I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17820

Exposure to Regulation and Income Inequality in Local Labor Markets: Evidence from the U.S. over the Past Half-Century

Andrey Stoyanov
Nick Zubanov

APRIL 2025

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17820

Exposure to Regulation and Income Inequality in Local Labor Markets: Evidence from the U.S. over the Past Half-Century

Andrey Stoyanov
York University

Nick Zubanov
University of Konstanz and IZA

APRIL 2025

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Exposure to Regulation and Income Inequality in Local Labor Markets: Evidence from the U.S. over the Past Half-Century*

Existing evidence points to a positive correlation between specific regulations and income inequality at a country or regional level, but little is known about how overall regulatory burden affects inequality at the local labor market level. Our study fills this gap by measuring local exposure to regulation from the industry-relevant articles of U.S. Code of Federal Regulation linked to local industry employment structure in 741 commuting zones (CZs) in the U.S. over the period 1970-2019. Relating our exposure to regulation measure to the CZ-level income inequality, computed from the Census records, we find that heavier regulation is followed by higher income inequality, lower average income and higher unemployment in the affected CZs. The implied effect estimates are sizeable and robust to various checks. We contribute to inequality research by identifying previously unknown, local effects of regulation on income inequality, exploring mechanisms through which they may occur, and demonstrating how available data can be used to produce more granular measures of exposure to regulation.

JEL Classification: L5, D63, E24

Keywords: regulation, income, inequality, employment, local labor market

Corresponding author:
Nick Zubanov
Department of Economics
University of Konstanz
Universitätsstrasse 10
78464 Konstanz
Germany

E-mail: nick.zubanov@uni-konstanz.de

* The authors thank Ontario/Baden-W¨urttemberg Academic Exchange Program for financial support of their collaboration.

1 Introduction

Like many other developed economies, the U.S. has seen an increase in the amount and scope of regulation (McLaughlin and Sherouse, 2019) coinciding with growing income inequality (Saez and Zucman, 2020). Are these two trends connected beyond coincidence? Various measures of regulation have been found negatively correlated with a number of (macro)economic indicators, including output (e.g. Besley and Burgess, 2004; Dawson and Seater, 2013; Coffey et al., 2020), startup activity (Djankov et al., 2002), and employment and earnings (Botero et al., 2004; Feldmann, 2009; Bailey and Thomas, 2017). Its relationships with startup activity, employment and earnings suggest that regulation may also matter for income inequality, which is the topic of this study.

The above correlations notwithstanding, the effect of regulation on income inequality is theoretically unclear and warrants empirical investigation. On the one hand, regulation may reduce inequality by creating institutional and legal frameworks that facilitate equal access to, and benefits from, gainful economic activities and limit discrimination or market power abuse (Baldwin et al., 2011). On the other hand, as it increases the costs of compliance (Bombardini et al., 2024), regulation may reduce profitability, growth and demand for labor by existing firms (Besley and Burgess, 2004), especially small firms that tend to bear higher per-worker compliance costs, or make new firm entry or self-employment more difficult (Chambers and O'Reilly, 2022a). Lower demand for labor and higher barriers to entrepreneurship caused by regulation may, in turn, exacerbate inequality by limiting access to valuable economic opportunities, especially to the more economically disadvantaged for whom those barriers are relatively high.

Existing studies of the relationship between regulation and income inequality, most of which find a positive link between the two, tend to focus on specific aspects of regulation and to be carried out at the country or state level. Starting with the scope of regulation, there is research on economic effects of regulation in labor relations (Besley and Burgess, 2004; Koeniger et al., 2007; Calderón and Chong, 2009), banking and finance (Beck et al., 2010;

Agnello et al., 2012; Delis et al., 2013; de Haan and Sturm, 2017) or environment (Jha et al., 2019; Huang and Yao, 2023), but there is less work on the effects of the overall regulatory burden (Chambers and O'Reilly, 2022b), or "red tape" which is the focus of this study.[1] Yet, the growing public concern over excessive red tape (TheEconomist, 2022) calls for research on its consequences, which may not be captured in the effects of legal acts targeting specific activities like environmental pollution, labor relations, or banking. While many specific regulatory initiatives may improve welfare by addressing specific market failures, the red tape that is meant to support their implementation is likely to have wider, and more negative, economic effects by increasing transaction costs across the board.

Turning to the level of analysis, most of the existing studies are carried out at the country or state level (Jha et al. (2019) and Huang and Yao (2023) studying the inequality effects of environmental regulations at the county and community level are rare exceptions). Country- or state-level studies cannot by design capture the effects of regulation at more disaggregate levels, such as local labor markets, without accounting for which an analysis of distributional effects of regulation would be quite incomplete. Local labor market dimension of inequality is important because measures of inequality as well as its economic effects may differ by level of analysis, owing to geographic differences in income levels, costs of living (Moretti, 2011; Black et al., 2014), and industry mix.

Zooming on the industry mix, local labor markets may react differently to country-wide regulatory shocks because of their varying exposure to industry-specific regulation. A useful analogy here is the heterogeneity in the effects of the global "China shock" on local employment and earnings depending on the local presence of industries affected by U.S.-China trade liberalization found by Autor and Dorn (2013) and Autor et al. (2013, 2016). As regulation can be industry-specific (Al-Ubaydli and McLaughlin, 2017), aggregating labor

[1] One example from tens of thousands pieces of regulation covered in our analysis is the *Affiliate Marketing Rule* issued by the Federal Trade Commission (Part 680, Subchapter F, Chapter I, Title 16 of the U.S. Code of Federal Regulations) for the purpose of implementing Section 241 of the *Fair and Accurate Credit Transactions Act* (Public Law 108-159). Our analysis does not cover laws passed by the legislative branch, such as the *Fair and Accurate Credit Transactions Act* itself, or legal determinations related to this Act within the common or case law.

markets with different industry structures into larger areas would blur their varying reactions to a given amount of nationwide regulation. Another reason for studying inequality at the local labor market level is that public attitudes towards inequality are driven by inequality perceptions (Gimpelson and Treisman, 2018; Kuhn, 2019; Knell and Stix, 2020) which are informed predominantly by local comparisons (Newman et al., 2015, 2018).

To our knowledge, there is no study on the effects of exposure to regulation broadly defined on inequality at the local labor market level. We fill this research gap by exploring the relationship between local exposure to regulation and locally measured Gini index of income inequality (our primary outcome measure) and income and unemployment (secondary outcomes). Following Autor and Dorn (2013) and Autor et al. (2013), we define local labor markets as commuting zones (CZs) – areas empirically identified by Tolbert and Killian (1987) and Tolbert and Sizer (1996) that have strong home-to-work commuting ties within, and weak ties between them. There are 741 CZs in the U.S. whose geographical connection to counties enables us to calculate socio-economic outcomes at the CZ level from the U.S. Census and American Community Survey data from 1970 onwards. We believe CZ is an appropriate unit of analysis for our purposes: localities defined by shared labor markets, of which commuter traffic is an indicator, seem more suitable for studying labor market consequences of regulation than those defined by administrative boundaries.

We compute our CZ-level measures of exposure to regulation using data from the RegData Project at George Mason University's Mercatus Center, available at https://www.quantgov.org. The information we use has been gathered automatically from each article of regulation in the U.S. Code of Federal Regulations (CFR) in each year since 1970 with the text search code developed by Al-Ubaydli and McLaughlin (2017). Their code counts restrictive terms, namely, "shall", "must", "may not", "prohibited" or "required", in each article and the article's relevance to each specific industry, based on the frequency of industry-specific terms in it, thus enabling a straightforward computation of the aggregate regulation applicable for each industry-year. We calculate measures of exposure to regulation for each CZ-year as averages of the industry-year measures, weighed by industry employment shares in a given CZ-year. Our primary measure of exposure to regulation is the total number of restrictive words relevant to a given CZ-year, and our secondary measure is their share in the total regulatory word count.

To preview our results, we find that CZs whose industry employment structure renders them more exposed to current regulation have higher income inequality, lower average income and higher unemployment a decade after. Our baseline regression results imply that, in a hypothetical pair of otherwise identical CZs, the one with a one standard deviation (sd.) higher regulatory burden applicable to it would have a one sd. higher Gini index, 0.05 sd. lower income, and 1/3 sd. higher unemployment rate a decade later. Thus, the implied effects are substantial and support the emerging consensus for reducing red tape to address economic inefficiencies that bring negative social consequences such as inequality.

We further advance the research on the effect of regulation on inequality by providing a rich characterization to our main results which adds depth to our understanding of how regulation may affect the economy. Specifically, first, the effects we find last between one and two decades, suggesting that the costs of regulation are not short-lived but markets eventually adjust to it. Second, the effects are driven primarily by newly passed regulations, presumably requiring more adjusting to, rather than changes in the existing red tape that is more likely to have been internalized already. Finally, the effects differ by the activity areas that regulation targets. While most of the existing literature focuses on specific regulations in the areas of environment, finance or labor, we find that the overall regulatory burden, not attributable to any specific area, is more prevalent and costly than area-specific regulations.

Our results are not driven by possible endogeneity of regulation to local conditions through local industry employment structure. Instrumenting contemporary industry employment shares with their lags and baseline (1970) levels produces qualitatively similar results to those from our main regression specification. Another potential source of endogeneity is locally-based special interests that may influence nationwide regulation. To account for these, we rerun our analysis excluding CZs with relatively few dominant industries and CZs that make up relatively large shares in nationwide employment in specific industries, finding that our results are quite robust to both manipulations.

We validate our approach to measuring local exposure to regulation with a series of placebo tests in which we rerun our analysis multiple times on simulated, rather than actually observed, industry relevance of regulations and industry shares in local employment. Our placebo tests reveal that the relationships we find between local exposure to regulation and local labor market outcomes depend critically on the local relevance of regulation as determined by its industry relevance and local employment structure. Specifically, we find through simulations that random deviations in either of these aspects from their true, locally observed values destroy the links between regulation and local labor market outcomes.

As an extra empirical justification for the local level of our analysis, we rerun our regressions at the more aggregate, state level. This exercise brings much weaker and less significant results, suggesting that one may overlook important consequences of local exposure to regulation by over-aggregating.

Our study makes three contributions to the existing literature. First, we extensively document hitherto unknown effects of exposure to regulation on income inequality at the local labor market level. The value of this contribution is in showing that local effects of regulation on inequality are stronger than those found in existing studies carried out at higher levels of aggregation. One cannot infer these local effects from a given aggregate effect, owing to varying local exposure to regulation that we find. Yet it is the local manifestations of income inequality that tend to be registered more clearly and acted upon more strongly. The second, and related, contribution of our study is in shedding light on a channel through which regulation may affect local income inequality – its relevance to industries present in local employment. A given piece of regulation is the more consequential for a given labor market, the more relevant it is to industries prevalent in that market. Our empirical results are consistent with this hypothesis.

Our third contribution is methodological, and lies in demonstrating how existing data sources, RegData and Census, can be leveraged to compute more finely-grained measures of income inequality and regulation and their relationship than what existing research has done so far. RegData is a valuable source of information on regulation that has supported numerous studies, some of which are related to inequality. So far, however, the level of analysis has been industry-occupation-year (e.g., Bailey et al., 2019; Mulholland, 2019) or state-year (e.g., Chambers et al., 2019; Chambers and O'Reilly, 2022b; Choudhury, 2023). We hope our work will stimulate further research on the local consequences of regulation for inequality and other socioeconomic indicators using RegData.

2 Data

This section describes our data and measures of local exposure to regulation, and reports descriptive statistics.

Communing Zones

Our unit of analysis is local labor market. We follow research on the consequences of macroeconomic shocks (Autor and Dorn, 2013; Autor et al., 2013, 2016) in defining local labor markets as commuting zones (CZs) – geographic areas characterized by strong commuting ties within them, and weaker ties in between. CZs were first delineated by Tolbert and Killian (1987) and updated by Tolbert and Sizer (1996) based on 1990 commuting flows. We use data produced by the latter update which identifies 741 commuting zones in the U.S.

Outcome variables

Given our focus on income inequality, our main outcome variable is the Gini index of income inequality, which is a commonly used inequality metric. Our secondary outcome variables are unemployment rate and log income level. We compute our outcome variables at CZ-year level using individual data from the U.S. decennial censuses from 1970 to 2000, and from the American Community Surveys for 2010 and 2019. We chose to use data from

2019 instead of 2020 to abstract from the likely consequences of the COVID-19 pandemic for local labor markets, but we still refer to the most recent nine-year period in our sample as a decade for simplicity. Aggregating individual data from the Census up to the CZ level is nontrivial, since CZs do not always coincide with counties (the level at which the place of residence is recorded in the Census). We therefore employ crosswalk tables from Autor et al. (2013) that map counties into CZs with weights based on fractions of county population commuting within a given CZ, estimated from 1990 commuting flows.[2]

Measuring local (CZ-level) exposure to regulation

While there are commonly accepted income inequality metrics such as the Gini index, which we use, the vastness and complexity of regulation defies common definitions. In this study, we focus on regulation in the form of rules issued by the U.S. government agencies acting on the mandate from the legislature for the purposes of implementing and enforcing compliance with the laws passed by the legislature. One can think of these "red tape" regulations as part of administrative law.

We quantify regulation so defined using RegData database accessible at https://www.quantgov.org. RegData contains information extracted from texts of the annual editions of the U.S. Code of Federal Regulations (CFR) with the text searching code developed by Al-Ubaydli and McLaughlin (2017). CFR consists of 50 Titles, each containing regulations focusing on a specific topic (e.g., Agriculture in Title 7, or Employees's benefits, Title 20). Each title is further broken into chapters, containing rules for a specific government agency, and parts, regulating their particular functions. We treat CFR part as the elementary unit of regulation

[2] For an illustration of this approach, consider an economy with two equally populated counties, A and B, two industries, I1 and I2, and two CZs, Z1 and Z2. County A has the average income 8, 10% of its workforce are employed in I1 and 90% in I2, 30% of its workforce commute within Z1 and 70% within Z2. County B has the average income 10, employs 80%/20% of its workforce in industry I1/I2, and 40%/60% of its population commutes within Z1/Z2. The average income in Z1 is then $\frac{8\times0.3+10\times0.4}{0.3+0.4} \approx 9.14$, and in Z2 it is $\frac{8\times0.7+10\times0.6}{0.7+0.6} \approx 8.92$, reflecting the relative prevalence of the population from the poorer county A in commuting zone Z2. The CZ-level share of industry I1 employment is $\frac{0.1\times0.3+0.8\times0.4}{0.3+0.4} \approx 0.5$ in Z1 and $\frac{0.1\times0.7+0.8\times0.6}{0.7+0.6} \approx 0.42$. (The shares for industry I2 employment are one minus the above.) Reassuringly, the economy-level averages for income and employment shares are the same whether one aggregates country- or CZ-level averages.

in our analysis. There are a total of 13,679 distinct CFR parts in our data sample.

The first key input from RegData in our calculations is the *number of restrictive words* in the text of CFR part k in year t's CFR, denoted W_{kt} in equation (1). The restrictive words are "must", "may not", "shall", "prohibited" or "required", which are commonly used in legal documents to prohibit an action or impose compliance. Descriptive statistics in Table 1 for the nearly 37 thousand CFR part-years present during 1970-2019 show that an average part contains 117.2 restrictive words in an average year, which comprise about 1% of its word count. 5% of part-years contain no restrictive words, and a further 5% have one or two restrictive words in their entire text. Both the number and the share of restrictive words vary by CFR part, reflecting the variation in the length as well as restrictiveness of CFR parts, but the share of restrictive words varies much less in comparison with their number.

The second key input from RegData is the *relevance weight* of part k to industry i in year t, denoted a_{ikt} in (1). We use Al-Ubaydli and McLaughlin (2017)'s measure of industry relevance, calculated based on the frequency with which an industry-specific collection of search terms appears in the text of a given part in a given year. Importantly, the search terms determining industry relevance do not contain any of the restrictive words listed above, so a given part would not be scored as more or less relevant for a given industry based on its restrictiveness.[3] Descriptive statistics in Table 1 for over 3 million CFR part-industry-years show that an average part is barely relevant for an average specific industry in an average year. The industry relevance weights distribution is highly skewed: for instance, 20% of observations have industry relevance weight at or above 0.1% (that is, one in 1,000 words in a part-year is a search term for a specific industry), and only 4% are at or above 1%. 7% of observations have zero industry relevance weight, and 20% have exceedingly small industry relevance weight of 0.001% or less.

To calculate our measures of exposure to Federal regulation at the CZ level, we combine

[3]Appendix B in Al-Ubaydli and McLaughlin (2017) contains this and other details on the computation of industry relevance weights. We take their weights without changing or otherwise adapting their text search code.

the number of restrictive words in a CFR part (W_{kt}) and its industry relevance weights (a_{ikt}) from RegData with industry employment shares computed from the Census data. Our primary measure of local exposure to regulation is the number of restrictive words from all CFR parts weighed by their relevance weights for each industry and then further weighed by the industries' employment shares in a given CZ:

$$RW_{ct} = \sum_i s_{cit} \sum_k a_{ikt} W_{kt}, \tag{1}$$

where s_{cit} is the share of industry i in total employment within CZ c in year t obtained from the Census data.[4] The quantity RW_{ct} in equation (1) can be interpreted as the total number of restrictive words in the CFR relevant to a certain CZ given its industry employment structure.

Because higher number of restrictive words may reflect greater volume as well as restrictiveness of regulation, we also calculate the share of restrictive words in the total volume of regulation applicable to a given CZ given its industry employment structure:

$$RS_{ct} = \frac{\sum\limits_i s_{cit} \sum\limits_k a_{ikt} W_{kt}}{\sum\limits_i s_{cit} \sum\limits_k a_{ikt} TW_{kt}}, \tag{2}$$

where TW_{kt} is the total word count in part k in year t.

Limitations of our measurement approach

Before we proceed with presenting our results, two limitations of our approach should be noted. First, word count-based regulation metrics from RegData will inevitably fail to fully capture the nuances of regulation or the degree of its strictness (Bombardini et al., 2024), especially when new legislation attempts to deregulate by giving businesses more freedom of choice (the extra "can", "permitted" or "allowed" in the new regulation texts would not affect our measure). Despite this limitation, however, we observe and report below that local labor market outcomes are meaningfully correlated with our measures of

[4] Because the industry of employment in the Census data is recorded using a different industrial classification system (NACE) from the one used in measuring regulatory burden (NAICS), we convert the NACE industry codes in the Census data into NAICS codes using the Industry Code Crosswalks provided by the US Census Bureau.

exposure to regulation, and that these correlations disappear once we use placebo measures. Therefore, despite their imperfections, RegData metrics are informative of the strictness of regulation.

Second, our measures RW and RS will underestimate the true local exposure to regulation when there is regulation applicable to all industries. For instance, if the regulation text contains restrictive words related to “every firm” or “every worker”, this part will have zero industry relevance. While such restrictions are clearly relevant, they will not affect the measured exposure to regulation. However, not counting universally applicable regulation will not distort our regression estimates when we include time fixed effects, which we do (Section 3). In consequence of including time fixed effects, our analytical approach will produce estimates of the *differential* effect of a given amount of regulation between CZs that differ in their exposure to it through their different industry employment mixes. This differential effect does not have to equal the average effect of the same amount of regulation on all CZs, which is not separately identifiable from the time fixed effects: it may be that regulation has a neutral or even positive socio-economic effect on average, but CZs more exposed to it are worse off than those less exposed. In other words, we do not know how costly regulation is on average, but we can say how much costlier it is to one local labor market that is more exposed to it than another, otherwise identical, one.

This said, we are not alone in our approach with all its benefits and limitations. Any empirical study that measures exposure to regulation from RegData and includes time fixed effects in its regression specification (e.g., Chambers et al., 2019; Chambers and O'Reilly, 2022b; Choudhury, 2023) would have the same limitations as the one outlined here. The progress our study makes lies in further localizing exposure to regulation and in tracing its differential effects by locality, rather than in estimating the average effect of regulation, which would require a plausibly exogenous variation in the universally applicable regulation over time.

3 Results

3.1 Descriptive statistics

The 741 CZs observed over 6 time periods spanning 5 decades (years 1970, 1980, 1990, 2000, 2010, and 2019) make 4446 CZ-year observations. Table 1 shows that an average CZ has the unemployment rate of 6% and log average income of 10.1 in an average year. The income is rather unequally distributed within CZs, with the Gini index of income inequality at 0.428 in the average CZ-year. Put differently, if our CZ-years were countries, the average one in our sample would be among the top 25% in the world in terms of income inequality. CZ-years vary considerably in income level and unemployment rate. Although the Gini index varies less strongly, analysis of variance reveals significant differences across CZs and years, with a clear upward trend in CZ-level income inequality.[5] Most of the variation in log average income is within CZs across years, but the spatial and temporal variations in the Gini index and unemployment rate are about equal in magnitude.

Its industry employment structure together with industry relevance of CFR regulation render an average CZ exposed to a total of 12,672 restrictive words in an average year (measure RW, defined in equation 1). Restrictive words make up about 1% of the applicable regulations' total word count (measure RS in equation 2). Their number varies greatly, with its standard deviation (6,616) more than half its mean. Most of the variation in exposure to regulation is over time: as we show later in Section 3.4, changes in regulation add about 5000 restrictive words in an average CZ-decade. Removing the time trend in the exposure to regulation still leaves considerable variance across CZs (1680, and increasing over time), which reflects growing differences in industry employment structures between CZs. The number and share or restrictive words are positively correlated ($r = 0.54$), implying that

[5] The seemingly limited variation in the Gini index by CZ-year is about a quarter of the variation in the Gini index across 214 countries and territories of the World over the same time period (1970-2019), whether measured in terms of standard deviation or inter-quartile range (Source: own calculations based on World Bank data, tinyurl.com/ytufm76a). It is thus a relatively large variation, especially considering higher possibility of local segregation by income level that would suppress differences in income inequality between CZs.

a more voluminous regulation also tends to be more restrictive. However, the standard deviation of the restrictive word share is much smaller, about 5% of its average, than that of the restrictive word count, which is about 50% of its average.

Table 1 about here.

CZ-level exposure to regulation strongly correlates with income inequality, whether measured along geographical or temporal dimension. Figure 1 pictures the variation in the share of restrictive words in the applicable regulation (equation (2)) across CZs in 2000, a year quite representative of our data (Panel A), and the contemporary variation in the Gini index of income inequality (Panel B). The two are strongly positively correlated ($r = 0.58$).

Figure 1 about here.

Figure 2 plots deviations in the Gini index from its CZ and year averages against ten-year-lagged deviations in the number and share of restrictive words from their respective CZ and year averages (deviations are implemented to filter out persistent heterogeneity between CZ and time trends). There is a positive correlation between the Gini index in CZs and their exposure to regulation a decade before.

Figure 2 about here.

3.2 Baseline regression estimates

We estimate the relationships of our interest with the following regression model:

$$Y_{c,t} = \alpha Y_{c,t-10} + \beta R_{c,t-10} + \gamma_c + \gamma_t + u_{c,t} \tag{3}$$

where $Y_{c,t}$ is the focal labor market outcome (Gini index, log income, or unemployment rate) in CZ c and year t, and $R_{c,t-10}$ is the focal measure of exposure to regulation (the number RW or share RS of restrictive words, or both) a decade ago, and $u_{c,t}$ is the idiosyncratic error term which we cluster by CZ. We add time (γ_t) and CZ (γ_c) fixed effects to isolate unobserved spatial and temporal heterogeneities in our data, such as universally applicable regulations, and include the decade-lagged dependent variable Y_{ct-10} to account for persistency in labor market outcomes over time. The key regressor R_{ct-10} is decade-lagged to account for the time it takes for regulation to take effect, and to reflect the frequency with which the labor market outcomes we work with are measured from the Census data. As we noted earlier, owing to the presence of time fixed effects, the focal regression coefficients β measures differential effect of regulation between CZs differently exposed to a given amount of regulation in a given year, rather than the average effect of regulation on all CZs.

Table 2 shows effects of local exposure to regulation on log income for different percentiles of within CZ-year income distribution estimated from equation (3). Higher exposure to regulation is followed by lower income at the lower end of income distribution a decade later. The upper end of income distribution is affected more mildly or even positively in some specifications. With lower income earners losing more than higher earners, one should expect a positive effect of exposure to regulation on income inequality.

Table 2 about here.

Indeed, Table 3 demonstrates that CZs exposed to more regulation end up facing higher levels of income inequality a decade later (columns 1 and 2), which is coupled with lower average income (columns 3 and 4) and higher unemployment rate (columns 5 and 6). As higher income inequality following more exposure to regulation is not accompanied by either higher average income or employment, there seems to be no margin along which more regulation would make the affected CZs economically better off relative to less affected ones. While these results cannot say whether regulation may bring other, non-monetary benefits, or how costly (or beneficial) regulation may be on average, they do imply that localities more exposed to a given amount of regulation suffer more.

Table 3 also shows that our two measures of exposure to regulation are statistically significant on their own right, suggesting that both absolute (volume) and relative (share) strictness of regulation matter. However, because the share of restrictive words varies much less than the number, the implied effects of its one standard deviation change are much smaller than those associated with a standard deviation change in the number of restrictive words. For instance, a one standard deviation increase in the locally applicable number of restrictive words is linked to an increase in the Gini index by 0.03, whereas a similarly large increase in the share of restrictive words increases the Gini index by a quarter of that amount. For this reason, we will focus on local exposure to regulation measured with the number of restrictive words (RW) in what follows.

Table 3 about here.

Effects by CFR title

As we stated in the introduction, much of the existing literature focuses on environmental, banking, and labor regulation. To better connect with this literature, we now separate the overall regulation into sections coming from CFR titles focusing on protecting the environment (Title 40), banking and finance (Titles 12 and 13), labor relations (Titles 20 and 29) and the rest, and rerun our analysis with the including the numbers of restrictive words falling into each rubric. Table 4 reports the results. We observe from the means of the title-specific restrictive words, calculated with (1) (in tens of thousands), that about three-quarters of the locally relevant regulation comes from CFR titles not associated with environment, finance or labor, the topics usually studied in the economics of regulation literature. Yet, due to the volume of this unclassified regulation, its effects account for most of the previously observed effects of overall regulation on local labor market outcomes. Another interesting result illustrated by Table 4 is that not all regulation has the same effects on all the outcomes of our interest: while regulation from different CFR titles has broadly similar (positive) effects on inequality, labor- and finance-focused regulation has more benign effects on income and employment than the same quantities of regulation that focus on the environment or other issues. All in all, however, the deleterious local labor market effects of unclassified regulation, which represents the bulk of locally relevant regulation, dominate the effects of more narrowly focused regulation.

Table 4 about here.

3.3 The medium- vs. long-run consequences of regulation

Our baseline regression specification (3) links current local labor market outcomes to decade-lagged exposure to regulation, producing what could be interpreted as medium-run effect estimates. To capture possible longer-run effects of regulation, we augment (3) with exposure to regulation two decades ago, $R_{c,t-20}$. The regression estimates reported in Table 5 show that two decade-lagged exposure to regulation is neither economically nor statistically significant for the current Gini index and log average income, suggesting that the effects of local exposure to regulation on these outcomes, while long-lasting, are not permanent. The estimated effects on unemployment rate, however, seem to last longer, and flip signs, largely cancelling each other.

Table 5 about here.

Leaving deeper interpretations of these varying dynamic patterns for further research, we note the more immediate implication of the results in Table 5: exposure to regulation may have different local labor market effects depending on when the regulation was introduced. As we show next (Table 6), while new regulations and changes in the existing ones contribute roughly equally to the observed variation in the exposure to regulation over time, new regulations have larger consequences than changes in the existing ones.[6]

3.4 The components of the growth in local exposure to regulation, and their importance

The growth in regulatory burden over time is a well-known fact (TheEconomist, 2022), and is also reflected in our measure of local exposure to regulation (1). Our measurement approach allows to identify five sources of its growth: (i) changes in industry employment shares over time, (ii) changes in industry relevance weights of CFR parts over time, (iii) introduction of new CFR parts, (iv) revocation of existing CFR parts, and (v) variation in restrictive word count in continuing CFR parts over time. To quantify these sources and appreciate their importance in affecting local labor market outcomes, we decompose the observed change in exposure to regulation within CZs over the five decades in our data. Separating the applicable regulation into sets existing in both years (denoted E), introduced between $t-10$

[6]One could take advantage of annual observations in RegData to estimate richer specifications measuring higher-frequency dynamics in the effects of exposure to regulation. However, yearly exposure to regulation measures are highly correlated (even ten year lags are correlated with 0.97), leading to near multicollinearity between different lags of regulation and spurious estimates.

and t (N) and repealed during the same period (X), and rearranging terms, we obtain:

$$
\begin{aligned}
RW_{ct} - RW_{ct-10} \;&=\; \sum_i s_{cit}\left(\sum_{k\in E} a_{ikt}W_{kt} + \sum_{k\in N} a_{ikt}W_{kt}\right) - \\
&-\; \sum_i s_{cit-10}\left(\sum_{k\in E} a_{ikt-10}W_{kt-1} + \sum_{k\in X} a_{ikt-10}W_{kt-10}\right) = \\
&=\; \underbrace{\sum_i s_{cit}\sum_{k\in E}\left(a_{ikt} - a_{ikt-10}\right)W_{kt}}_{R^A_{ct}\text{ - change due to industry relevance}} + \underbrace{\sum_i s_{cit-10}\sum_{k\in E} a_{ikt-10}\left(W_{kt} - W_{kt-10}\right)}_{R^W_{ct}\text{ - change due to restrictive word count}} + \\
&+\; \underbrace{\sum_i \left(s_{cit} - s_{cit-10}\right)\sum_{k\in E} a_{ikt-10}W_{kt}}_{R^S_{ct}\text{ - change due to industry employment shares}} + \\
&+\; \underbrace{\sum_i s_{cit}\sum_{k\in N} a_{ikt}W_{kt}}_{R^N_{ct}\text{ - change due to new regulations}} - \underbrace{\sum_i s_{cit-10}\sum_{k\in X} a_{ikt-10}W_{kt-10}}_{R^X_{ct}\text{ - change due to repealed regulations}} \qquad (4)
\end{aligned}
$$

The first column of Table 6 lists averages and standard deviations of the decadal changes in locally applicable regulation during our study period, overall (panel A) as well as by component (panel B) calculated from the decomposition in equation (4). The number of restrictive words in locally applicable regulations increased by just under 5000 ($= 0.493\times10^4$) in an average CZ-decade, half of which growth is accounted for by new regulations (R^N_{ct}). The other half is shared between changes in the restrictive word count, industry relevance weights or industry employment shares applicable to continuing regulation. No title, part or chapter of the CFR were ever repealed, so the contribution of repealed regulation is zero and the corresponding term in decomposition (4), $R^X_{ct} = 0$, is not included in further analysis.

Table 6 about here.

Columns (1), (2) and (3) of Table 6 report the estimated effects of the total change in local exposure to regulation (Panel A) and its components, estimated from (4), on the local labor market outcomes (panel B). The effect estimates come from a modified version of our main regression specification (3) in which we use decade-lagged decadal changes in exposure to regulation coming from the sources of variation identified in equation (4):

$$Y_{c,t} = \alpha Y_{c,t-10} + \beta_1 R^A_{c,t-1} + \beta_2 R^W_{c,t-1} + \beta_3 R^S_{c,t-1} + \beta_4 R^N_{c,t-1} + \gamma_c + \gamma_t + u_{ct} \tag{5}$$

The estimated effects of the total change in local exposure to regulation are not far from the baseline estimates in Table 3 coming from equation (3) in levels. Turning to the effects by source of the change in regulation, newly introduced regulations make the strongest individual contribution to the estimated total effects of change in local exposure to regulation on Gini index and log income. The effects of the other sources of change in local regulation are smaller and less precisely estimated. In case of unemployment, changes in both new and existing regulations are similarly weighty (coefficients 0.104 and 0.163, respectively), but given the larger variation in new regulations across CZ-years than in changes in existing regulations (st.dev. 0.1 vs. 0.03, Table 6), differences in new regulations are still the most important factor explaining the variation in unemployment rates across local labor markets.

The importance of newly introduced rules relative to other sources of growth in the exposure to regulation (Table 6) and the fading away of the effects of earlier regulations (Table 5) suggest that, although labor market effects of exposure to regulation are long-lasting, there is some adjustment over time. These results are consistent with firms' strategic adaptation to regulation through outsourcing compliance functions or building internal capacity, whichever arrangement minimizes the transaction costs involved. Changes in the existing regulations are more likely to be adequately handled within the existing arrangements, with a marginal increase in transaction costs, whereas new regulations may require new capabilities that are relatively expensive to acquire. Consequently, compared to an equivalent change in the existing regulations, new regulations result in higher transaction costs, lower business profitability, and lower demand for labor. These effects are especially strong for lower-wage workers for whom the compliance costs relative to wages are higher, and in the short run, before new regulations can be addressed more efficiently.

3.5 The implied magnitude of the effects of regulation

Performing simple calculations with the baseline regression results in Table 3, we see that a standard-deviation increase in our measure of local exposure to regulation (0.66×10^4 words, Table 1) is linked to an increase in the Gini index by 0.03 (= 0.66×0.051, or just over 1 standard deviation of the Gini index in our data), a decrease in log average income by 0.04 (5% of its standard deviation), and an increase in unemployment rate by 0.01 (a third of its standard deviation) a decade later. These implied effects are massive, but how should one think of variation in measured local exposure to regulation in terms of the actually introduced rules and their restrictiveness?

As an illustration, first, consider the following hypothetical but realistic scenario. The existing regulation increases by $\Delta W = 100,000$ restrictive words through some combination of factors explicated in the decomposition equation (4). While this is a large increase, implying an extra ten million words in CFR given their average share in the text (1%, Table 1; for comparison, there are about 800 thousand words in The King James Bible), it is close to the average decadal increase in the number of restrictive words in our sample. This new regulation is equally relevant to half of the industries with the relevance weight 0.015 for each, and has zero relevance to the other half, so that the average industry relevance of this regulation (0.0075) is close to the sample mean (Table 1). There are two CZs, CZ1 and CZ2, with 80% and 20% of their respective workforces employed in the affected industries.

Given the above scenario, the newly introduced regulation will increase our measure of local exposure to regulation (equation (1)) by $100,000 \times 0.015 \times 0.8 = 1200$ relevant restrictive words in CZ1, and by 300 relevant restrictive words in CZ2, reflecting their different industry employment structures. Based on the regression results in Table 3, the implied effects on the Gini index, log average income and unemployment rate in the more affected CZ1 are $0.051 \times 1200/10,000 = 6.1 \times 10^{-3}$ (or about 25% of standard deviation), $-0.057 \times 1200/10,000 = 6.8 \times 10^{-3}$ (0.9% of s.d.) and $0.02 \times 1200/10,000 = 2.4 \times 10^{-3}$ (8.3% of s.d.), respectively. For the less affected CZ2, the effects of the same increase in Federal regulation will be four times lower, owing to its lesser exposure to this new regulation.

As another illustration, consider a newly introduced regulation counting 50,000 restrictive words. This scenario is also realistic given that about half of the change in local exposure to regulation is due to newly introduced rules. Holding the industry relevance and CZ employment structure the same as in the previous example, CZ1's exposure to regulation will increase by 600 relevant restrictive words, and CZ2's by 150. Based on the regression estimates by component of the decadal growth in regulation (Table 6), this new regulation would increase the Gini index in the more affected CZ1 by $0.086 \times 600/10000 = 5.2 \times 10^{-3}$ (21% of s.d.), decrease log average income by $-0.298 \times 600/10000 = 0.018$ (2.3% of s.d.), and increase unemployment by $0.104 \times 600/10000 = 6.24 \times 10^{-3}$ (21.5% of s.d.). Our calculations thus suggest that the implied local labor market effects of exposure to regulation are substantial and driven primarily by newly introduced red tape.

4 Endogeneity concerns

An important challenge to interpreting our regression results in terms of the effects of regulation is potential endogeneity of local exposure to regulation to other local factors that may also affect the outcomes of our interest. We here probe the robustness of our results to two sources of endogeneity: local employment structure, and political connections.

Local employment structure

We begin by focusing on what we believe is the most plausible source of endogeneity: CZ-level industry employment structure. Indeed, this is the only local source of variation in exposure to regulation, since the other sources – industry relevance, restrictiveness and volume of regulation – apply equally to all CZs. To appreciate potential endogeneity of local industry employment structure, it is helpful to draw analogy with the "China shock" literature (Autor and Dorn, 2013; Autor et al., 2013, 2016). Industries more exposed to imports from China, or, in our story, more regulation, lose more output and employment. As these industries shrink, their shares in local employment decrease, leading to underestimated local exposure to regulation as well as its variation (recall equation (1)), and resulting in upward-biased effect estimates. While using lags of exposure to regulation in empirical analysis, as we do, should go some way towards addressing the reverse causality outlined above, there may be other sources of endogeneity, for instance, industries shrinking in expectation of more regulation in the future.

To address the above endogeneity concerns, following the China shock literature, we make use of shift-share instruments (Borusyak et al., 2025). Specifically, we instrument our measure of local exposure to regulation defined in (1) as a function of contemporary employment shares, $RW_{ct}(s_{cit}, a_{ikt}, W_{kt})$, with the same expression but evaluated at decade-lagged employment shares, $RW_{ct}(s_{ci,t-10}, a_{ikt}, W_{kt})$ in one specification, and at employment shares as of 1970, $RW_{ct}(s_{ci,1970}, a_{ikt}, W_{kt})$, in another specification. We apply the instrumental variable estimator to both specifications. Table 7 reports the results. As evidenced by the large first-stage F-statistics, the instruments are strong. The instrumental variable regression estimates are broadly similar to our baseline results, copied for ease of comparison.

Table 7 about here.

Local political connections

Another challenge to interpreting our regression results as causal effect estimates is potential feedback from locally important interest groups transmitted to federal regulators through locally connected politicians. For instance, politicians with ties to CZs dominated by certain industries or to CZs making up large shares in certain industries nationwide may wish to sponsor regulation favorable to those industries. An example of such activity is offered in Lake and Millimet (2016) who find U.S. Congress representatives to be more likely to vote for a nationwide free-trade agreement if the expected redistribution of the implied gains from trade is more favorable to the residents of their congressional districts.

To assess the sensitivity of our results to the possible endogeneity of federal regulation to local interests, we rerun our analysis excluding CZs where such interests may be relatively powerful. We use two exclusion criteria, in separate specifications. First, we exclude CZs that had particularly high shares in nationwide employment for at least one industry at least once. Second, we exclude CZs that had particularly high local employment concentration by industry at least once. We used the 99th and 95th percentiles of the nationwide distribution of the focal variable as the cutoff points for both exclusion criteria.

Table 8 about here.

The results, presented in Table 8, are broadly similar to our baseline findings, showing no strong evidence that our conclusions so far are driven by regulation being endogenous to local interests. To provide some background to these results, while CZs surely differ in industry employment structure, even those in the top of the distribution are neither particularly highly concentrated nor represent large shares in nationwide employment for any single industry. For instance, CZs in the top 5% of nationwide distribution of the Herfindahl-Hirschman (HHI) industry employment concentration index (our measure of CZ-level industry employment concentration) have an HHI 0.052 or above, which corresponds to the "effective number of industries" equal to $1/0.052 = 19$ or fewer, which is not particularly low given the total of 80 industries in the classification we use. Similarly, the top 5th percentile of the distribution of single CZ shares in nationwide employment in any industry is 0.6%, suggesting relatively low dominance of specific CZs in nationwide employment. Thus, given the low industry employment concentration within and across CZs that we observe in our data, local interests at CZ level are unlikely to have strongly affected nationwide regulation.

5 Validating our measurement approach

Having explored the extent to which our regression results may be causally interpreted, we now turn to probing the validity of our approach to measure exposure to regulation. Our measure exploits the variation in the relevance of regulation across industries and the variation in industry employment shares across CZs. To gauge the importance of these two sources of variation for identifying the effects of regulation on local labor market outcomes, we perform placebo treatment tests (Eggers et al., 2023). In our specific implementation, we rerun our analysis on simulated data produced by replacing the actually observed industry relevance and shares in local employment with random draws from their distributions. We repeat this procedure multiple times, noting the frequency with which our baseline results are reproduced. Low reproduction frequency would indicate the importance of the industry relevance and employment shares in identifying the effects of our interest, validating our measurement approach.

Industry relevance

We begin with the industry relevance weight a_{ikt} in equation (1) as computed by Al-Ubaydli and McLaughlin (2017). How would our baseline regression results change if we replaced the actual a_{ikt} with a randomly drawn placebo value? Given the very large number of regulatory parts in our data, for computational simplicity, we redefine industry relevance with a categorized relevance weight (CRW) dummy equal 1 if the actual a_{ikt} exceeds a certain threshold, and 0 otherwise, and calculate the exposure to regulation measure with the so defined CRW. Columns 1, 3 and 5 in panel A of Table 9 report the results from the baseline regression specification (3) based on the CRWs defined for three different thresholds in a_{ikt} (0.01, 0.05, 0.09) for the Gini index, log income and unemployment rate, respectively. The results are qualitatively similar to our baseline estimates in Table 3. Their being smaller in magnitude reflects both larger mean of and measurement error in the CRW-based measure of exposure to regulation as compared to the original one in equation (1).

Table 9 about here.

We then produce simulated CRWs by replacing the actual CRWs with a random 0/1 draw from the same distribution, calculate the implied exposure to regulation with (1), and apply regression (3) to the simulated data, repeating the procedure 100 times.[7] Columns 2, 4 and 6 in panel A of Table 9 report the means of the regression coefficients estimated on the simulated CRWs, the number of times those coefficients exceeded the ones estimated with the actual CRWs, and the number of times the estimates from the simulated data were statistically significant at 1% level. Figure 3 plots kernel densities of the effects of exposure to regulation estimated on the simulated CRWs against the effect estimated on the actual data (thick vertical line). The regressions with simulated CRWs produce estimates that center around zero. Although the estimates based on simulated CRWs are statistically significant 10 to 24% of the time, depending on the specification and owing presumably to the variation in industry employment shares, they hardly ever match or exceed in magnitude those obtained from the same specifications applied to the actual data. The differences between the regression results on the real vs. simulated data are especially strong for the Gini index and unemployment rate, but even for log income, the estimates with simulated CRWs are substantially smaller than those obtained on real data. The failure of the same regression specifications applied to simulated CRW data to reproduce the results based on the actual CRWs suggests the importance of industry relevance as a channel through which regulation affects local economies. It also validates the industry relevance measure developed by Al-Ubaydli and McLaughlin (2017) and our approach that relies on this measure.

Figure 3 about here.

Industry employment structure

[7] The number of repetitions, 100, may seem relatively low as compared to other studies relying on stochastic simulation, but our simulations required significant computational effort. It took us three weeks to perform the simulation exercise for industry relevance weights with 100 repetitions.

In doing placebo treatment tests for industry shares in local employment, we follow the same procedure as described above, except that we replace the actual employment shares (s_{cit} in (1)), rather than category dummies, with random draws from their observed distribution. Columns 2, 4 and 6 in panel B of Table 9 report the means of the regression coefficients estimated from equation (3) applied to data with simulated industry employment shares, the number of times those coefficients exceeded the ones estimated with the actual data, and the number of times the estimates from the simulated data were statistically significant at 1% level. For comparison, columns 1, 3 and 5 copy the corresponding estimates on the actual data, first reported in Table 3. Figure 4 plots kernel densities of the effects of exposure to regulation on unemployment rate, log income and the Gini index estimated with the simulated industry employment shares. Similar to industry relevance, the effect estimates with placebo employment shares center around zero, hardly ever exceed the estimates obtained on the actual data, and are never statistically significant. The failure to reproduce our findings on simulated data validates our measurement approach by showing that the variation in industry relevance of regulation combined with the variation in industry employment mix in equation (1) is reliably correlated with the variation in inequality, unemployment and income across local labor markets.

Figure 4 about here.

Results at the state vs. CZ level

In a separate but related inquiry, we ask how local the exposure to regulation should be to matter. Table 10 reports the results of rerunning the same regression specification (3) we have used on the CZ-year data before with the outcomes and exposure to regulation now aggregated to the state-year. We observe much weaker relationships (columns 1, 4, 7). This may have been due to lack of statistical power or the ten-year time lag of the exposure to regulation variable being too long to detect a relationship. However, rerunning the same analysis with shorter, one or two-year lags of exposure to regulation, we obtain qualitatively similar results for the Gini index and, puzzlingly, a positive relationship between recent exposure to regulation and log income.

Table 10 about here.

One reason for failing to replicate at the state level the relationships we observed at the CZ level may be that states differ less than CZs in industry employment structure. For an illustration, consider Herfindahl-Hirschman industry employment concentration index (HHI) we have used earlier in this study. In year 2000, the variance in HHI across CZs was 2.4 times larger than across states, suggesting that aggregating data up to the state level would result in a large loss of variation in industry employment structure. To the extent that exposure to regulated industries through employment is a mechanism that enables local labor market effects of regulation, suppressing the variation in industry employment structure would prevent detecting those effects and render the measure of exposure to regulation through employment less relevant. Yet, regulation does seem to have local consequences, ignoring which could lead to its costs being under-appreciated. Measuring local consequences of regulation requires a suitably granular measurement method such as the one we propose in this study.

6 Conclusion

Overall, our results suggest that the costs of federal regulation are not spread evenly but are higher in locations more exposed to it through local employment structure. Identification of varying local effects of federal regulation is a nontrivial task, and there is little empirical work on this topic to date. We develop an approach to measuring local effects of regulation using variation in exposure to regulated industries through local employment. Specifically, our measure of local exposure to regulation incorporates the restrictiveness and industry relevance of regulation, proxied with the number of restrictive words and frequency of industry-specific terms in the regulation text, respectively, and industry weights in local employment.

We find that our measure of local (CZ-level) exposure to regulation is robustly and positively correlated with the local Gini coefficient of income inequality and unemployment rate, and negatively with log average income a decade after. Further analyses reveal that these correlations are driven primarily by newly passed, general-purpose regulation rather than changes in the existing or domain-specific (environment, financial or labor) red tape. A causal interpretation of our regression results is warranted by their robustness to potential endogeneity of industry shares in local employment. Failure to reproduce significant effects in placebo tests validates our measurement approach and suggests that local labor markets are more, or less, affected by federal regulation depending on their exposure to regulated industries through local employment structure.

Our approach is not without limitations: it cannot fully capture the strictness of a text with a given number of restrictive words or detect the effects of regulation applicable to all industries. These noted limitations notwithstanding, we offer an innovative and widely applicable method of quantifying locally differential effects of federal regulation that can facilitate research in various directions beyond our present study. For instance, our method can be applied to measuring exposure to regulation coming from different sources (e.g., certain parts in CFR or other compendia of regulatory texts) and at different levels of aggregation, including individual firms. It is also flexible enough to be redesigned to study the effects of other aspects regulation as may be captured by the text search terms other than the standard restrictive words we used.

Figure 1: Share of restrictive words in applicable regulation (Panel A) and income inequality (Panel B) by commuting zone in year 2000

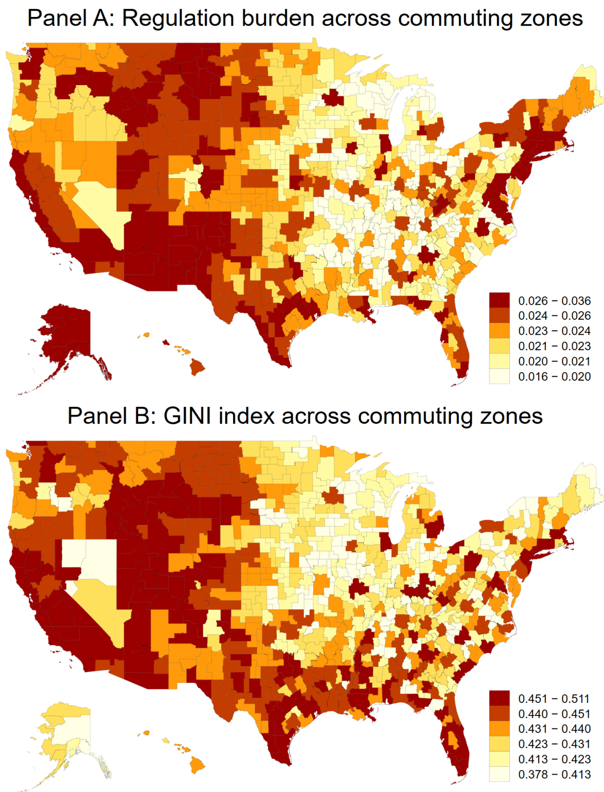


Figure 2: Gini index of income inequality and regulatory burden a decade ago

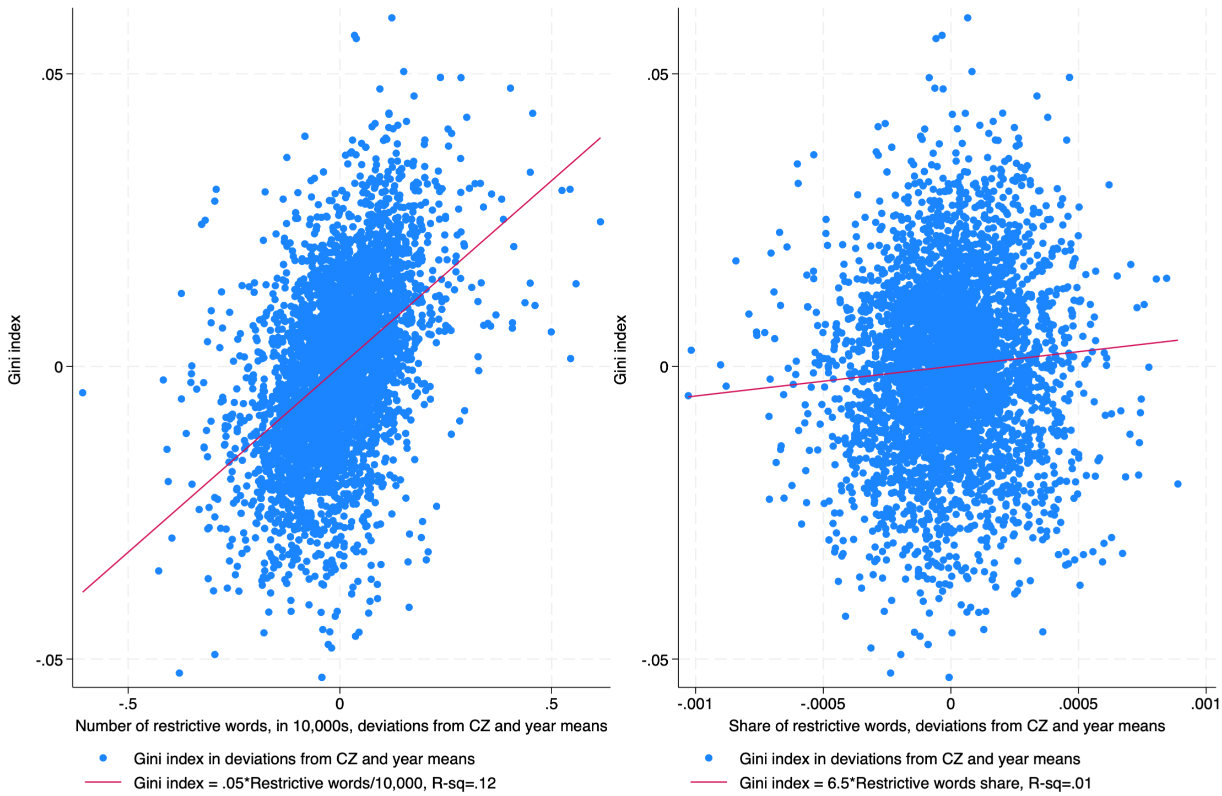


Figure 3: Kernel density of the effects of regulation estimated on data simulated by replacing the true industry relevance with random draws from its distribution (thick vertical lines - effect estimates on actual data)

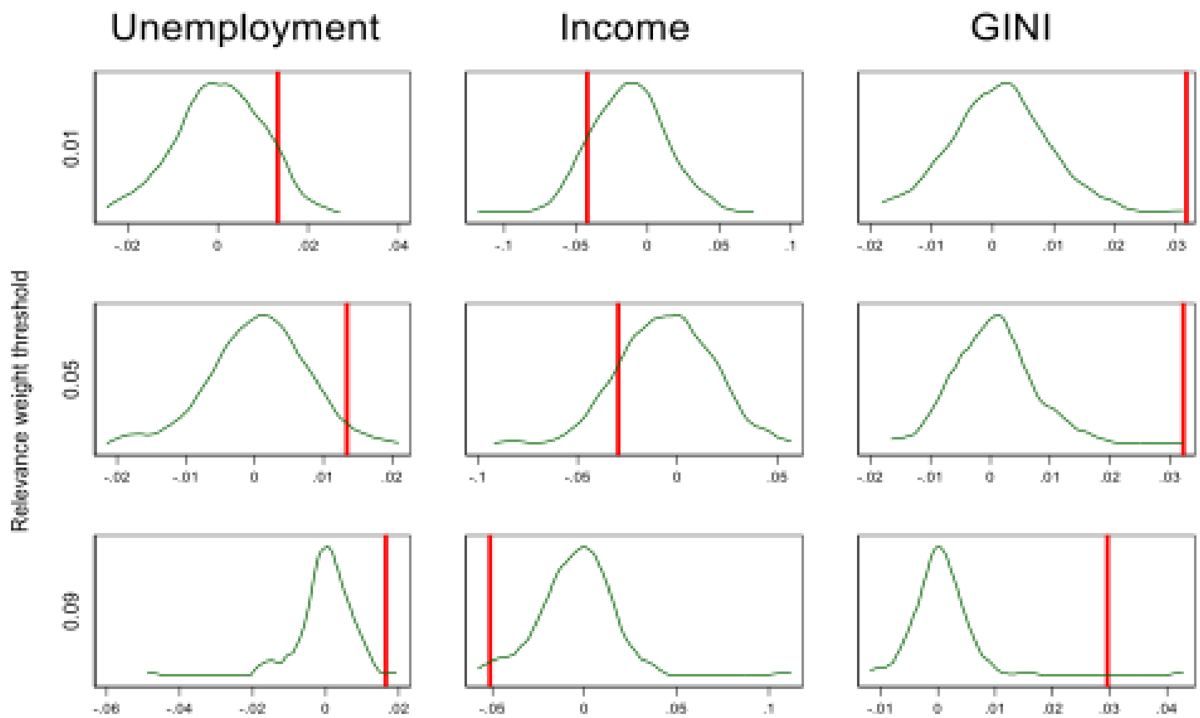
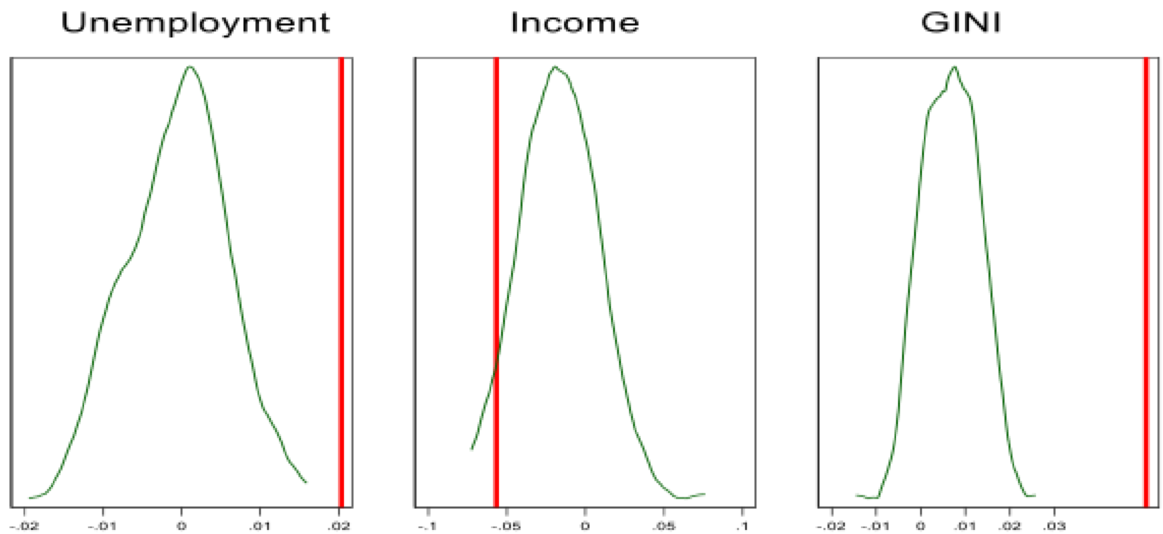


Figure 4: Kernel density of the effects of regulation estimated on data simulated by replacing the true CZ-level employment shares with random draws (thick vertical lines-effect estimates on actual data)



References

Agnello, L., Mallick, S. K., and Sousa, R. M. (2012). Financial reforms and income inequality. *Economics Letters*, 116(3):583–587.

Al-Ubaydli, O. and 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(1):109–123.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–97.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The China Syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, 103(6):2121–68.

Autor, D. H., Dorn, D., and Hanson, G. H. (2016). The China Shock: Learning from labor-market adjustment to large changes in trade. *Annual Review of Economics*, 8(1):205–240.

Bailey, J. B. and Thomas, D. W. (2017). Regulating away competition: The effect of regulation on entrepreneurship and employment. *Journal of Regulatory Economics*, 52(3):237–254.

Bailey, J. B., Thomas, D. W., and Anderson, J. R. (2019). Regressive effects of regulation on wages. *Public Choice*, 180(1):91–103.

Baldwin, R., Cave, M., and Lodge, M. (2011). Why regulate? In *Understanding Regulation: Theory, Strategy, and Practice*. Oxford University Press.

Beck, T., Levine, R., and Levkov, A. (2010). Big bad banks? The winners and losers from bank deregulation in the United States. *The Journal of Finance*, 65(5):1637–1667.

Besley, T. and Burgess, R. (2004). Can labor regulation hinder economic performance? Evidence from India. *The Quarterly Journal of Economics*, 119(1):91–134.

Black, D. A., Kolesnikova, N., and Taylor, L. J. (2014). Local labor markets and the evolution of inequality. *Annual Review of Economics*, 6(1):605–628.

Bombardini, M., Trebbi, F., and Zhang, M. B. (2024). Measuring the costs and benefits of regulation. Working Paper 32955, National Bureau of Economic Research.

Borusyak, K., Hull, P., and Jaravel, X. (2025). A practical guide to shift-share instruments. *Journal of Economic Perspectives*, 39(1):181–204.

Botero, J. C., Djankov, S., Porta, R. L., Lopez-de Silanes, F., and Shleifer, A. (2004). The regulation of labor. *The Quarterly Journal of Economics*, 119(4):1339–1382.

Calderón, C. and Chong, A. (2009). Labor market institutions and income inequality: An empirical exploration. *Public Choice*, 138(1):65–81.

Chambers, D., McLaughlin, P. A., and Stanley, L. (2019). Regulation and poverty: An empirical examination of the relationship between the incidence of federal regulation and the occurrence of poverty across the US states. *Public Choice*, 180(1):131–144.

Chambers, D. and O'Reilly, C. (2022a). The economic theory of regulation and inequality. *Public Choice*, 193(1):63–78.

Chambers, D. and O'Reilly, C. (2022b). Regulation and income inequality in the United States. *European Journal of Political Economy*, 72:102101.

Choudhury, S. (2023). The causal effect of regulation on income inequality across the U.S. states. *European Journal of Political Economy*, 80:102471.

Coffey, B., McLaughlin, P. A., and Peretto, P. (2020). The cumulative cost of regulations. *Review* *of Economic Dynamics*, 38:1–21.

Dawson, J. W. and Seater, J. J. (2013). Federal regulation and aggregate economic growth. *Journal of Economic Growth*, 18(2):137–177.

de Haan, J. and Sturm, J.-E. (2017). Finance and income inequality: A review and new evidence. *European Journal of Political Economy*, 50:171–195.

Delis, M. D., Hasan, I., and Kazakis, P. (2013). Bank regulations and income inequality: Empirical evidence. *Review of Finance*, 18(5):1811–1846.

Djankov, S., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2002). The regulation of entry. *The Quarterly Journal of Economics*, 117(1):1–37.

Eggers, A. C., Tuñón, G., and Dafoe, A. (2023). Placebo tests for causal inference. *American Journal of Political Science*, 68(3):1106–1121.

Feldmann, H. (2009). The unemployment effects of labor regulation around the world. *Journal of Comparative Economics*, 37(1):76–90.

Gimpelson, V. and Treisman, D. (2018). Misperceiving inequality. *Economics & Politics*, 30(1):27–54.

Huang, B. and Yao, Y. (2023). Does environmental regulation matter for income inequality? New evidence from Chinese communities. *Journal of the Association of Environmental and Resource Economists*, 10(5):1309–1334.

Jha, A., Matthews, P. H., and Muller, N. Z. (2019). Does environmental policy affect income inequality? Evidence from the Clean Air Act. *AEA Papers and Proceedings*, 109:271–76.

Knell, M. and Stix, H. (2020). Perceptions of inequality. *European Journal of Political Economy*, 65:101927.

Koeniger, W., Leonardi, M., and Nunziata, L. (2007). Labor market institutions and wage inequality. *Industrial and Labor Relations Review*, 60(3):340–356.

Kuhn, A. (2019). The subversive nature of inequality: Subjective inequality perceptions and attitudes to social inequality. *European Journal of Political Economy*, 59:331–344.

Lake, J. and Millimet, D. L. (2016). An empirical analysis of trade-related redistribution and the political viability of free trade. *Journal of International Economics*, 99:156–178.

McLaughlin, P. A. and Sherouse, O. (2019). RegData 2.2: A panel dataset on US federal regulations. *Public Choice*, 180(1):43–55.

Moretti, E. (2011). Local labor markets. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4, pages 1237–1313. Elsevier.

Mulholland, S. E. (2019). Stratification by regulation: Are bootleggers and Baptists biased? *Public Choice*, 180(1):105–130.

Newman, B. J., Johnston, C. D., and Lown, P. L. (2015). False consciousness or class awareness? Local income inequality, personal economic position, and belief in American meritocracy. *American Journal of Political Science*, 59(2):326–340.

Newman, B. J., Shah, S., and Lauterbach, E. (2018). Who sees an hourglass? Assessing citizens' perception of local economic inequality. *Research & Politics*, 5(3).

Saez, E. and Zucman, G. (2020). The rise of income and wealth inequality in America: Evidence from distributional macroeconomic accounts. *Journal of Economic Perspectives*, 34(4):3–26.

TheEconomist (2022). *Enthusiasm for regulation, often in areas like the climate, shows no sign of* *flagging*, 10 january edition.

Tolbert, C. M. and Sizer, M. (1996). U.S. commuting zones and labor market areas: A 1990 update. Staff Reports 278812, United States Department of Agriculture, Economic Research Service.

Tolbert, Charles M., I. and Killian, M. S. (1987). Labor market areas for the United States. Staff Reports 277959, United States Department of Agriculture, Economic Research Service.

Table 1. Summary Statistics.

	(1) Mean	(2) 25th pctile	(3) Median	(4) 75th pctile	(5) Std	(6) N
			Commuting zone level			
Gini index	0.428	0.413	0.427	0.441	0.025	4,446
Unemployment	0.062	0.040	0.055	0.076	0.029	4,446
Log Income	10.10	9.52	10.26	10.70	0.783	4,446
Regulatory burden based on the applicable number of restrictive words	12,672	7,128	12,346	17,841	6,616	4,446
Regulatory burden as % of applicable restrictive words	1.104	1.074	1.110	1.144	0.062	4,446
			Regulatory item level			
Number of restrictive words	117.2	9.0	32.0	95.0	667.2	36,936
Share of restrictive words	0.011	0.006	0.011	0.015	0.007	36,936
Cumulative industry relevance weight	0.697	0.176	0.472	1.059	0.661	36,936
			Regulatory item-industry level			
Individual industry relevance weight	0.0078	0.0002	0.0004	0.0008	0.0662	3,287,304

Notes: Descriptive statistics are reported for years 1980, 1990, 2000, 2010, and 2019, which are included in our regression sample.

Table 2. Effects of exposure to regulation on percentiles of income distribution.

	(1)	(2)	(3)	(4)	(5)
	Percentile of log income distribution				
	10	25	50	75	90
Lag Restrictions (word count)	-0.432*** (-9.52)	-0.258*** (-11.18)	-0.121*** (-6.30)	-0.034* (-1.92)	-0.024 (-1.33)
Lag Restrictions (percent)	-2.699*** (-9.86)	-1.002*** (-7.72)	-0.138 (-1.36)	0.173 (1.57)	0.234*** (2.65)

Sample: Commuting zones at decadal intervals from 1970 to 2020.
Notes: Standard errors are clustered at the commuting zone level; *t*-statistics are reported in parenthesis. All specifications include time and commuting zone fixed effects.

Table 3. Baseline results at commuting zone level.

	(1)	(2)	(3)	(4)	(5)	(6)
	Gini index		Log Income		Unemployment	
Lag Restrictions (word count)	0.051*** (12.94)	0.044*** (11.24)	-0.057*** (-4.07)	-0.049*** (-3.52)	0.020*** (6.49)	0.021*** (6.19)
Lag Restrictions (percent)		0.121*** (7.54)		-0.186** (-2.31)		-0.009 (-0.42)
Lag Dep. Var.	0.167*** (6.76)	0.167*** (6.82)	0.420*** (13.68)	0.434*** (14.32)	-0.098*** (-4.04)	-0.098*** (-4.04)
R2	0.726	0.732	0.991	0.992	0.777	0.777
N	3,705	3,705	3,705	3,705	3,705	3,705

Sample: Commuting zones at decadal intervals from 1970 to 2020.
Notes: Standard errors are clustered at the commuting zone level; *t*-statistics are reported in parenthesis. All specifications include time and commuting zone fixed effects.

Table 4. Result by area of regulation: Environment, Labor, Finance, and all other.

	Mean (St.Dev.)	(1) Gini index	(2) Log Income	(3) Unemployment
Lag Restrictions (Environment)	0.094 (0.088)	0.020* (1.76)	-0.317*** (-6.92)	0.011 (1.07)
Lag Restrictions (Finance)	0.028 (0.015)	0.081 (1.08)	1.649*** (3.95)	-0.066 (-0.84)
Lag Restrictions (Labor)	0.128 (0.082)	0.087*** (3.74)	0.191** (2.04)	-0.185*** (-6.26)
Lag Restrictions (Other)	0.823 (0.371)	0.048*** (8.40)	-0.097*** (-3.76)	0.061*** (9.33)
R2		0.727	0.992	0.782
N		3,705	3,705	3,705

Notes: Standard errors are clustered at the commuting zone level; *t*-statistics are reported in parenthesis. All specifications include time and commuting zone fixed effects. The total number of regulations from the baseline regression is disaggregated into Environmental, Financial, Labor, and all other categories by CFR title.

Table 5. Baseline results with two lags of exposure to regulation.

	(1)	(2)	(3)	(4)	(5)	(6)
	Gini index		Log Income		Unemployment	
Restrictions (first lag)	0.051*** (12.94)	0.050*** (9.19)	-0.057*** (-4.07)	-0.126*** (-7.50)	0.020*** (6.49)	0.030*** (6.34)
Restrictions (second lag)		0.005 (0.82)		-0.020 (-0.93)		-0.020*** (-3.38)
R2	0.726	0.745	0.991	0.989	0.777	0.821
N	3,705	2,964	3,705	2,964	3,705	2,964

Notes: Standard errors are clustered at the commuting zone level; *t*-statistics are reported in parenthesis. All specifications include time and commuting zone fixed effects.

Table 6. Effects of changes in local exposure to regulation by source, as defined in the decomposition equation (4).

	(1) Gini index	(2) Log Income	(3) Unemployment
		Panel A	
Total change (mean=0.493, std=0.111)	0.035*** (6.45)	-0.078*** (-4.52)	0.025*** (5.46)
		Panel B	
Components of total change (equation 4):			
Change in industry relevance weights (mean=0.109, std=0.084)	0.002 (0.30)	-0.030 (-1.33)	-0.030*** (-4.61)
Change in regulation intensity for existing regulations (mean=0.070, std=0.030)	-0.034 (-1.18)	-0.201 (-1.44)	0.163*** (4.93)
Change in employment shares (mean=0.066, std=0.096)	0.019** (2.57)	-0.012 (-0.46)	0.018*** (2.95)
New regulations (mean=0.249, std=0.101)	0.086*** (6.99)	-0.298*** (-6.86)	0.104*** (7.18)
N	2,964	2,964	2,964

Notes: This table reports the estimates from regressions of local labor market outcomes on the total decadal change in local exposure to regulation (Panel A) and on the components of this change separately (Panel B), as identified in decomposition equation (4).
Sample: Commuting zones at decadal intervals from 1970 to 2020.
Standard errors are clustered at the commuting zone level; *t*-statistics are reported in parenthesis. All specifications include time and commuting zone fixed effects, as well as the lagged dependent variable.

Table 7. Results with and without instrumental variables for CZ industry employment shares.

	(1) Gini index	(2) Log Income	(3) Unemployment
	OLS (no instrumenting)		
Lag Restrictions (word count)	0.051***	-0.057***	0.020***
	(12.94)	(-4.07)	(6.49)
	IV with 10-year lagged employment shares		
Lag Restrictions (word count)	0.052***	-0.123***	0.019***
	(6.22)	(-4.00)	(2.90)
First stage F-stat	105.56	234.75	146.72
	IV with 1970 employment shares		
Lag Restrictions (word count)	0.032**	-0.278***	0.019
	(2.38)	(-5.45)	(1.47)
First stage F-stat	39.99	117.43	55.80

Sample: Commuting zones at decadal intervals from 1970 to 2020. Notes: Standard errors are clustered at the commuting zone level; *t*-statistics are reported in parenthesis. All specifications include time and commuting zone fixed effects.

Table 8. Results excluding commuting zones with highly concentrated industry employment structure.

	(1) Gini index	(2) Log Income	(3) Unemployment
Excluding CZs with high industry shares in countrywide employment:			
CZs with at least one industry in top 5% of countrywide employment at least once	0.039*** (6.60) 2,461	-0.068*** (-3.01) 2,461	0.013** (2.57) 2,461
CZs with at least one industry in top 1% of countrywide employment at least once	0.045*** (9.86) 3,273	-0.070*** (-4.30) 3,273	0.016*** (4.29) 3,273
Excluding CZs with high industry employment concentration:			
HHI in the top 5% countrywide at least once	0.051*** (11.31) 2,925	0.000 (0.01) 2,925	0.022*** (5.57) 2,925
HHI in the top 1% countrywide at least once	0.050*** (11.98) 3,495	-0.044*** (-3.13) 3,495	0.021*** (6.01) 3,495

Notes: Standard errors are clustered at the commuting zone level; *t*-statistics are reported in parenthesis. All specifications include time and commuting zone fixed effects.

Table 9. Results of validating our measure of exposure to regulation with placebo tests.

		(1)	(2)	(3)	(4)	(5)	(6)
		Gini index		Log Income		Unemployment	
		Real	Simulated	Real	Simulated	Real	Simulated
		Panel A: Simulated industry relevance weights					
Number of restrictive words	CRW=1 if relevance weight ≥ 0.01	0.032*** (11.847)	0.001 [0/13]	-0.042*** (-3.590)	-0.014 [14/13]	0.013*** (6.885)	0.001 [10/24]
	CRW=1 if relevance weight ≥ 0.05	0.032*** (10.242)	0.001 [0/15]	-0.030*** (-3.904)	-0.007 [16/18]	0.012*** (6.272)	0.001 [6/21]
	CRW=1 if relevance weight ≥ 0.09	0.030*** (9.486)	0.001 [1/10]	-0.051*** (-4.243)	-0.004 [2/16]	0.011*** (6.533)	0.000 [1/24]
		Panel B: Simulated industry shares in local employment					
Number of restrictive words		0.051*** (12.94)	0.007 [0/0]	-0.057*** (-4.07)	-0.016 [6/0]	0.020*** (6.49)	-0.001 [0/0]

Notes: In panel A, the specifications in rows differ by the value of the relevance weight above which a given item of regulation is considered relevant for a given industry. For instance, in the first row, a given item of regulation is considered relevant for all industries whose relevance weights for that item are above 0.01. These industries get their original relevance weight replaced with a categorized relevance weight CRW=1, while the rest of the industries get CRW=0.

Columns (1), (3) and (5) report the results based on the actual data: CRWs in panel A, industry employment shares in panel B. Columns (2), (4) and (6) report the results from the same specifications estimated on simulated rather than actual data. The first cell reports the mean of the coefficient estimates across 100 simulated samples. The first number in square brackets indicates the number of simulated samples producing coefficient estimates equal or larger than the estimates from the real data; the second is the number of simulated samples producing statistically significant estimates at 1% level.

Table 10. Baseline results at the state level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Gini index			Log Income			Unemployment	
10-year lag Restrictions (word count)	0.018 (0.68)			0.056 (1.35)			0.016* (1.92)		
1-year lag Restrictions (word count)		0.023 (1.02)			0.108** (2.36)			0.023** (2.46)	
2-year lag Restrictions (word count)			0.024 (1.05)			0.096** (2.12)			0.023** (2.50)
10-year lag Dep. Var.	0.369*** (3.03)	0.337*** (3.25)	0.333*** (3.22)	0.375*** (4.11)	0.296*** (3.12)	0.309*** (3.23)	-0.166** (-2.13)	-0.168** (-2.39)	-0.171** (-2.41)
R2	0.891	0.895	0.896	0.994	0.994	0.994	0.833	0.840	0.840
N	248	248	248	248	248	248	248	248	248

Notes: States at decadal intervals from 1970 to 2020. Standard errors are clustered at the state level; *t*-statistics are reported in parenthesis. All specifications include time and state fixed effects.