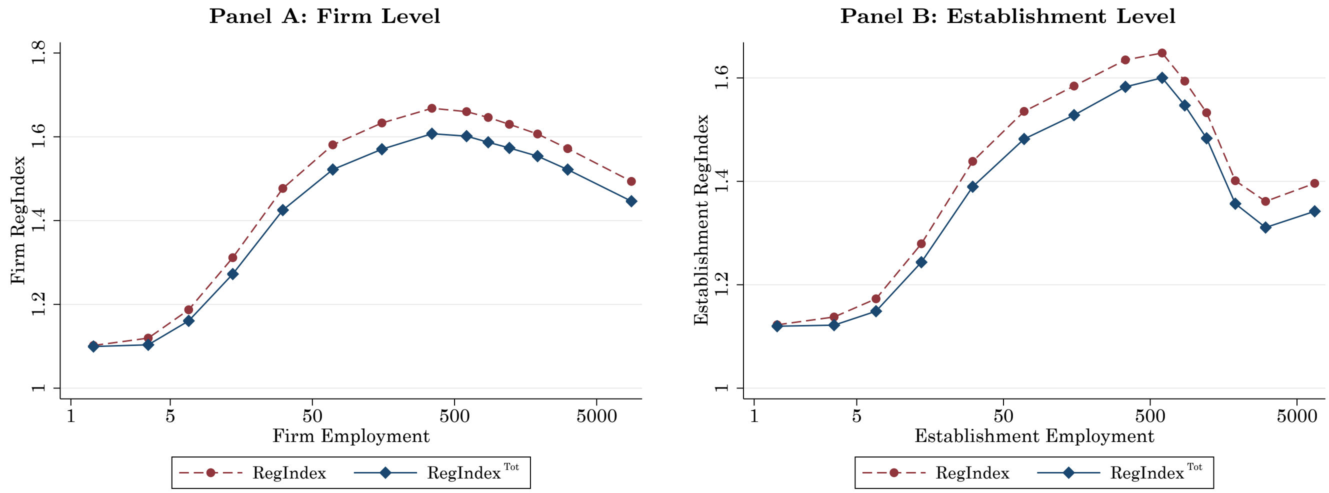


Figure 8: **RegIndex and Size: Multi-State vs. Single State Firms**

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 in log scale.

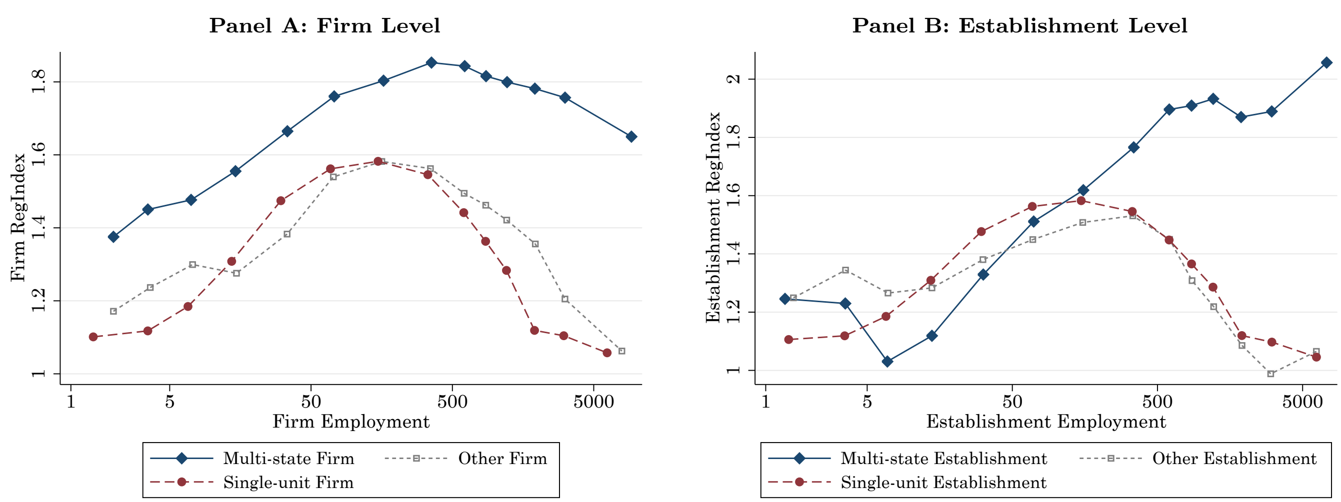


Figure 9: **Regulatory Compliance Specialization and Size**

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 in log scale.

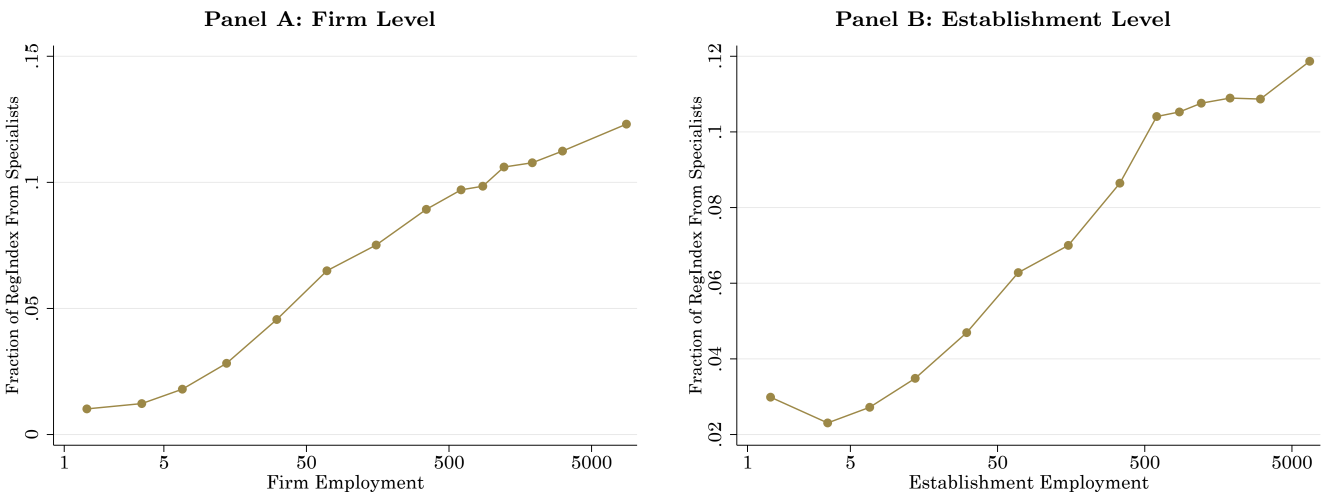


Figure 10: **Decomposing RegIndex: Regulatory Requirements Versus Enforcement**

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 in log scale.

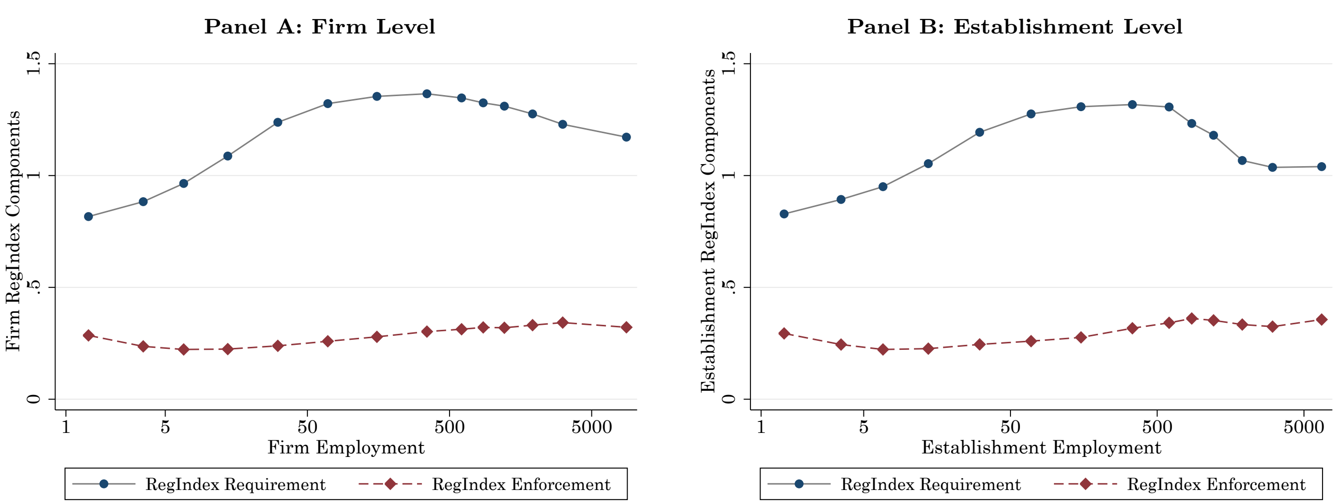


Table 1: **Summary Statistics**

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: **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 NAICS6. *, **, 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 4: **Regulatory Compliance Within Occupation**

This table reports the robustness of Table 3 by showing the relationship between firms' requirements on regulatory compliance skills and their size *within an occupation*. To do so, we 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). We next measure each job posting's textual content's requirement on regulation-related tasks by applying our exact methodology in Section 4 on BGT's definition of "skills." We take a simple average of the regulation-related dummy for each skill within a job posting to measure the job posting's average regulation-relatedness (in percentage). We then regress the job posting's regulation-relatedness on the firm's employment (in thousands) and the squared of the employment, while controlling for year-occupation fixed effects. Column (1) reports results for all SOC 6-digit occupations. Column (2) requires results for occupations that have positive regulation-task intensity (RTI) in our definition in Section 3 (see Table A.2), while Column (3) reports the results for 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 $-\frac{b}{2a}$.

	All Occ (1)	Occ (RTI>0) (2)	Occ (RTI = 0) (3)
Emp	0.722***	0.536***	0.858***
	(0.204)	(0.168)	(0.290)
Emp^2	-0.445***	-0.341***	-0.520***
	(0.131)	(0.102)	(0.181)
max	1.558***	2.073***	1.188***
	(0.092)	(0.071)	(0.128)
argmax	0.811***	0.787***	0.825***
	(0.048)	(0.048)	(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 5: **Changes in RegIndex and Shocks**

This table reports the results of regressing 3-year changes in a firm's log RegIndex on instrumental 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 instrumental variables for the NAICS 6-digit industry's enforcement shocks and regulatory-requirement shocks. Section 5 provides details on the construction of the instrumental 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	$\geq$ 750	1-19	20-399	400-749	$\geq$ 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

Online 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.[45]

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.[46] 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 down-weighting 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 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

[45]Our regulation-related tasks can be downloaded at www.miaobenzhang.com/Regulation_Related_Tasks.xlsx.
[46]Our findings are robust to alternative levels of downweighting second-tier keywords.

O*NET.[47] 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.[48]

B Details on Instrumental Variables Construction

The regulatory agencies employed for the construction of the instrumental 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 Corporation (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). Online Appendix Figure A.4 plots the time series for $\Delta \log enf_{kt}$ and $\Delta \log reg_{kt}$ for each main agency in our sample.

Our procedure identifies 227 out of 612 occupations of federal employees as "regulation-related." In the Online 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 shocks which is potentially a more volatile supply-side instrument for regulatory enforcement. Online Appendix Tables A.7-A.10 show virtually no substantive differences

[47] 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%.

[48] 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.

between results using the alternative shocks from our main results.

Our IV's exhibit desirable properties that are intuitive. First, Online Appendix Table A.3 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 shocks co-move over time. Online Appendix Figure A.4 traces the time series of the two shocks for each main regulatory agency. We observe that there is substantial independent variation in each of the two shock 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, Online Appendix Table A.4 shows that the overall correlation between the instrumental variables is 12 percent at the establishment level.

C Measuring Enforcement Employment

As a robustness check to using our main measure of enforcement shock 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.[49] 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 12 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.7-A.10 present results using

[49]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.

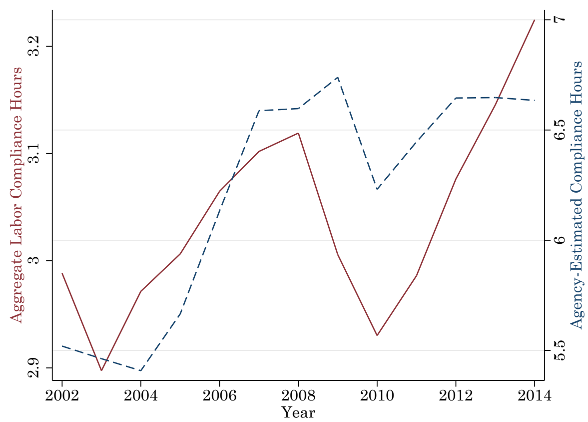
enforcement shocks, $iv(\Delta \log(p_{it}))$, based on this alternative measure.

D Additional Figures and Tables

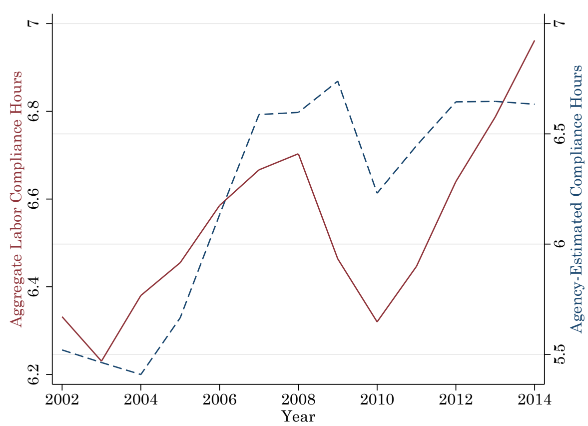
Figure A.1: **Validation of Regulation Index Using Agency-Estimated Compliance Hours—All Versions**

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: **RegData Across States**

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

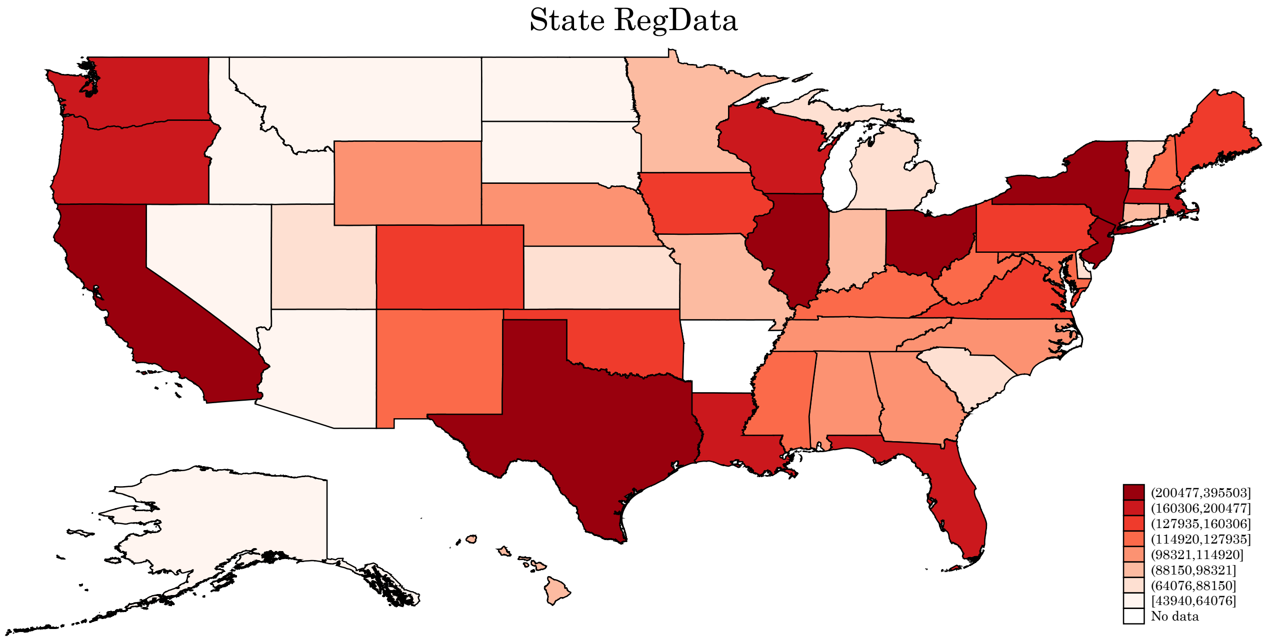
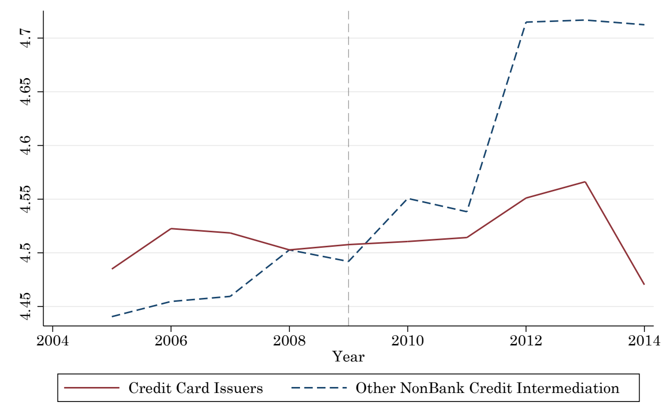


Figure A.3: **Case Studies of Industry Regulatory Shocks Using RegData**

This figure plots the response of industries' RegData measure to five industry-level regulatory shocks. 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 shocks 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.

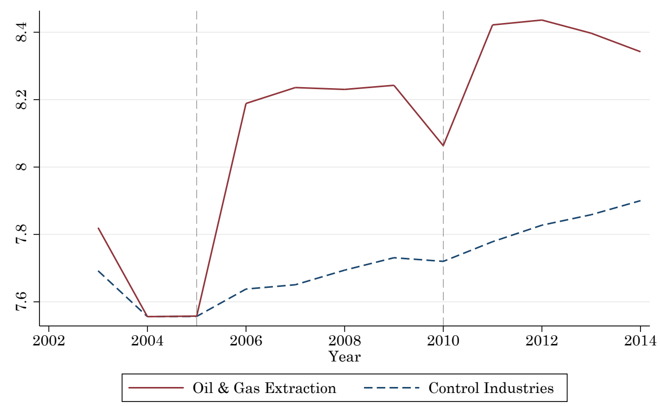
Case 1: Credit Cards

(CARD Act 2009)



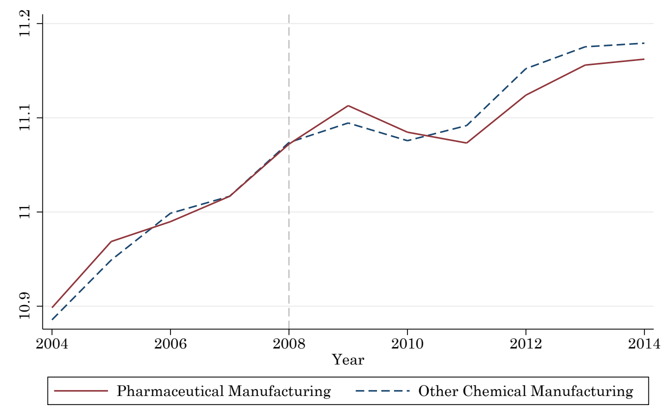
Case 2: Oil & Gas

(EPAct 2005 and Executive Orders 2010)



Case 3: Pharmaceutical

(FDA Guidance 2008)



Case 4: Hospitals

(Affordable Care Act 2010)

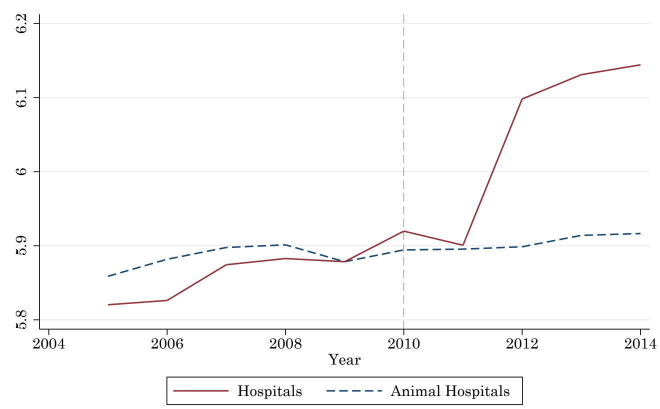


Figure A.4: **Shocks to Regulatory Requirements and Enforcement by Agency**

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

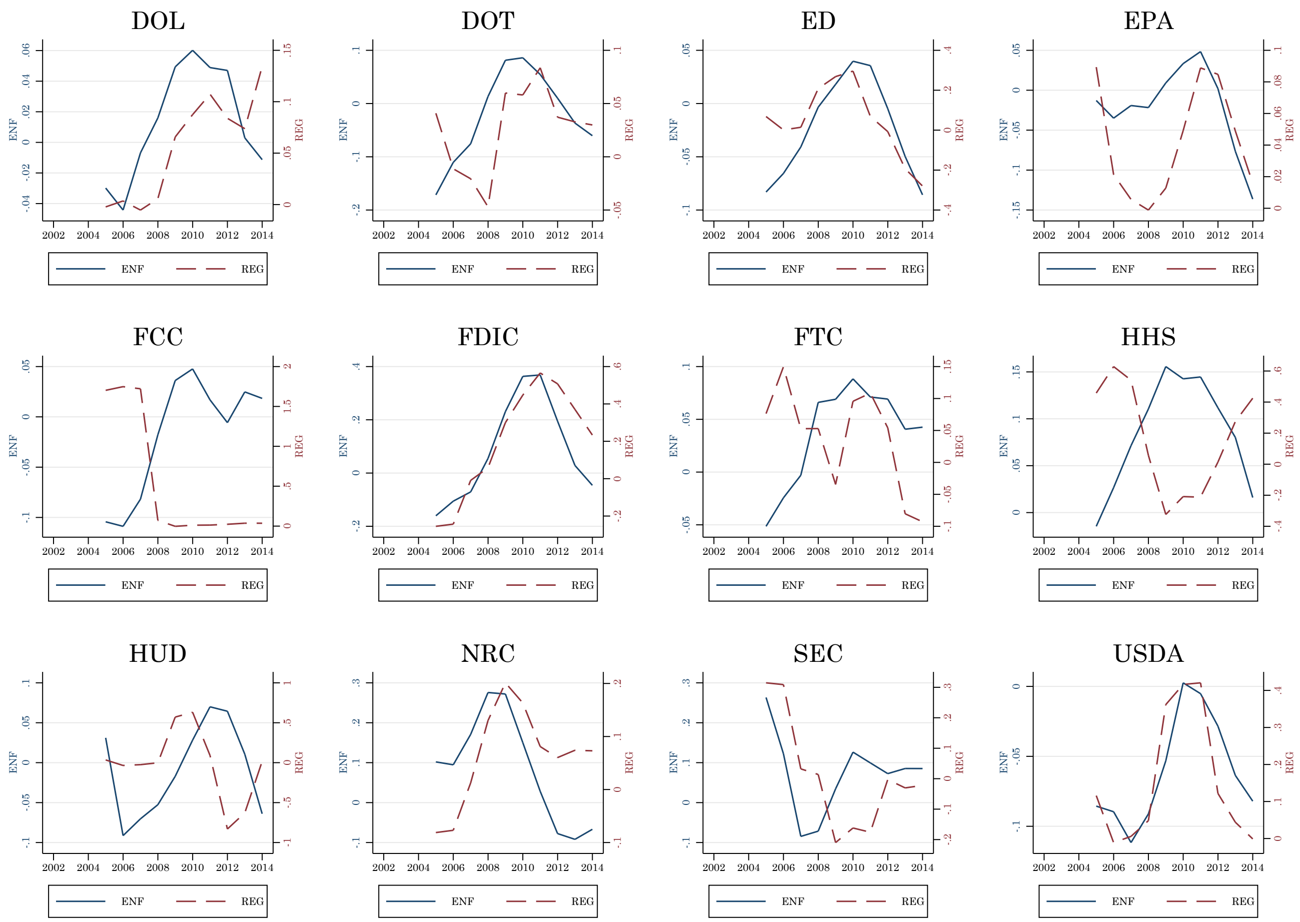


Figure A.5: **Levels of Regulatory Requirements vs. Enforcement Measures by Agency**

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

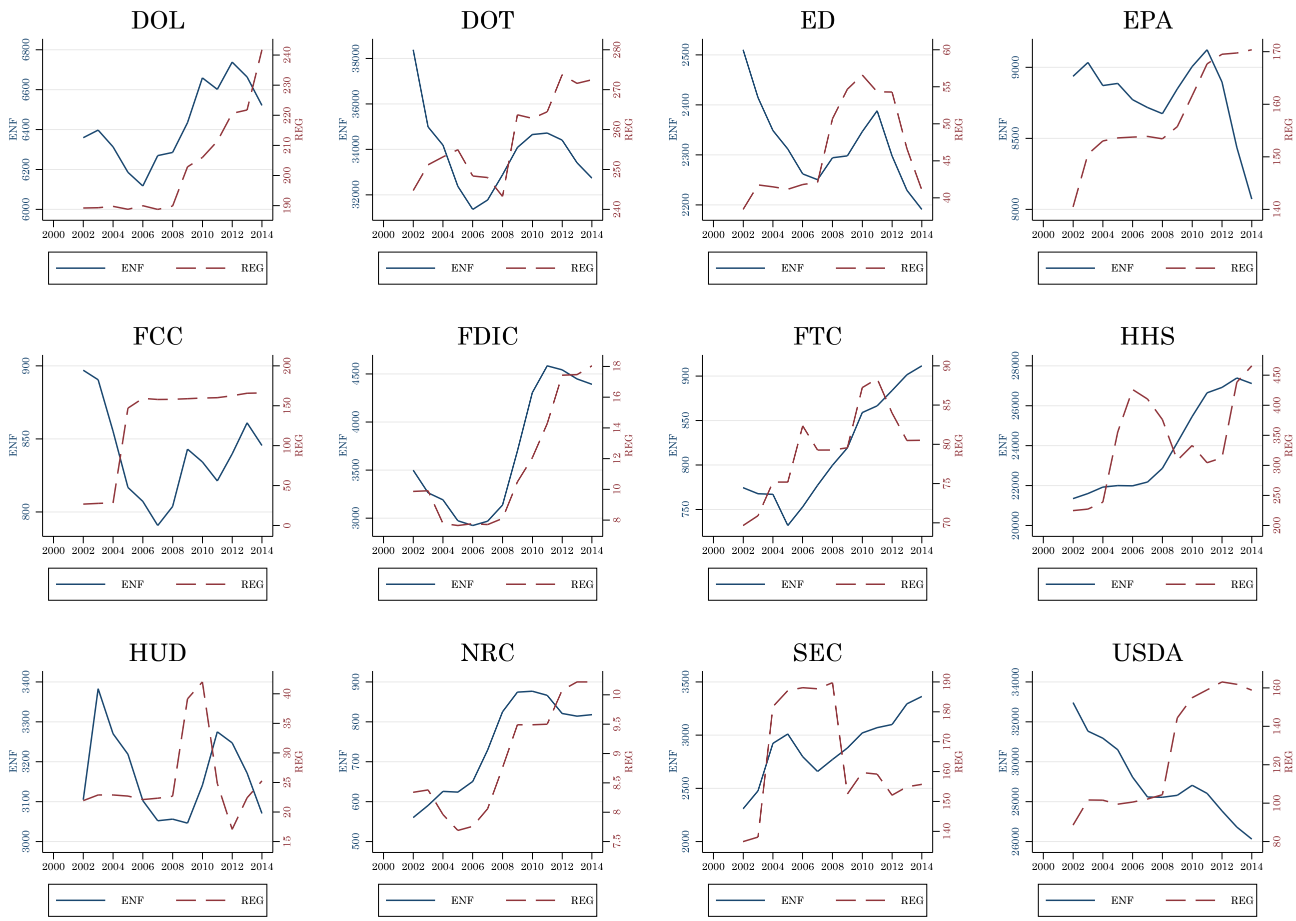


Table A.1: **Examples of Regulation-related Tasks**

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 Shocks**

This table reports the response of establishments' RegIndex to major industry regulatory shocks 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 shock 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:	Credit Cards	Oil and Gas	Oil and Gas	Pharmaceutical	Hospitals
Case:	CARDAct 2009	EPAct 2005	Executive Order 2010	FDA Guidance 2008	ACA 2010
	(1)	(2)	(3)	(4)	(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: **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.2: **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.3: **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.4: **Correlation of Instrumental Variables**

This table reports the Pearson correlation between the instrumental variable for enforcement shocks, $iv(\Delta \log(p_{it}))$, and the instrumental variable for regulatory-requirement shocks, $iv(\Delta \log(\tilde{R}_{it}))$, 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 instrumental 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.5: **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 instrumental variables in Table 5 in each NAICS 1-digit sector. Equation (7) provides the regression specification. See Table 5 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)
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.5: **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 instrumental variables in Table 5 in each NAICS 1-digit sector. Equation (7) provides the regression specification. See Table 5 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.6: **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 instrumental variables in Table 5 in each NAICS 1-digit sector. Equation (7) provides the regression specification. See Table 5 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.6: **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 instrumental variables in Table 5 in each NAICS 1-digit sector. Equation (7) provides the regression specification. See Table 5 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.7: **Robustness—Regulatory Requirements vs. Enforcement: An Instrumental Variable Approach**

This table reports the robustness check of Panel A of Table 5 by reconstructing the instrumental variable for enforcement shocks using only enforcement-related regulatory occupations in each agency. See Online Appendix C and Table 5 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.8: **Robustness—Regulatory Requirements vs. Enforcement in Subsamples by Size**

This table reports the robustness check of Panel B of Table 5 by reconstructing the instrumental variable for enforcement shocks using only enforcement-related regulatory occupations in each agency. See Online Appendix C and Table 5 for details.

	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)
$iv(\Delta\log(p_{it}))$	0.103	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.9: **Robustness—Regulatory Requirements vs. Enforcement for Firms in Each Sector**

This table reports the robustness check of Table A.5 by reconstructing the instrumental variable for enforcement shocks using only enforcement-related regulatory occupations in each agency. See Online Appendix C and Table A.5 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.10: **Robustness—Regulatory Requirements vs. Enforcement for Establishments in Each Sector**

This table reports the robustness check of Table A.6 by reconstructing the instrumental variable for enforcement shocks using only enforcement-related regulatory occupations in each agency. See Online Appendix C and Table A.6 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)