NBER WORKING PAPER SERIES

WHO BENEFITS FROM WORKER REPRESENTATION ON CORPORATE BOARDS?

Christine Blandhol
Magne Mogstad
Peter Nilsson
Ola L. Vestad

Working Paper 28269
http://www.nber.org/papers/w28269

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2020, Revised June 2025

The project received financial support from the Norwegian Research Council (grant number 275123). We thank Devan Rawlings for outstanding research assistance. We are grateful for useful comments from seminar participants at the CESifo, the University of Oslo, the University of Essex, and the Research Institute of Industrial Economics (IFN). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Who Benefits from Worker Representation on Corporate Boards?
Christine Blandhol, Magne Mogstad, Peter Nilsson, and Ola L. Vestad
NBER Working Paper No. 28269
December 2020, Revised June 2025
JEL No. G34, G38, J31, J54, J58

ABSTRACT

We study a size-contingent law in Norway that grants workers the right to board representation in firms with 30 or more employees. To analyze the impact of the law, we embed the regulation into an equilibrium model of the labor market. We show how to use behavioral responses to the regulation to identify (i) the direct effects of the policy on regulated firms and workers, (ii) the distortions from firms adjusting their size to avoid the regulation, and (iii) the equilibrium effects in the labor market. We evaluate these effects on firm profits and production, as well as on worker compensation, including both wages and non-wage amenities.

Christine Blandhol
Princeton University
Department of Economics
christine.blandhol@gmail.com

Magne Mogstad
University of Chicago
Department of Economics
and NBER
magne.mogstad@gmail.com

Peter Nilsson
Stockholm University
Department of Economics
Peter.nilsson@iies.su.se

Ola L. Vestad
Statistics Norway
ola.vestad@ssb.no

1 Introduction

Corporations are made up of multiple stakeholders, including workers, managers, and shareholders. The rights and interests of these stakeholders vary markedly across western market economies. On the one hand, both the US and the UK follow a corporate governance system of shareholder primacy. In it, it is the shareholders, and the shareholders alone, that elect the corporate board, which directly or indirectly manages the corporation on behalf of the shareholders. On the other hand, many European countries follow a shared governance system that grants workers formal authority in the corporate decision-making, often in the form of worker representation on corporate boards. As many economies have seen significant declines in the labor share of income, worker representation on corporate boards has gained popularity as a way to ensure the interests and views of the workers. For example, recent polls suggest that a majority of American voters want workers to hold seats on corporate boards,[1] and leading politicians both in the US and the UK are advocating a system of shared governance.[2] Yet, there is limited scientific evidence on how such a shared governance system actually affects firms and workers.

This paper contributes to a growing empirical literature on whether, to what extent, and why employees benefit from worker representation on corporate boards, and how such representation affects firm outcomes (e.g., Gorton and Schmid, 2004; Jäger, Schoefer and Heining, 2021). We study a size-contingent law in Norway that grants workers the rights to board representation in firms with 30 or more employees. Building on Garicano, Lelarge and Van Reenen (2016), we embed this regulation into an equilibrium model of the labor market. We show how to use the behavioral responses to the regulation to identify (i) the direct effects of the policy on regulated firms and workers, (ii) the distortions from firms adjusting their size to avoid the regulation, and (iii) the equilibrium effects in the labor market. We evaluate these effects on both worker compensation (including wages and amenities) and firm profits.

Section 2 describes the institutional context and key empirical facts about worker rep-

[1]See https://www.vox.com/2018/4/6/17086720/poll-corporate-board-democracy-worker-council-codetermination-union-labor.

[2]A prominent example from the US is Senator Elizabeth Warren, who proposed a federal bill in 2018 that would give workers in large corporations the right to elect two fifths of all board seats; https://www.vox.com/2018/8/15/17683022/elizabeth-warren-accountable-capitalism corporations. Worker representation has also been high on the political agenda in the UK, with Theresa May pledging to have workers represented on corporate boards in her campaign for the 2016 Conservative Party leadership election; https://www.theguardian.com/politics/nils-pratley-on-finance/2016/jul/11/theresa-may-plan-workers-boardroom-reform-extraordinary-tories.

resentation in Norway. We document that approximately 40% of Norwegian workers are employed in firms with at least one worker representative on the corporate board. The prevalence of representation rises sharply with firm size: it is around 3% among workers in firms with fewer than 30 employees and nearly 80% in firms with more than 200 employees. Even though board representation is increasing in firm size, we find no evidence of a clear discontinuity at the 30-employee regulation threshold.

We discuss two potential explanations for this lack of discontinuity in worker representation at the threshold, both of which inform our modeling assumptions. First, as emphasized by Jäger et al. (2021), the mere existence of representation rights may influence firm behavior, even if workers do not elect a representative. The reason is that workers can then credibly threaten to demand worker representation, affecting bargaining power and outcomes between workers and firms. Accordingly, our baseline model assumes that the policy directly affects all firms above the threshold, regardless of whether representation is implemented.

Second, the absence of a discontinuity may reflect "measurement error" in the sense that the observed number of employees in the data might not reflect the latent expected firm size that workers and firms consider when making their decisions. To account for this, we allow for measurement error in the observed number of employees, thereby permitting the decisions of firms and workers to depend on the latent expected firm size, not the observed one.

In Section 3, we describe the data and sample construction. By linking several administrative data sources we are able to construct a matched panel dataset of all workers, firms, and corporate boards for the period 2004-2014. This panel data allows us to measure the worker representation status of firms and to follow workers over time, even if they switch between firms. We are also able to estimate the distribution of firm size and wages, particularly around the 30-employee threshold.

In Section 4, we develop an equilibrium model of the labor market with a size contingent law that grants workers the rights to board representation. The model allows for firm-level wage dispersion through three channels. First, firms differ in the amenities they offer to workers. Vertical differentiation of amenities leads to compensating differentials which we allow to be correlated with productivity. Second, workers have idiosyncratic preferences over firm amenities, generating upward-sloping residual labor supply curves. As a result, more productive firms pay higher wages as they hire more workers. Third, the size-contingent policy may directly affect firm-level productivity, amenities, and (fixed and variable) labor costs, thereby influencing the production and the wage setting.

We show that there exists an interval of productivity levels over which firms find it optimal to bunch at the policy threshold to avoid the regulation. Firms can be classified into three groups based on their productivity: (i) unregulated firms with less than 29 employees, (ii) bunching firms that cap size at 29 employees to avoid regulation, and (iii) regulated firms that optimally operate above the threshold despite the cost.

Next, we decompose the effect of the policy on average profit and wages in the economy into three components. The direct policy effect captures, for each regulated firm, the difference between actual profit (wage) and the profit (wage) they would have had in the absence of the regulation, keeping the wages and amenities at other firms fixed. The size distortion effect captures the profit and wage loss among bunching firms which distort their size to avoid the regulation. The equilibrium effect captures the effect on profit, for each firm in the economy, of changes in an economy-wide wage index capturing amenities and wages paid by other firms.

In Section 5, we present key data moments, and discuss how they are used to identify and estimate the model using GMM. The first key data moments are the information from the distribution of firm size. The Pareto tail index of the firm size distribution is directly informative of the distribution of firm productivity, the elasticity of residual labor supply, and the correlation between productivity and amenities. The degree of bunching around the policy threshold is informative about the net cost to firms of the regulation. Intuitively, if the policy imposes a significant burden on firms, they would bunch at the threshold to balance the cost of regulation against deviations from their optimal size.

The second set of key moments are the changes in wages and value-added around the threshold. In the absence of bunching, the change in these outcomes at the threshold would recover the direct effect of the policy. In the presence of bunching, the change in these outcomes captures both the direct effect of the policy and a term reflecting the sorting of different types of firms around the threshold. We show that it is possible to isolate the direct effect of the policy by combining these moments with other moments from the data, including information from the firm size distribution.

The final key moment is the wage premium associated with working in a firm with worker representation. We estimate this premium using a mover design that compares wage changes for pairs of co-workers, where one worker moves to a firm with worker representation and the other to a firm without. The estimated wage premium is 4.3%. We show that this worker representation premium captures a combination of the direct effect of the policy and an "omitted variable bias" due to the correlation between worker representation and firm size. Since residual labor supply curves are upward-sloping, larger firms must pay workers a higher wage.

We estimate the model jointly using GMM, where the objective function captures model deviations from the moments discussed above. In Section 6, we use the estimated model to draw inference about the effects of the policy on firms and workers. We begin by estimating the net cost of the regulation for marginal firms (i.e. those located at the threshold). Since we find little if any bunching at the policy threshold, we conclude that the net cost is small.

Because we allow the policy to affect both fixed and variable costs, the impact on the average regulated firm may differ from that on marginal firms. The policy may also influence firms below the threshold due to equilibrium effects in the labor market. As a result, the model is essential for understanding the policy's broader economic impact and for computing aggregate effects. In the baseline model, the total effect of the policy on average firm profit is a precisely estimated decrease of 0.50%, equivalent to approximately 108 USD per worker per year. This modest total effect combines a small negative direct impact on regulated firms and a small positive equilibrium effect. In a model that allows for measurement error in firm size, the direct effect becomes positive and borderline statistically significant. While we can rule out negative effects on aggregate profits of the board representation policy at conventional levels of significance, it becomes difficult to draw firm conclusions about the magnitude of positive effects, if any. For workers, in the baseline model, the policy reduces welfare by the equivalent of a uniform 0.04% wage cut, driven entirely by the direct effect on the wage at regulated firms. Allowing for measurement error in firm size increases this welfare loss to the equivalent of a uniform 0.11% wage cut.

One possible explanation for the lack of significant effects is that minority board representation offers workers limited influence over firm strategy or compensation policies. This interpretation is consistent with Bertrand et al. (2019), who find that gender quotas on corporate boards had very little discernible impact on women's careers in firms beyond its direct effect on the women who made it into boardrooms.

Our paper contributes to a growing literature on co-determination in corporate governance. A series of papers have studied worker representation on the supervisory boards of German corporations, using (changes in) the regulations governing workers right to representation to estimate the effect of the policy on marginal firms and workers. Gorton and Schmid (2004) exploit a discontinuity in the rights to worker representation, with workers' share of the seats on the supervisory board increasing from one third to one half at a firm size threshold of 2,000 employees. More recent contributions using the same firm size threshold for identification include Lin et al. (2018), Kim et al. (2018), and Redeker (2022). These studies extend the work of Gorton and Schmid (2004) in several ways, including studying different outcomes. Jäger, Schoefer and Heining (2021) study the effects of a reform that abolished the rights to worker representation for newly incorporated firms while leaving older firms unaffected. They find that worker representation on supervisory boards have no effect on affected firms' wages and productivity. Recent work by Harju, Jäger and Schoefer (2024) examines the introduction of formal information-sharing institutions for larger firms in Finland. These institutions take the form of representation both on corporate and advisory boards, primarily designed to increase information sharing between workers and management. They find no effects of formal information-sharing institutions on voluntary job separations, at most small positive effects on other measures of job quality, and positive effects on value added per worker. While the above studies have focused on the effects of co-determination laws on directly affected firms and workers, little is known about the aggregate and equilibrium effects of such policies. One exception is Jäger, Noy and Schoefer (2022), who examine the relationship between co-determination laws and macroeconomic outcomes using cross-country data.

2 Institutional Context and Worker Representation Policy

Norway is a small open economy with a well-educated and healthy population. The labor market is characterized by a relatively compressed wage distribution driven by smaller returns to skill compared with other OECD countries. Although there is relatively low inequality in labor income, capital income and wealth are highly concentrated. Collective bargaining agreements are common in the private sector. However, there is substantial room for individual firms to adjust wages. In the most common form of collective agreements, general guidelines and wage floors are established at the industry-level while wages for individual workers are set at the firm-level.[3]

Regulations about worker representation. In 1972, workers' right to representation on the corporate board was established as an institution with legal regulations provided by the Norwegian Limited Liability Companies Act ("Aksjeloven"). Worker representatives serve on the board with the same rights and responsibilities as regular board members, and can propose and vote for firm-wide bonus schemes, executive pay schemes, and influence the overall strategy of the firm. For this reason, firms may have an incentive to discourage or

[3] See Mogstad et al. (fothcoming) for a detailed discussion of the Norwegian labor market and wage setting.

prevent the election of a worker representative on the board if the interests of the shareholders and workers are not aligned.

Under Norwegian law, workers' right to representation on the corporate board is a discontinuous function of firm size: In firms with less than 30 employees, workers have no rights to representation on the board; in firms with between 30 and 50 employees, workers can demand at least one representative on the corporate board; in firms with between 50 and 200 employees, workers can demand at least two representatives and up to one third of the board members; and in firms with more than 200 employees, worker representation on the corporate board is mandatory unless workers are represented on the firm's corporate assembly or on the corporate board of a different firm within the same corporate group (a group of firms with common ownership).

Since firm size is a choice, the firm can avoid the threat of worker representation by choosing a firm size below the cut-off of 30 employees. Firms expecting to be only temporarily above a given firm size threshold can also apply to be exempt from the regulation, and the general regulation does not apply to firms operating in certain industries.[4]

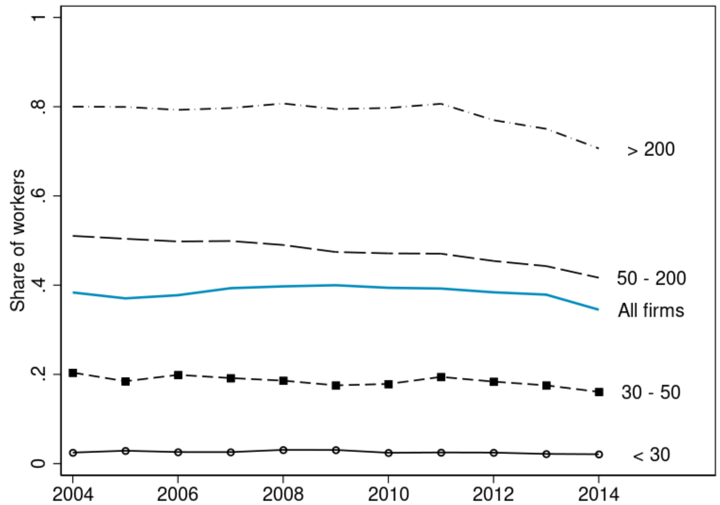
Adoption of worker representation in firms. Empirically, worker representation on the corporate board is a pervasive institution in the Norwegian labor market. Figure 1 shows that about 40% of workers are employed by firms with at least one worker representative on the corporate board. The prevalence of worker representation increases with firm size: The share of workers represented on the corporate board ranges from about 3% for workers in firms with fewer than 30 employees to nearly 80% in firms with more than 200 employees.[5] In Figure 2a, we zoom in on take-up around the first policy cutoff. Even though board representation is increasing in firm size, we find no evidence of a clear discontinuity at the 30-employee regulation threshold.

The smoothness in take-up around the policy cut-off has at least two potential explanations which will inform our modeling choices. First, while the *right* to worker representation changes discontinuously at 30 employees, workers do not necessarily have to exercise this right for the policy discontinuity to have an impact. As emphasized by Jäger et al. (2021), a policy granting workers representation rights is likely to affect firm behavior even if the work-

[4]Exempt industries include newspapers and media, bank and insurance, and off-shore drilling. These industries are subject to other regulations.

[5]Among the workers in firms with more than 200 employees and no worker representative on the corporate board, 51 percent are represented on the corporate board of a different firm within the same corporate group. Worker representation on the corporate assembly is not observed in our data.

Figure 1: Share of workers represented at the corporate board by employer firm size.



Notes: This figure shows the share of workers employed by a firm with at least one worker representative on the corporate board. Firms are categorized by the number of employees in January of year $t-1$, and each line represents a separate firm size category. The figure is constructed using the full sample defined in Section 3.2.

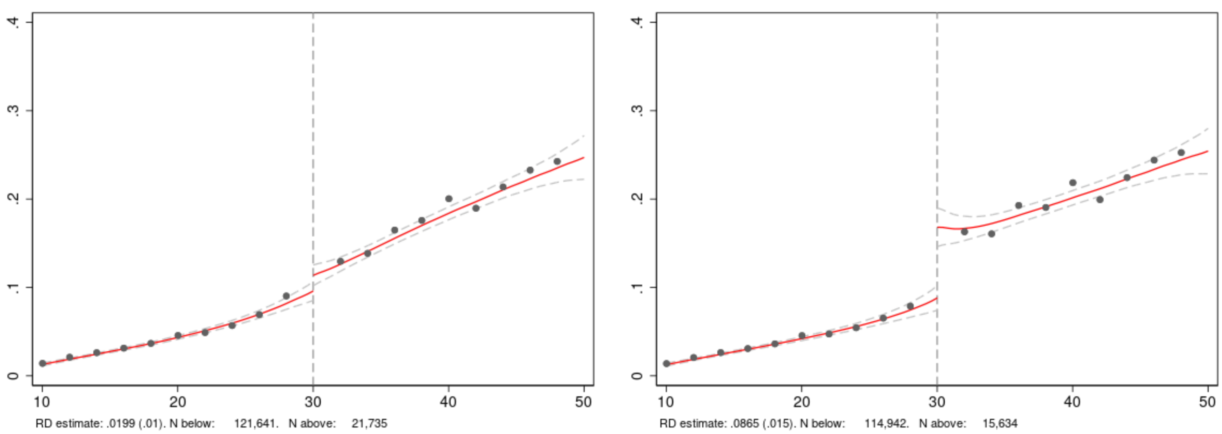
ers do not elect a representative. The reason is that workers can then credibly threaten to demand worker representation, affecting the bargaining power and outcomes between workers and firms.[6] In the main analysis, we therefore assume that the policy directly affects all firms above the 30 employee threshold, not just those that implement worker representation. In Appendix C, we specify and estimate an alternative model in which the policy discontinuity only has a direct effect on firms that implement worker representation.

Another possible explanation for the smoothness in take-up around the policy cut-off is "measurement error" in the sense that the observed number of employees in the data might not reflect the latent expected firm size that workers and firms consider when making their decisions. Unless the firm chooses to adopt representation voluntarily, workers would need to know that the firm employs more than 30 workers and petition for representation before they can elect a representative. If the firm is only temporarily above the cutoff, workers may not notice, or by the time they have noticed, the firms can argue that they are only temporarily above the cutoff and apply for an exemption from the regulation. Since there are significant fluctuations in the number of employees throughout the year, the observed

[6]This point is often referred to as "bargaining in the shadow of the law". As emphasized by Stevenson and Wolfers (2006), legal rights may influence outcomes even when they are not exercised.

number of employees in the data might not reflect the latent expected firm size that workers and firms consider when making their decisions. In Figure 2b, we remove some of the measurement error in firm size by restricting to firms that remained consistently on each side of the cutoff for two consecutive years, $t-2$ and $t-1$. In this selected sample of firms, there is a significant discontinuity in the take-up of the policy at the cut-off. While this is a selective subsample of the firms, the presence of a discontinuity is consistent with measurement error in the observed distribution of firm size. In our model and empirical analysis, we will therefore allow for measurement error in the observed number of employees, thereby permitting the decisions of firms and workers to depend on the latent expected firm size, not the observed one.

Figure 2: Share of firms with a worker representative by firm size



(a) All firms

(b) Firms with persistence in firm size

Notes: These figures plot the share of firms with at least one worker representative on the corporate board in each firm size bin. The dashed vertical lines denote the regulation's firm size cut-off of 30 employees. The sample used is the discontinuity sample as defined in Section 3.2. In Panel (b), the discontinuity sample is restricted to workers in firms with between 10 and 50 employees in year $t-1$ which remained above or below the cut-off in years $t-1$ and $t-2$. Standard errors are clustered at the firm-level.

3 Data and Descriptive Evidence

Below we describe our data and sample selection. Details about the data sources and each of the variables are given in Appendix Section A.

3.1 Data

We use several administrative data sources that we are able to link together using unique identifiers for firms and workers. This results in a matched panel dataset with detailed information on the characteristics and outcomes of Norwegian workers, firms, and corporate boards for the period 2004-2014.

Our main source of data on employment and workers' compensation is a matched employer-employee panel data set, consisting of annual tax records of the universe of workers that are matched to non-pecuniary information about employment from the Norwegian Labor and Welfare Administration. This register is used in the administration of sickness benefits and therefore subject to extensive quality controls. The dataset includes information on total earnings, contracted hours, and the number of days worked at each job. Earnings include fixed salary, bonus, overtime, and vacation and severance pay, but exclude sickness benefits. We construct hourly wages using annual earnings and contracted hours, adjusting for the number of days worked.

Our firm data draws on several administrative registers maintained by the Bronnoysund Register Center. We obtain information from firms' income statements and balance sheets, including revenue and cost of inputs, from the Register of Company Accounts. This register covers the universe of limited liability firms – the most common legal entity type of firms and also the population affected by the regulation governing the rights to worker representation. In our analysis, we define value added as revenue net of cost of inputs. We merge the income statement and balance sheet data with information on the industry and geographic location of each firm from the Central Register of Establishments and Enterprises.

Lastly, starting in 2004 we are able to merge in administrative data on the composition of boards of directors from the Register of Legal Entities. Firms are required by law to report the identity and the role of each director, including whether each director was elected by and among the employees or by the shareholders. This allows us to observe the worker representation status of each firm from 2004 onward, and to measure adoption of worker representation for firms adopting worker representation on the corporate board for the first time.

3.2 Sample selection and descriptive statistics

We construct our baseline sample using workers between 25 and 60 years old and define the highest-paying job in each year as the worker's main job. We exclude firms operating in industries which are exempt from the regulation governing worker representation (1.5% of firms, 6.6% of workers), and we also drop a small number of observations with missing value added, industry, or region (5% of firms, 5% of workers). With these restrictions, our sample consists of about 1.5 million workers and 128 thousand firms.

We estimate the moments used to inform the model parameters using two different subsamples. Table 1 summarizes the sample restrictions, and compares the characteristics of the full sample and each subsample. The first subsample, which we refer to as the *discontinuity sample,* restricts the full sample to workers employed by firms in a window of 20 employees on either side of the firm size discontinuity in the rights to representation at the extensive margin, which occurs at 30 employees. We exclude firms with more than 50 employees because there is another discontinuity in the rights to representation at this firm size.[7] The discontinuity sample will be used to estimate discontinuities in the firm size, wage, and value added distribution which are informative about the effects of the policy through the lens of the model.

The second subsample, which we refer to as the *movers sample*, restricts the full sample to workers observed transitioning from firms without worker representation, to firms with a worker representation on the corporate board (treatment group workers) and their co-workers switching between firms with the same representation status (control group workers). To make sure that the treatment and control group workers are moving from similar types of firms, we restrict the sample to firms with at least one treatment and one control group worker who are switching jobs in the same year. The worker representation status of each firm is measured two years before the worker switches to another firm. We focus on the first full-time job-to-job transition for each worker and consequently exclude moves where the individual is observed claiming unemployment insurance benefits in the year of the move.

Table 1 presents summary statistics for the full sample, as well as for the discontinuity and mover samples. By construction, the discontinuity sample excludes the smallest and largest firms, resulting in a substantially lower average firm size. Consistent with the well-documented positive correlation between firm size, wages, and value-added per worker, these latter two variables are also slightly lower in the discontinuity sample. In contrast, the mover sample more closely resembles the full sample in terms of observable characteristics.

[7]Even though the regulation creates several discontinuities in the rights to representation, we focus on the discontinuity in the rights to worker representation created by the 30 employees threshold, i.e. on the extensive margin of worker representation. This is the most relevant margin for policymakers interested in the effects on wages of introducing worker representation on the corporate board of a given firm.

Table 1: Firm characteristics

	Full sample	Discontinuity sample	Mover sample
Worker representative on the corporate board	0.38	0.09	0.05
Firm size	1,027.94	34.30	428.02
Hourly wage (NOK)	241.75	227.02	234.83
Value added pr. worker (NOK)	1,289.69	1,069.31	1,068.03
Nr of workers	1,561,750	627,141	9,011
Nr of firms	127,748	28,589	4,874

Notes: This table presents summary statistics for the main analysis samples. All variables are defined in Appendix Table A.1. The characteristics of the mover sample are measured two years before the move.

4 Model

We now develop an equilibrium model of the labor market with a size contingent law that grants workers the rights to board representation.

4.1 Agents, preferences, and technology

The economy is composed of workers with mass E_0, indexed by i and a unit mass of firms indexed by j. For tractability, we assume that workers and firms face exogenous birth-death processes which ensures stationarity in the productivity distribution of firms and the preference distribution of workers.

Worker. We assume that worker i's utility from working at firm j in period t is given by:

$$u_{ijt} = \beta \log (w_{jt}) + a_{jt} + \varepsilon_{ijt}$$

where w_{jt} is the wage offered at firm j in period t, a_{jt} captures the quality of workplace amenities, and ε_{ijt} reflects worker i's idiosyncratic preferences for working at firm j. We assume that the cross-sectional distribution of ε_{ijt} in period t is Type 1 Extreme Value (T1EV).

This preference specification allows workers to view firms as imperfect substitutes. The term a_{jt} introduces vertical differentiation among employers, while ε_{ijt} captures horizontal differentiation. The relative importance of idiosyncratic workplace amenities to wages is governed by β; as β increases, wages weigh more heavily in worker decisions. Horizontal differentiation among employers generates monopsony power, as firms face a residual labor supply curve that is not perfectly elastic.

Technology. The value added Y_{jt} generated by firm j in period t is determined by a production function with constant returns to scale:

$$Y_{jt} = R_{jt} E_{jt}$$

where R_{jt} is the firm's productivity (TFP), and E_{jt} is labor employed by the firm. The specification of the value added function abstracts from capital, or equivalently, assumes that capital can be rented at some fixed price.[8]

Firms are heterogeneous in productivity R_{jt} and in the amenities a_{jt} they offer to workers. We follow Garicano et al. (2016) in assuming that the cross-sectional distribution of productivity in period t is a Pareto distribution:

$$F_R(R) = 1 - \left(\frac{R_{min}}{R}\right)^{\gamma_R}$$

where R_{min} is the minimal productivity in the economy and γ_R is the Pareto tail index. The vertical component of amenities is a function of productivity:

$$a_{jt} = \theta \log(R_{jt})$$

where $\theta \in (-\beta, +\infty)$. If $\theta > 0$, workers prefer working at more productive firms, even holding wages constant, because these firms offer better amenities.

Labor Market. We consider an environment where all labor is hired in a spot market and ε_{ijt} is private information to the worker. As a result, firm j offers a unique wage to all workers at time t. Then, in period t, the number of workers that choose to work at firm j is given by:

$$E\left(w_{jt}, R_{jt}, \tilde{W}_t\right) = \frac{\exp\left(\beta \log(w_{jt})\right)}{\exp\left(\tilde{W}_t\right)} R_{jt}^{\theta} E_0$$

where $\tilde{W}_t \equiv \log\left[\sum_{k=1}^{J} \exp\left(a_{kt} + \beta \log\left(w_{kt}\right)\right)\right]$ is the wage index of the economy. Because

[8] More precisely, we can derive the same linear specification of the value added production function under the assumptions that (i) the production function is Cobb-Douglas in labor and capital with constant returns to scale, and (ii) capital can be rented at a fixed price.

of our assumption that the productivity and preference distributions are stationary, we can drop the time subscript on the wage index. We also assume that each firm is strategically small so that, when setting their wage to maximize profit, firms ignore the effect of their own wage on the wage index.

4.2 Size-dependent regulation

We introduce a size-dependent regulation in the economy. Firms choose their wage and employment level taking into account the effect of the regulation on their profit. We denote the presence of worker representation on the corporate board by $B_{jt} \in \{0, 1\}$. The size-dependent probability of having a worker representative on the board is:

$$\mathbb{P}(B_{j,t} = 1 | E_{jt}) = \begin{cases} P_0, & \text{if } E_{jt} \leq E_c \\ P_1, & \text{if } E_{jt} > E_c \end{cases}$$

where $P_0 < P_1$, reflecting that firms exceeding the threshold E_c are more likely to have a worker representative on the board due to the policy. Under the size-dependent regulation, the firm's profit is given by:

$$\pi\left(R_{jt}\right) \equiv \max_{w} \left\{ \begin{matrix} (R_{jt} - w) \times E\left(w, R_{jt}, \tilde{W}\right) & \text{if } E\left(w, R_{jt}, \tilde{W}\right) \leq E_c \\ (\delta R_{jt} - \tau w) \times \kappa E\left(w, R_{jt}, \tilde{W}\right) - F & \text{if } E\left(w, R_{jt}, \tilde{W}\right) > E_c \end{matrix} \right\}$$

where δ, κ, τ, and F are parameters that capture the effect of the policy.

The policy affects firms in two main ways. First, having a worker representative on the board could affect total factor productivity (TFP). We allow the effect to be positive or negative. A positive effect could arise if worker representation improves information sharing between workers and the management as in Freeman and Lazear (1995). The effect may instead be negative if worker representation constrains managerial discretion or delays restructuring efforts, potentially lowering productivity. Specifically, we model the productivity effect of worker representation as a parameter δ which shifts TFP above the policy threshold:

$$TFP = \begin{cases} R_{jt} & \text{if } E_{jt} \leq E_c \\ \delta R_{jt} & \text{if } E_{jt} > E_c \end{cases}$$

The second way in which the policy may affect firms is that worker representation could influence the provision of costly amenities to workers. On the one hand, the amenities are valued by workers which we model as a parameter κ that shifts the relationship between productivity and amenities above the firm size cut-off:

$$\exp(a_{jt}) = \begin{cases} R_{jt}^{\theta} & \text{if } E_{jt} \leq E_c \\ \kappa R_{jt}^{\theta} & \text{if } E_{jt} > E_c \end{cases}$$

Thus, we allow worker representation to affect hard-to-measure outcomes which are valued by workers such as better workplace safety and health or worker job satisfaction. On the other hand, the cost of providing these amenities may have a negative effect on firm profits. The cost is composed of a fixed component, F, and a variable component, τ, which scales with employment:

$$\text{Labor Cost} = \begin{cases} w_{jt}E_{jt} & \text{if } E_{jt} \leq E_c \\ \tau w_{jt}E_{jt} + F & \text{if } E_{jt} > E_c \end{cases}$$

4.3 Equilibrium

Consider the firm's problem below the threshold and define R_c as the productivity level such that E_c is an interior solution. Formally, letting $w_0^u(R) \equiv \text{argmax}_w \left\{(R - w) \times E\left(w, R, \tilde{W}\right)\right\}$, we define R_c such that:

$$E_c \equiv E\left(w_0^u(R_c), R_c, \tilde{W}\right)$$

Solving for R_c, we have:

$$R_c \equiv \left(\frac{1+\beta}{\beta}\right)^{\frac{\beta}{\beta+\theta}} \left(\frac{E_c}{E_0} \exp\left(\tilde{W}\right)\right)^{\frac{1}{\beta+\theta}}$$

Though we do not restrict the sign of each individual effect of the policy (i.e. $\kappa \lesseqgtr 1, \delta \lesseqgtr 1$, $\tau \lesseqgtr 1$, and $F \lesseqgtr 0$), we assume that the net effect of the policy on firm profits, at the employment threshold E_c, is (weakly) negative. Formally, let us define the following two profit functions:

$$\pi_0\left(R, \tilde{W}\right) \equiv \max_w \,(R - w) \times E\left(w, R, \tilde{W}\right) \text{ s.to. } E\left(w, R, \tilde{W}\right) \leq E_c \tag{1}$$

$$\pi_1\left(R, \tilde{W}\right) \equiv \max_w \,(\delta R - \tau w) \times \kappa E\left(w, R, \tilde{W}\right) - F \tag{2}$$

Assumption 1. *At the employment threshold, the net effect of the policy on firm profits is negative:*

$$\pi_0\left(R_c, \tilde{W}\right) \geq \pi_1\left(R_c, \tilde{W}\right)$$

which is equivalent to:

$$\frac{F}{R_c E_c} \geq \left[\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta} - 1\right]\frac{1}{1+\beta}.$$

Under Assumption 1, we can derive equilibrium employment and wages as functions of firm productivity. Due to the cost imposed by the policy, there exists an interval of productivity levels over which firms find it optimal to bunch at the policy threshold to avoid these costs. Provided that the net magnitude of the scalable policy parameters (δ, κ, and τ) is not too large, this bunching interval is bounded. The following proposition formalizes this result:

Proposition 1. *Under Assumption 1, if $\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta} \geq 1$, there exists some productivity level $R_r \in [R_c, +\infty)$ such that:*

$$E_{jt} = \begin{cases} \left(\frac{\beta}{1+\beta}\right)^{\beta} \frac{E_0}{\exp\left(\tilde{W}\right)} R_{jt}^{\beta+\theta} & \forall R_{jt} \leq R_c \\ E_c & \forall R_c \leq R_{jt} < R_r \\ \kappa\left(\frac{\delta}{\tau}\frac{\beta}{1+\beta}\right)^{\beta} \frac{E_0}{\exp(W)} R_{jt}^{\beta+\theta} & \forall R_{jt} \geq R_r \end{cases}$$

and

$$w_{jt} = \begin{cases} \frac{\beta}{1+\beta} R_{jt} & \forall R_{jt} \leq R_c \\ \left(\frac{E_c}{E_0}\exp\left(\tilde{W}\right)\right)^{\frac{1}{\beta}} R_{jt}^{-\frac{\theta}{\beta}} & \forall R_c \leq R_{jt} < R_r \\ \frac{\delta}{\tau}\frac{\beta}{1+\beta} R_{jt} & \forall R_{jt} \geq R_r \end{cases}$$

If $\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta} < 1$, then no firm ever chooses a size above E_c.

The proof is given in Appendix B.1. In the rest of the paper, we focus on the empirically relevant case where the bunching interval is bounded.

4.4 Equilibrium firm size distribution

Let us define the employment level associated with the upper limit of the bunching interval as

$$E_r \equiv \kappa\left(\frac{\delta}{\tau}\frac{\beta}{1+\beta}\right)^{\beta} \frac{E_0}{\exp\left(W\right)} R_r^{\beta+\theta} \geq E_c.$$

Given the distribution of firm productivity $F_R(.)$, we can derive the corresponding equilibrium distribution of firm size, which follows a "broken power law":

$$Pr\left(E_{jt} < E\right) = \begin{cases} 1 - XE^{-\bar{\beta}}, & E < E_c \\ 1 - XTE_r^{-\bar{\beta}}, & E_c \leq E < E_r \\ 1 - XTE^{-\bar{\beta}}, & E \geq E_r \end{cases}$$

where $\overline{\beta} \equiv \frac{\gamma_R}{\beta+\theta}$, $X \equiv R_{min}^{\gamma_R}\left(\frac{\beta}{1+\beta}\right)^{\frac{\gamma_R\beta}{\beta+\theta}}\left(\frac{E_0}{exp(\tilde{W})}\right)^{\frac{\gamma_R}{\beta+\theta}}$, and $T \equiv \kappa^{\frac{\gamma_R}{\beta+\theta}}\left[\frac{\delta}{\tau}\right]^{\frac{\beta\gamma_R}{\theta+\beta}}$. The associated equilibrium firm size density, which we will take to the data, is given by:

$$\chi^*(E) = \begin{cases} \overline{\beta}XE^{-\left(1+\overline{\beta}\right)} & \text{if } E < E_c \\ X\left(E_c^{-\overline{\beta}} - TE_r^{-\overline{\beta}}\right) & \text{if } E = E_c \\ 0 & \text{if } E_c < E < E_r \\ \overline{\beta}TXE^{-\left(1+\overline{\beta}\right)} & \text{if } E \geq E_r \end{cases} \tag{3}$$

The "broken power law" distribution of firm size is of the same parametric form as in Garicano et al. (2016), despite being derived from a different model. In Appendix B.3.1, we derive the joint distribution of firm size E_{jt} and worker representation B_{jt}, which forms the basis of our maximum likelihood estimation.

Measurement error in firm size. We consider two versions of the model. The baseline model assumes firm size is observed without error. Following Garicano et al. (2016), we also consider a model with measurement error, where observed employment E_{jt} is given by:

$$E_{jt} = E_{jt}^* \exp(\varepsilon), \quad \varepsilon \sim \mathcal{N}(0, \sigma),$$

where E_{jt}^* is the true employment of the firm, and σ is the variance of the measurement error. The joint distribution of firm size and worker representation under measurement error is derived in Appendix B.3.2. We estimate both the baseline model and the model incorporating measurement error and compare the results throughout the analysis.

4.5 Total effects of the regulation

Next, we consider the aggregate effects of the policy on firm profits, average worker wages, and average worker welfare.

4.5.1 Aggregate effects of the policy on firms

A key insight from Proposition 1 is that firms can be classified into three groups based on their productivity: (i) *unregulated firms* with productivity $R \in [R_{\min}, R_c]$, (ii) *bunching firms* with $R \in (R_c, R_r)$, and (iii) *regulated firms* with $R \in [R_r, +\infty)$. Define $\pi_0^u(R, \tilde{W})$ as the profit function for unregulated firms, $\pi_0^c(R, \tilde{W})$ as the profit for bunching firms, and $\pi_1\left(R, \tilde{W}\right)$ as the profit for regulated firms. Similarly, let ω_0^u, ω_0^c, and ω_1 be the share of firms in each group. Average profits in the economy under the policy, Π_1, can then be expressed as:

$$\Pi_1 \equiv \omega_0^u \times \underbrace{\mathbb{E}\left[\pi_0^u(R, \tilde{W}_1) \mid R_{min} \leq R < R_c\right]}_{\text{Unregulated firms}} + \omega_0^c \times \underbrace{\mathbb{E}\left[\pi_0^c(R, \tilde{W}_1) \mid R_c \leq R < R_r\right]}_{\text{Bunching firms}}$$
$$+ \omega_1 \times \underbrace{\mathbb{E}\left[\pi_1(R, \tilde{W}_1) \mid R \geq R_r\right]}_{\text{Regulated firms}}$$

where $\tilde{W}_1$ denotes the wage index in the economy under the policy. Similarly, let $\tilde{W}_0$ be the counterfactual wage index in an economy without the policy, with counterfactual aggregate profits given by:

$$\Pi_0 \equiv \mathbb{E}\left[\pi_0^u(R, \tilde{W}_0)\right].$$

Using these definitions, the impact of the policy on aggregate profits can be decomposed into three components:

$$\Pi_1 - \Pi_0 = \underbrace{\omega_1 \times \mathbb{E}\left[\pi_1(R, \tilde{W}_1) - \pi_0^u(R, \tilde{W}_1) \mid R \geq R_r\right]}_{\text{Direct policy effect } \equiv \Delta\Pi^{DE}} \tag{4}$$
$$+ \underbrace{\omega_0^c \times \mathbb{E}\left[\pi_0^c(R, \tilde{W}_1) - \pi_0^u(R, \tilde{W}_1) \mid R_c \leq R < R_r\right]}_{\text{Size Distortion Effect} \equiv \Delta\Pi^{SD}} + \underbrace{\mathbb{E}\left[\pi_0^u(R, \tilde{W}_1) - \pi_0^u(R, \tilde{W}_0)\right]}_{\text{Equilibrium Effects } \equiv \Delta\Pi^{EE}}$$

The *direct policy effect* reflects the change in profit for firms directly subject to the regulation, i.e. those firms with firm size above the threshold. For each regulated firm, the direct policy effect is the difference between actual profits and the profits they would have had in the absence of the regulation, keeping the wage index and productivity fixed. The *size distortion effect* captures the profit loss incurred by firms which distort their size to avoid the regulation. Specifically, some firms may choose to remain smaller than they would be in the absence of the policy, thereby reducing aggregate profits. The *equilibrium effect* captures the effect on profit of changes in the wage index due to differences in wages paid and the amenities offered by firms in the two economies.

4.5.2 Aggregate effects of the policy on workers

We define $w_0^u(R)$ as the wage paid by unregulated firms, $w_0^c(R, \tilde{W})$ as the wage paid by bunching firms, and $w_1(R)$ as the wage paid by regulated firms. The average wage under the policy is given by

$$\begin{aligned} W_1 \equiv \omega_0^u &\times \mathbb{E}\left[w^u(R) \mid R_{min} \leq R < R_c\right] \\ &+ \omega_0^c \times \mathbb{E}\left[w_0^c(R) \mid R_c \leq R < R_r\right] \\ &+ \omega_1 \times \mathbb{E}\left[w_1(R) \mid R \geq R_r\right]. \end{aligned}$$

Let $W_0 \equiv \mathbb{E}\left[w_0^u(R)\right]$ be the counterfactual average wage in the absence of the policy. As with profits, the impact of the policy on average wages can be decomposed as:

$$W_1 - W_0 = \underbrace{\omega_0^c \times \mathbb{E}\left[w_0^c(R, \tilde{W}_1) - w_0^u(R) \mid R_c \leq R < R_r\right]}_{\text{Size Distortion Effect}} + \underbrace{\omega_1 \times \mathbb{E}\left[w_1(R) - w_0^u(R) \mid R \geq R_r\right]}_{\text{Direct Policy Effect}}.$$

In Section 6, when presenting the results of this decomposition of wage effects, we replace the productivity distribution $F_R(.)$ with an employment-weighted distribution across firms to make the results representative of the average worker as opposed to the average firm.

Worker Welfare. Under our assumptions on worker preferences, the expected utility across workers is given by:

$$\mathbb{E}\left[\max_j \{u_{ijt}\}\right] = C_E + \tilde{W},$$

where the expectation is taken over the distribution of worker preferences, and $C_E \approx 0.5772$ is Euler's constant. The effect of the policy on average worker welfare can be expressed as:

$$EV = \exp\left(\frac{\tilde{W}_1 - \tilde{W}_0}{\beta}\right) - 1$$

where EV indicates the proportional increase in the wages at all firms that would yield a welfare change equal to the policy (i.e. an average welfare change equal to $\tilde{W}_1 - \tilde{W}_0$).

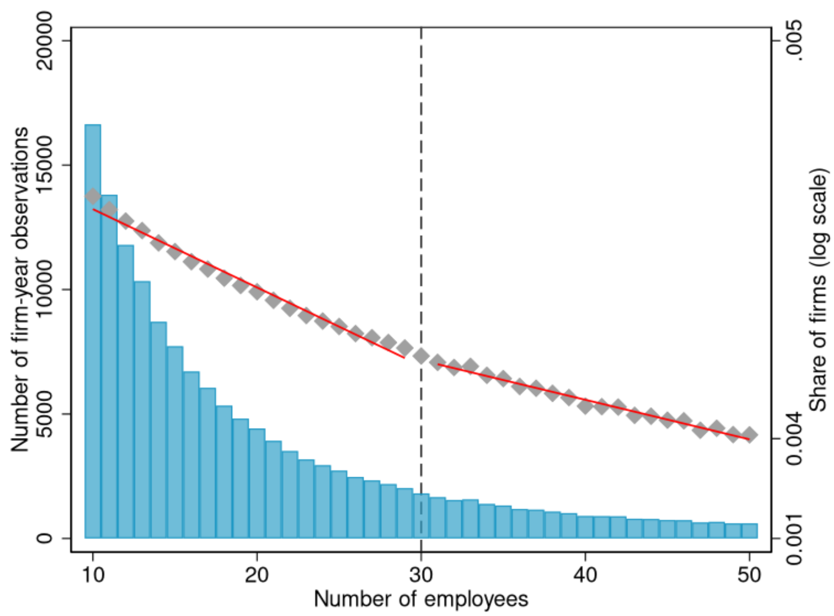
5 Recovering the Model Parameters

In this section, we estimate and present key data moments, and then discuss how they are used to identify and estimate the model using GMM.

5.1 Information from the distribution of firm size

Two features of the firm size distribution inform the model parameters: (i) the degree of bunching around the policy threshold and (ii) the Pareto tail index of the firm size distribution.

Figure 3: Firm size distribution around the policy threshold.



Notes: This figure plots the distribution of firm size around the policy threshold. The sample used is the discontinuity sample defined in Section 3.2. The bars and the y-axis on the left reports the number of firm-year observations in each firm-size bin. The scatter points and the y-axis on the right plots the log share of firms in each firm size bin. The red line fits a straight line to the scatter plot on each side of the policy cut-off.

Bunching at the policy threshold. We begin by examining the degree of bunching around the policy threshold. Intuitively, if the policy imposes a significant burden on firms, they would bunch at the threshold to balance the cost of regulation against deviations from their optimal size. Figure 3 plots the number of firm-year observations by employment level. The scatter points display the log share of firms in each size bin. The roughly linear relationship between log share and firm size suggests that the firm size distribution is well approximated by a Pareto distribution. Importantly, there is no visible bunching around the threshold, indicating that firm size does not substantially respond to the regulation. This finding aligns with evidence from similar size-dependent worker representation policies in Germany and Finland (Lin et al., 2018; Harju et al., 2024), and suggests that the net cost of the policy for firms near the threshold is limited.

This intuition is formalized in the model, where the equilibrium firm size distribution follows a broken Pareto distribution. If the policy imposes costs at the threshold, firms avoid an interval of firm sizes (E_c, E_r), with larger compliance costs leading to a wider interval. Additionally, the missing mass of firms above the threshold results in a level shift in the distribution, captured by T, which is informative about the policy parameters δ, κ, and τ:

$$T = \kappa^{\frac{\gamma_R}{\beta+\theta}} \left(\frac{\delta}{\tau}\right)^{\frac{\beta\gamma_R}{\beta+\theta}}.$$

The shape of the firm size distribution around the threshold is also informative about the variance of the measurement error (σ). As σ increases, the model allows for smoother bunching at the threshold.

Pareto tail index of the firm size distribution The Pareto tail index of the firm size distribution provides information about three parameters: the Pareto tail index of the firm productivity distribution (γ_R), the elasticity of residual labor supply to firms (β), and the relationship between productivity and amenities (θ). Specifically, the slope of the firm size distribution in log-log space in Figure 3 corresponds to the composite parameter $-(1+\bar{\beta})$ from the equilibrium density function $\chi^*(E)$.

Estimates of the firm size distribution moments. We estimate all the moments of the firm size distribution $(\overline{\beta}, T, E_r)$ and σ jointly using maximum likelihood estimation.[9] The estimates and standard errors are reported in Panel A of Table 2. Consistent with Figure 3, we find no evidence of significant bunching at the policy threshold. Specifically, we cannot reject the hypothesis that there is no level shift in the distribution ($T = 1$), and we estimate the bunching interval $E_r - E_c$ to be close to zero in both the specifications with and without measurement error.

[9] Using the restriction that the density of firm size integrates to one, X is a direct function of other parameters of the firm size distribution. See Appendix B.4 for more details.

Table 2: Data Moments

Description	Baseline	Measurement error
(A) Firm size distribution		
Pareto tail index of the firm size distribution ($\bar{\beta}$)	1.047	1.355
	(0.006)	(0.010)
Bunching shifter (T)	0.998	0.991
	(0.001)	(0.005)
Bunching employment (E_r)	29.043	29.000
	(0.050)	(0.173)
Prob. of worker representation above the cut-off (P_1)	0.170	0.195
	(0.003)	(0.003)
Prob. of worker representation below the cut-off (P_0)	0.031	0.027
	(0.000)	(0.001)
Variance of measurement error (σ)		0.123
		(0.005)
(B) Wage distribution		
Worker representation pay premium	0.043	0.043
	(0.004)	(0.004)
RD of wages at the cut-off	0.000	-
	(0.008)	
Average difference in wages across the cut-off	-	0.018
		(0.003)
(C) Value-added distribution		
RD of firm value added at the cut-off	-0.006	-
	(0.024)	
Average difference in firm value added across the cut-off	-	0.408
		(0.008)

Notes: This table presents the estimated data moments. Panel (A) shows the MLE estimates of the shape parameters of the firm size distribution. The sample used is the discontinuity sample as defined in Section 3.2. Panel (B) shows the estimated moments of the wage distribution as defined in the text. The worker representation pay premium is estimated using the mover sample, while wages around, above, and below the policy cut-off are estimated using the discontinuity sample. Panel (C) shows the estimated moments from the distribution of value added. Both moments are estimated using the discontinuity sample.

5.2 Information from the distribution of wages

We use two different moments from the distribution of wages to inform the model parameters. First, we estimate the wage premium associated with working in a firm with worker representation using workers moving between firms. Second, we examine the gap in average wages over the bunching interval. These two moments are informative not only about the direct effects of the policy, but also about the firm-size premium, which is key to disentangling the policy's effects from differences in unobserved firm productivity and worker preferences.

5.2.1 Worker representation pay premium

In Appendix B.2, we decompose the worker representation pay premium into two components: the *direct effect* of the policy on wages and a *firm-size wage premium*. For clarity, we present here the empirically relevant special case in which there is no bunching:

$$\begin{aligned}
&\mathbb{E}\Big[\log(w_{jt}) \mid B_{jt}=1\Big] - \mathbb{E}\Big[\log(w_{jt}) \mid B_{jt}=0\Big] \\
&= \Big[\mathbb{P}(E_{jt} > E_c \mid B_{jt}=1) - \mathbb{P}(E_{jt} > E_c \mid B_{jt}=0)\Big] \\
&\quad \times \Bigg\{ \underbrace{\log\left(\frac{\delta}{\tau}\right)}_{\text{Direct effect of regulation}} + \underbrace{\mathbb{E}[\log(R) \mid R \geq R_r] - \mathbb{E}[\log(R) \mid R < R_r]}_{\text{Firm-size premium from monopsony and amenities}} \Bigg\}.
\end{aligned} \tag{5}$$

The pay premium is the product of two terms: (i) the difference in the probability of being above the firm-size threshold between firms with and without a worker representative, and (ii) the difference in average log wages between firms above and below the threshold. In the absence of bunching, the latter term decomposes into the direct effect of the policy, $\log(\delta/\tau)$, and a firm-size wage premium, reflecting the fact that larger firms typically pay higher wages.

In our model, the firm-size wage premium captures two mechanisms: (i) firms face upward-sloping residual labor supply curves, leading them to pay higher wages as they increase their size, and (ii) firm-specific amenities, which are correlated with size whenever $\theta \neq 0$, generate compensating differentials.

In practice, observed wage differences across firms may also reflect differences in worker quality. Because our model abstracts from worker heterogeneity, we estimate the wage premium net of such differences using a mover design that controls for time-invariant unobserved worker characteristics. Specifically, we implement a difference-in-differences (DiD) framework that compares wage changes for pairs of co-workers, where one worker switches to a firm with worker representation and the other to a firm without.

To visualize the design, we recenter the data so that all moves occur at time zero and estimate the following time-varying DiD specification:

$$w_{i,s} = \alpha_i + \pi_s + \sum_{k \neq -2} \tau_k \mathbf{1}[s = k, G(i) = 1] + \epsilon_{i,s}, \tag{6}$$

where s denotes the year relative to the move, α_i is a worker fixed effect that captures the returns to time-invariant worker characteristics (such as ability), π_s captures time effects, and $G(i)$ is an indicator for switching to a firm with worker representation. The error term $\epsilon_{i,s}$ captures transitory wage shocks. We measure representation status of both origin and destination firms at $s = -2$, ensuring it is predetermined relative to the move. The coefficients of interest τ_k measure the wage effect of working in a firm with worker representation in year k after the move.

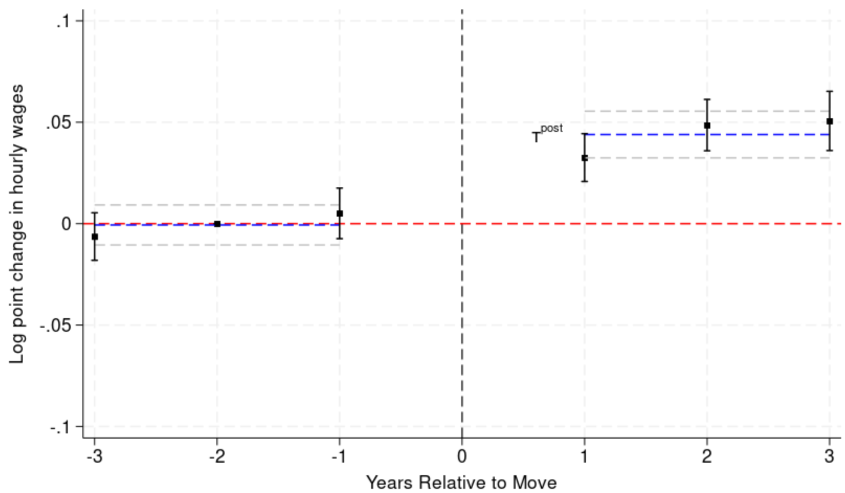


Figure 4: Graphical evidence: Worker representation pay premium

Notes: This figure plots the estimated τ_k coefficients (along with 95% confidence intervals) from the specification in equation (6), for $k \in \{-3, \ldots, 3\}$. The parameter τ^{post} indicates the impact in the post-treatment period ($k > 0$). The treatment group includes workers switching from a firm without representation to a firm with representation, while the comparison group includes workers switching between two firms without representation. The sample used is the movers sample as defined in Section 3.2.

To construct a single estimate of the worker representation pay premium, we aggregate all post-move years ($k > 0$) into a single indicator. We exclude $k = 0$, as wages that year reflect an employment-duration weighted average of wages in the new and old job. Panel B of Table 2 reports the resulting estimate: a 4.3% premium, net of time-invariant worker quality. According to the decomposition in equation (5), this estimate is primarily informative of two key composite parameters: the ratio δ/τ and the composite parameter $\beta + \theta$, which governs the firm-size wage premium.

5.2.2 Wages over the bunching interval

The gap in average wages around the policy threshold depends not only on the direct effect of the policy δ/τ, but also on the composition of firms with different productivity and amenities at these employment levels. Formally, we have:

$$\begin{aligned}&\mathbb{E}\left[\log w_{jt} \mid E_{jt}=E_r\right]-\mathbb{E}\left[\log w_{jt} \mid E_{jt}=E_c\right]= \\ &\qquad \underbrace{\log\left(\frac{\delta}{\tau}\right)}_{\text{Direct Effect}} \underbrace{-\frac{1}{\gamma_R}\log\left(T\right)+\frac{1}{\beta+\theta}\log\left(\frac{E_r}{E_c}\right)+\frac{\theta}{\beta}\mathbb{E}\left[\log\left(\frac{R}{R_c}\right) \mid R_c \leq R \leq R_r\right]}_{\text{Effect of differences in firms' productivity and amenities}} \qquad (7)\end{aligned}$$

In the limit as the bunching interval shrinks (i.e., $E_r \to E_c$), equation (7) simplifies to a regression discontinuity (RD) estimate of log hourly wages around the threshold.[10] If there is no bunching (i.e., if $E_r = E_c$, $T = 1$, and $R_r = R_c$), then the composition effect goes to zero and the RD recovers the *direct effect* of the policy $\log(\delta/\tau)$.

We implement the RD design using a standard specification:

$$w_{j,t} = \pi + \xi B_{j(i)} + h(E_{j(i)}) + \nu_i \qquad (8)$$

where w_i is the log hourly wage of worker i in year t, $j(i)$ denotes the employer of worker i in the same year, $E_{j(i)}$ is the firm's size measured in January of the preceding year, and $B_{j(i)} \equiv \mathbf{1}\{E_{j(i)} \geq 30\}$ is an indicator for whether the firm exceeds the regulatory threshold. The function $h(\cdot)$ is specified as a local linear regression on either side of the cut-off, using a triangular kernel and a bandwidth of 10 employees.

A key advantage of the RD design is its transparency in visualizing potential discontinuities in the wage distribution. Figure 5a plots log hourly wages against firm size. Each point represents the average wage within a firm-size bin, while the fitted local linear regressions highlight trends on either side of the threshold. The y-axis is standardized to ± 0.5 standard deviations of log hourly wages. Visually, there is no evidence of a jump in wages at the threshold. This is confirmed by the RD estimate from equation (8), reported in Table 2, which is 0.00% and not statistically significant at conventional levels.

Even if there were a sharp discontinuity in wages at the threshold based on latent expected

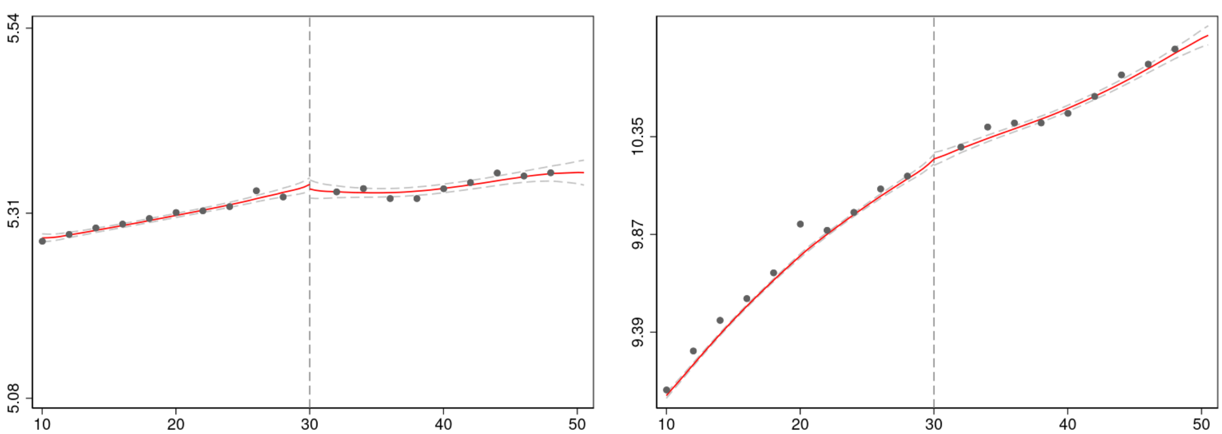
[10] In the limit, the LHS of equation (7) becomes:

$$\lim_{E \to E_c^+} \mathbb{E}\left[\log w_{jt} \mid E_{jt}=E\right] - \lim_{E \to E_c^-} \mathbb{E}\left[\log w_{jt} \mid E_{jt}=E\right].$$

firm size, the presence of measurement error would smooth out any discontinuity in observed firm size. However, if the policy effect were large, we would still expect to observe a distinct (but not discontinuous) change in outcomes in a neighborhood around the threshold, making a local comparison of average outcomes informative. For this reason, when estimating the model with measurement error, we replace the RD with a simple comparison of the average outcome within a bandwidth to the left and to the right of the threshold. We refer to Appendix E.3.2, for details about how we estimate this difference in average outcomes to the left and to the right of the threshold. Panel B of Table 2 reports the resulting estimate: a 1.8% higher average wage for firms above E_r. This positive estimate captures both a potential direct effect of the policy and differences in firm productivity and amenities at these levels of employment. The direct effect can be separately recovered once we estimated the relationship between firm size, productivity, and amenities.

Figure 5: Firm outcomes around the policy threshold

(a) (log) Hourly Wage (b) (log) Value added



Notes: These figures plot the distribution of (log) hourly wages and value added around the policy threshold. The sample used is the discontinuity sample as defined in Section 3.2. The dashed vertical lines denote the regulation's firm size cut-off of 30 employees. Each graph sets the scale of the y-axis equal to ± 0.5 standard deviation of the respective variable. Standard errors are clustered at the firm-level.

5.2.3 Information from the distribution of value added

The gap in average value added over the bunching interval is a combination of the direct effect of the policy through the productivity effect δ and a composition effect:

$$\mathbb{E}\left[\log Y_{jt} \mid E_{jt}=E_r\right]-\mathbb{E}\left[\log Y_{jt} \mid E_{jt}=E_c\right]= \underbrace{\log(\delta)}_{\text{Direct Effect}} + \underbrace{\log\left(\frac{E_r}{E_c}\right)+\mathbb{E}\left[\log\left(\frac{R_r}{R}\right) \mid R_c \leq R \leq R_r\right]}_{\text{Effect of differences in firms productivity and amenities}} \tag{9}$$

Thus, the difference in log value added is particularly informative about the policy parameter δ, which captures the regulation's effect on productivity. Additionally, it reflects the impact of endogenous sorting, as firms with different productivity levels adjust their size around the threshold.

Figure 5b illustrates the effect of the regulation on log value added around the threshold. We apply the same estimation strategy as in equation (8), but with log value added as the dependent variable. The estimated discontinuity in value added is an economically small -0.006% and is not statistically significant at conventional levels.

When estimating the model with measurement error, we replace the RD with a simple comparison of the average log value-added within a bandwidth to the left and to the right of the threshold. We refer to Appendix E.3.2, for details about how we estimate this difference in average outcomes to the left and to the right of the threshold. Panel C of Table 2 reports the resulting estimate.

Fixed cost of the regulation. The indifference condition for the marginal firm with productivity R_r allows us to recover the fixed cost of the regulation F from the other identified parameters of the model:

$$F = E_c R_c \left[\left(\frac{E_r}{E_c}\right)^{\frac{1}{\beta+\theta}} \kappa^{\frac{1}{\beta+\theta}} \left(\frac{\delta}{\tau}\right)^{-\frac{\beta}{\beta+\theta}} \left[\frac{\delta E_r}{(1+\beta)E_c}-1\right]+\frac{\beta}{1+\beta}\right]$$

5.2.4 Externally calibrating the elasticity of residual labor supply

Table 2 shows that the number of parameters to be estimated exceeds the number of available moments, indicating that the model is under-identified. In principle, we could use additional data moments to achieve full identification. For example, following Lamadon et al. (2022), we could leverage the pass-through of value-added shocks to employment and wages. In our model, where amenities depend on productivity, the ratio of pass-through to employment and wages would recover $\beta + \theta$. Including this additional moment would result in a determined non-linear system of equations, rendering the model fully identified.

However, in practice, incorporating the additional moment to estimate all parameters jointly results in an objective function with very low curvature, making it challenging to reliably estimate the parameters. While the sum $\beta + \theta$ is precisely estimated, the data contains little information to separately distinguish between β and θ. Simulation exercises indicate that this issue is particularly severe in the empirically relevant case where the share of firms bunching at the regulation threshold is low. To address this, we fix β, the elasticity of residual labor supply, at a value consistent with the previous literature. This allows us to use six data moments to estimate the remaining six parameters: γ_R, θ, δ, τ, κ, and F.

Fortunately, a well-established literature provides estimates for β across different contexts. Kroft et al. (forthcoming) estimate β in the range $[3.5, 4.1]$, Lamadon et al. (2022) find a value of 4.6, Suárez Serrato and Zidar (2016) estimate an elasticity of residual labor supply around 4.2, and Card et al. (2018) use 4.0 as their preferred value in calibration exercises. Other studies based on experimental variation in wage offers for small tasks or survey experiments generally find estimates in the range $3.0 - 5.0$ (Caldwell and Oehlsen, 2018; Dube et al., forthcoming; Sokolova and Sorensen, 2021). Based on these findings, we set $\beta = 4.0$ for our main estimates and explore the sensitivity of our results to alternative values of β in Section 6.3.

5.3 Estimated Model Parameters

We estimate all model parameters (other than the externally set β) jointly using GMM and present the estimated values in Table 3 for both the baseline specification and the specification that accounts for measurement error.

The first panel of the table reports parameters that characterize the firm productivity distribution and its relationship with amenities. The estimated Pareto tail index of the productivity distribution (γ_R) is 13.17 under the baseline specification and 25.13 when allowing for measurement error. A higher value of γ_R implies a thinner tail, indicating less dispersion in firm productivity. The estimated covariance between productivity and amenities, measured as the ratio θ/γ_R^2, is 0.049 in the baseline case and 0.023 when accounting for measurement error. This suggests a moderate relationship between productivity and workplace amenities, with the magnitude decreasing slightly after correcting for potential mismeasurement of firm size. Lamadon et al. (2022) also find a positive covariance between TFP and workplace amenities. Since amenities increase firm size, a positive correlation reduces the firm-size wage premium. Conditional on firm size, higher amenities imply that a given firm has lower productivity, leading to lower expected wages.

The second panel of the table reports the parameters that capture the impact of the size-dependent regulation on firms. Under the baseline model, the estimates for the effect on productivity ($\delta - 1$), amenities ($\kappa - 1$), and labor costs ($\tau - 1$) are all negative (or zero), economically small (below 1%), and statistically insignificant. The small negative effects on productivity and labor costs affect both wages and profit in opposite directions suggesting that the net effect on the marginal firms and workers is also small.

Under the model with measurement error, the estimates are all positive and slightly larger in magnitude. Even though the effect on amenities is statistically insignificant, we can reject a null effect on labor costs and productivity. However, the positive effects on labor costs and productivity affect both wages and firm profit in opposite directions suggesting that the net effect on marginal firms and workers is relatively small.

Lastly, the estimated fixed cost (F) is -2.47 (thousand USD) in the baseline case and increases to 6.11 (thousand USD) under measurement error. These correspond, respectively, to -0.2 and 0.5% of the marginal firm's revenue ($R_c E_c$). Both estimates are statistically insignificant, and economically small.

6 Model-based insights

In this section, we use the estimated model to draw inference about the effects of the policy for firms and workers using the approach laid out in Section 4.5.

6.1 Effects of the policy on firms

Let us start with the direct effect of the policy on the profits of marginal firms (i.e., firms with productivity R_c). Using the estimates in Table 3, we evaluate:

$$\frac{\pi_1\left(R_c, \tilde{W}\right) - \pi_0\left(R_c, \tilde{W}\right)}{\pi_0\left(R_c, \tilde{W}\right)} = \left[\kappa\tau\left(\frac{\delta}{\tau}\right)^{1+\beta} - 1\right] - (1+\beta)\frac{F}{R_c E_c}$$

Consistent with the observed lack of bunching around the threshold, the direct effect of the policy on marginal firms' profits is approximately zero in the baseline model (0.00% with a standard error of 0.008), and slightly negative but close to zero in the model with measurement error (-0.02% with a standard error of 0.01). However, this net effect consists of both

Table 3: Model parameters

Description	Parameter	Baseline	Measurement error
(A) Shape parameters			
Pareto tail index of productivity distribution	γ_r	13.169	25.133
		(0.391)	(2.113)
Covariance between productivity and amenities	$\frac{\theta}{\gamma_r^2}$	0.049	0.023
		(0.001)	(0.001)
Elasticity of residual labor supply	β	4.000	4.000
(B) Parameters specific to the size-dependent policy			
Productivity	$\delta - 1$	-0.007	0.029
		(0.017)	(0.013)
Non-wage amenities	$\kappa - 1$	0.000	0.005
		(0.026)	(0.023)
Labor cost	$\tau - 1$	-0.007	0.032
		(0.014)	(0.011)
Fixed cost (1,000s USD)	F	-2.47	6.11
		(4.46)	(3.42)

Notes: This table presents the estimated parameter values. The parameters $(\gamma_R, \theta, \delta, \kappa, \tau)$ are estimated jointly using GMM. The fixed cost is backed out as described in the text. P_0, P_1 are estimated using MLE. Standard errors are constructed using 100 replications of block bootstrap where each block is a firm. The elasticity of residual labor supply β is calibrated as described in the text.

variable components that scale with size (δ, τ, and κ) and a fixed cost (F). Consequently, the direct effect of the policy on the average regulated firm may differ from its effect on the marginal firm. Furthermore, in equilibrium, the policy may also impact firms below the threshold, either because they adjust their size to avoid regulation or due to the policy's effect on the wage index. These indirect effects are captured by $\Delta\Pi^{SD}$ and $\Delta\Pi^{EE}$ in the decomposition of aggregate effects discussed in Section 4.5. Thus, the model is essential for extrapolating the policy's impact on the average firm and for calculating the aggregate effects.

We quantify the aggregate effects of the policy on firm profits, with the results presented in Table 4. Drawing on the decomposition in equation (4), we estimate a small and insignificant reduction of -0.50% in average firm profits as a result of the regulation. This corresponds to a reduction of about 108 USD per worker per year. This total effect is composed of a small positive equilibrium effect stemming from the change in the wage index and a negative, but insignificant direct effect of the policy on directly affected firms. Consistent with the absence of bunching in the data, the estimated size distortion effect is zero.

When allowing for measurement error in firm size, the estimated direct and equilibrium effects of the policy become slightly positive and borderline statistically significant. While we can rule out negative effects on aggregate profits of the board representation policy at conventional levels of significance, it becomes difficult to draw firm conclusions about the magnitude of positive effects, if any.

6.2 Effects of the policy on workers

Panel B of Table 4 presents the total effects of the regulation on workers. Compared to an economy without the regulation, average hourly wages are 0.03% lower in the baseline model and 0.20% lower in the model with measurement error. Both effects are economically small and not statistically different from zero. Using the decomposition in equation (4.5.2), we can separate the total effect into the size distortion and the direct policy effect. Consistent with the absence of bunching, we find no significant size distortion effect. The direct effect of the policy is small and statistically insignificant.

Since both wages and amenities enter worker preferences, the policy's impact on worker welfare—or total compensation—depends on both components. Under the baseline model, we find no impact of the policy on amenities, implying that the effect on worker welfare equals the wage effect. Under the model with measurement error, we estimate a small positive effect on amenities that partly offsets the small negative wage effect. The resulting change in worker welfare is -0.11%, meaning that the total change in worker welfare caused by the policy is equivalent to a 0.11% reduction in wages of all workers. These estimates are economically small and not statistically different from zero.

Finally, we present estimates of the components of the worker representation pay premium decomposition. The worker representation pay premium reflects both a potential direct effect of the regulation, $\log(\delta/\tau)$, and a composition effect arising because firms with a worker representative tend to be larger on average (and thus have higher productivity and wages). Our estimates imply that virtually all of the positive wage effect from moving to a firm with a worker representative is driven by the composition effect. This finding derives from (i) the absence of any discernible discontinuity in wages around the threshold and (ii) the absence of bunching ($R_r \approx R_c$). As a result, we estimate $\log(\delta/\tau) = 0$.

Table 4: Impacts of the regulation

		Baseline		Measurement error	
(A) Firm effects		Percent	USD / worker	Percent	USD / worker
Aggregate annual firm profits					
	Direct policy effect	-0.66%	-141.31	1.53%	293.56
		(1.19%)	(256.63)	(0.85%)	(162.36)
	+ Size distortion effect	0.00%	0	0.00%	0
		(0.00%)	(0.00)	(0.00%)	(0.00)
	+ Equilibrium effects	0.15%	32.86	0.44%	84.07
		(0.06%)	(13.96)	(0.23%)	(43.92)
= Total effect		-0.50%	-108	1.97%	378
		(1.20%)	(259.04)	(0.84%)	(160.17)
(B) Worker effects		Percent	USD / hour	Percent	USD / hour
Wage compensation					
	Direct policy effect	-0.03%	0.00	-0.20%	0.00
		(0.55%)	(0.00)	(0.39%)	(0.00)
	+ Size distortion effect	0.00%	0.00	0.00%	0.00
		(0.00%)	(0.00)	(0.00%)	(0.00)
	= Total effect	-0.03%	0.00	-0.20%	0.00
		(0.55%)	(0.00)	(0.39%)	(0.00)
Total compensation		-0.04%	-0.02	-0.11%	-0.04
		(0.02%)	(0.01)	(0.06%)	(0.02)
Worker representation pay premium					
	Direct policy effect	-0.10%		-0.40%	
		(0.64%)		(0.57%)	
	+ Composition	25.18%		13.35%	
		(0.73%)		(1.04%)	
	+ Size distortion effect	0.00%		0.00%	
		(0.00%)		(0.00%)	
	= Difference across threshold	25.09%		12.95%	
		(0.39%)		(0.68%)	
	× Gap in share of large firms	16.98%		32.62%	
		(0.26%)		(0.91%)	
	= Total effect	4.26%		4.22%	
		(0.00%)		(0.17%)	

Notes: This table presents the model-implied aggregate effects of the policy. Standard errors are constructed using 100 replications of block bootstrap where each block is a firm.

6.3 Robustness

We assess the robustness of our results along two dimensions: (i) the assumed elasticity of residual labor supply (β), and (ii) the assumption that the direct effects of the regulation apply to all firms above the 30-employee threshold, regardless of whether they actually adopt worker representation.

Sensitivity to the elasticity of residual labor supply. Appendix Table D.5 reports estimated aggregate effects under alternative values of the residual labor supply elasticity, $\beta \in \{3.0, 5.0\}$. For each value of β, we re-estimate all remaining model parameters and recompute the implied aggregate effects. The results are highly robust: the estimated impacts on profits, wages, and worker compensation remain nearly unchanged across the entire range of β.

Exposure conditional on actual adoption. Our baseline model assumes that the direct effects of the regulation apply to all firms above the threshold, independent of actual adoption. Appendix Section C.1 provides suggestive evidence in support of this assumption. Specifically, we examine firms that elect a worker representative for the first time and compare them to firms that adopt the policy in later years. We find no significant change in wages or value-added per worker around the time of election, suggesting that actual adoption does not materially affect outcomes. This supports the view that the policy's effects, if any, do not operate through the actual implementation of worker representation on the board.

Nevertheless, we consider an alternative version of the model in which the direct effects of the regulation apply only to firms that actually adopt worker representation. This *partial exposure* model may be relevant if, for instance, workers are imperfectly informed about their rights and the presence of a representative serves as a proxy for workers being aware of and acting on those rights. We refer to the original specification as the *full exposure* model.

While the equilibrium firm size distribution under partial exposure remains a broken power law (as in equation (3)), the mapping from observed moments to model parameters differs. Additional details are provided in Appendix C.2.

Appendix Table C.3 reports estimated aggregate effects under the partial exposure model. Results are virtually identical to those under full exposure.

7 Conclusion

We have studied a size-contingent law in Norway that grants workers the right to board representation in firms with 30 or more employees. To analyze the impact of the law, we embedded the regulation into an equilibrium model of the labor market. We showed how behavioral responses to the regulation identify (i) the direct effects of the policy on regulated firms and workers, (ii) the distortions from firms adjusting their size to avoid the regulation, and (iii) the equilibrium effects in the labor market. We evaluated these effects on firm profits and production, as well as on worker compensation, including both wages and non-wage amenities.

Our estimates suggest very small distortions as well as very small direct and equilibrium effects of the policy on firms and workers. One possible explanation for the lack of significant effects is that minority board representation offers workers limited influence over firm strategy or compensation policies.

When thinking about the generalizability of our results, it is useful to observe that developed countries differ in their system for worker representation. For example, Germany has a two-tiered board system, with both a supervisory board and an executive board. Norway, on the other hand, has the same system as in the US, the UK, and several other European countries, with a single-tiered board of directors. Despite the differences in institutions, our estimates of the effects of worker representation are in line with the results from the German setting. This finding suggests that the conclusion about the absence of impacts of worker representation on firms and workers generalizes across shared governance systems that differ markedly in the degree to which workers are given authority in the corporate decision-making.

References

Bertrand, Marianne, Sandra E Black, Sissel Jensen, and Adriana Lleras-Muney, "Breaking the glass ceiling? The effect of board quotas on female labour market outcomes in Norway," *The Review of Economic Studies*, 2019, *86* (1), 191–239.

Caldwell, Sydnee and Emily Oehlsen, "Monopsony and the Gender Wage Gap: Experimental Evidence from the Gig Economy," Technical Report, MIT Working Paper 2018.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," *Journal of econometrics*, 2021, *225* (2), 200–230.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

Dube, Arindrajit, Alan Manning, and Suresh Naidu, "Monopsony and Employer Mis-optimization Explain Why Wages Bunch at Round Numbers," *American Economic Review*, forthcoming.

Freeman, Richard B. and Edward P. Lazear, "An Economic Analysis of Works Councils," in "Works Councils: Consultation, Representation, and Cooperation in Industrial Relations," University of Chicago Press, January 1995, pp. 27–52.

Garicano, Luis, Claire Lelarge, and John Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, November 2016, *106* (11), 3439–3479.

Gorton, Gary and Frank A Schmid, "Capital, labor, and the firm: A study of German codetermination," *Journal of the European Economic Association*, 2004, *2* (5), 863–905.

Harju, Jarkko, Simon Jäger, and Benjamin Schoefer, "Voice at Work," *American Economic Journal: Applied Economics*, 2024.

Jäger, Simon, Benjamin Schoefer, and Jörg Heining, "Labor in the Boardroom," *The Quarterly Journal of Economics*, 2021, *136* (2), 669–725.

_ , Shakked Noy, and Benjamin Schoefer, "What does codetermination do?," *ILR Review*, 2022, *75* (4), 857–890.

Kim, E Han, Ernst Maug, and Christoph Schneider, "Labor representation in governance as an insurance mechanism," *Review of Finance*, 2018, *22* (4), 1251–1289.

Kroft, Kory, Yao Luo, Magne Mogstad, and Bradley Setzler, "Imperfect Competition and Rents in Labor and Product Markets: The Case of the Construction Industry," *American Economic Review*, forthcoming.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, "Imperfect competition, compensating differentials, and rent sharing in the US labor market," *American Economic Review*, 2022, *112* (1), 169–212.

Lin, Chen, Thomas Schmid, and Yuhai Xuan, "Employee representation and financial leverage," *Journal of Financial Economics*, 2018, *127* (2), 303–324.

Mogstad, Magne, Kjell G Salvanes, and Gaute Torsvik, "Income Equality in The Nordic Countries: Myths, Facts, and Lessons," *Journal of Economic Literature*, fothcoming.

Redeker, Nils, "The Politics of Stashing Wealth: The Decline of Labor Power and the Global Rise in Corporate Savings," *The Journal of Politics*, 2022, *84* (2), 975–991.

Serrato, Juan Carlos Suárez and Owen Zidar, "Who benefits from state corporate tax cuts? A local labor markets approach with heterogeneous firms," *American Economic Review*, 2016, *106* (9), 2582–2624.

Sokolova, Anna and Todd Sorensen, "Monopsony in Labor Markets: A Meta-Analysis," *ILR Review*, 2021, *74* (1), 27–55.

Stevenson, Betsey and Justin Wolfers, "Bargaining in the Shadow of the Law: Divorce Laws and Family Distress*," *The Quarterly Journal of Economics*, 02 2006, *121* (1), 267–288.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2021, *225* (2), 175–199. Themed Issue: Treatment Effect 1.

A Data Appendix

Table A.1: Variable definitions

Variable	Description
Labor market outcomes	*Source: Tax Records and the Norwegian Labor and Welfare Administration*
Earnings	Annual pre-tax labor earnings measured at the worker-firm level. Earnings include fixed salary, bonus, overtime, and vacation and severance pay, but exclude sickness benefits.
Hours worked	Number of contracted hours worked each calendar year.
Hourly wage	Annual earnings divided by hours worked. Outliers below the 5th percentile are winsorized.
Hiring rate	Nr. of hires in year t divided by average firm size in year t and $t-1$
Separation rate	Nr. of separations in year t divided by average firm size in year t and $t-1$
Firm accounts	*Source: The Register of Company Accounts*
Revenues	Total annual sales by each firm in a given year.
Cost of inputs	Cost of materials and intermediate inputs for each firm in a given year.
Value added	Revenues minus cost of inputs.
Value added pr. worker	Value added divided by firm size
Corporate boards	*Source: The Register of Legal Entities*
Director's role	Indicates each director's role on the board, e.g. chair vs regular directors.
Elected by	Indicates whether each director is elected by and among the employees or by shareholders.
Firm characteristics	*Source: Central Register of Establishments and Enterprises and the Norwegian Labor and Welfare Administration.*
Firm size	Number of employees measured as the number of registered workers at the beginning of each year, part-time workers are counted as 0.5
Firm age	Number of years since establishment.
Industry	2-digit code classifying a firm's main activity according to the Nomenclature of Economic Activities.

B Proofs and Derivations

B.1 Proof of Proposition 1

Preliminaries: Deriving profit, wage and employment for each productivity level

Firm j's profit maximization problem is:

$$\max_{w} \begin{Bmatrix} (R_{jt} - w) \times E\left(w, R_{jt}, \tilde{W}\right) & \text{if } E\left(w, R_{jt}, \tilde{W}\right) \leq E_c \\ (\delta R_{jt} - \tau w) \times \kappa E\left(w, R_{jt}, \tilde{W}\right) - F & \text{if } E\left(w, R_{jt}, \tilde{W}\right) > E_c \end{Bmatrix} \tag{B.1}$$

Let us first consider the profit maximization problem conditioning on $E_{jt} \leq E_c$. Equation (1) in the main text defines the profit function conditional on having employment below E_c:

$$\pi_0\left(R, \tilde{W}\right) \equiv \max_{w} (R - w) \times E\left(w, R, \tilde{W}\right) \text{ s.to. } E\left(w, R, \tilde{W}\right) \leq E_c$$

Assuming an interior solution, the first-order conditions are:

$$w = \frac{\beta}{1+\beta} R_{jt} \tag{B.2}$$

Plugging (B.2) into the labor supply function, we obtain profit maximizing employment at an interior solution:

$$E = \left(\frac{\beta}{1+\beta}\right)^{\beta} \frac{E_0}{\exp\left(\tilde{W}\right)} R_{jt}^{\beta+\theta} \tag{B.3}$$

Since employment is increasing in productivity R_{jt}, the solution to (B.1) is such that there is some productivity level R_c for which all firms with $R_{jt} \leq R_c$ are at an interior solution while firms with $R_{jt} > R_c$ are at a corner solution and choose $E = E_c$. This level of productivity R_c is given by:

$$R_c \equiv \left(\frac{1+\beta}{\beta}\right)^{\frac{\beta}{\beta+\theta}} \left(\frac{E_c}{E_0} \exp\left(\tilde{W}\right)\right)^{\frac{1}{\beta+\theta}}$$

Firms that are at a corner solution will offer wage w such that their size is $E = E_c$. We find the wage for these firms by inverting the residual labor supply curve:

$$w = \left[\frac{E_c}{E_0} \exp\left(\tilde{W}\right)\right]^{\frac{1}{\beta}} R_{jt}^{-\frac{\theta}{\beta}} \tag{B.4}$$

Plugging optimal employment and wage (as a function of productivity) into the definition of profit, we obtain the profit function:

$$\pi_0\left(R,\tilde{W}\right)=\begin{cases}\frac{1}{\beta}\left(\frac{\beta}{1+\beta}\right)^{1+\beta}\frac{E_0}{\exp\left(\tilde{W}\right)}R^{1+\beta+\theta} & \text{if } R\le R_c\\ \left[R-\left(\frac{E_c}{E_0}\exp\left(\tilde{W}\right)\right)^{\frac{1}{\beta}}R^{-\frac{\theta}{\beta}}\right]E_c & \text{if } R>R_c\end{cases}\tag{B.5}$$

Let us now consider the maximization problem conditional on $E_{jt}>E_c$. Equation (2) in the main text defines the profit function conditioning on having employment above E_c:

$$\pi_1\left(R,\tilde{W}\right)\equiv\max_{w}\left(\delta R_{jt}-\tau w\right)\times\kappa E\left(w,R,\tilde{W}\right)-F$$

The first-order conditions are:

$$w=\frac{\delta}{\tau}\frac{\beta}{1+\beta}R\tag{B.6}$$

Plugging (B.6) into the labor supply function, we obtain the profit maximizing employment:

$$E=\kappa\left(\frac{\delta}{\tau}\frac{\beta}{1+\beta}\right)^{\beta}\frac{E_0}{\exp\left(W\right)}R^{\beta+\theta}\tag{B.7}$$

Plugging into the definition of profit, we obtain the profit function:

$$\pi_1\left(R,\tilde{W}\right)=\kappa\frac{\tau}{\beta}\left(\frac{\delta}{\tau}\frac{\beta}{1+\beta}\right)^{1+\beta}\frac{E_0}{\exp\left(W\right)}R^{1+\beta+\theta}-F\tag{B.8}$$

Lastly, as a sort of normalization, it is useful to re-write the profit function as a function of the ratio R/R_c:

$$\overline{\pi}\left(R/R_c\right)=\max\left\{\overline{\pi}_0\left(R/R_c\right),\overline{\pi}_1\left(R/R_c\right)\right\}\tag{B.9}$$

where

$$\overline{\pi}_0\left(R/R_c\right)=\begin{cases}\frac{1}{1+\beta}\left(R/R_c\right)^{1+\beta+\theta}R_cE_c & \text{if } R/R_c\le 1\\ \left[R/R_c-\frac{\beta}{1+\beta}\left(R/R_c\right)^{-\frac{\theta}{\beta}}\right]R_cE_c & \text{if } R/R_c>1\end{cases}$$

and

$$\overline{\pi}_1\left(R/R_c\right)=\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta}\frac{1}{1+\beta}\left(R/R_c\right)^{1+\beta+\theta}R_cE_c-F$$

Proof of Proposition 1

Proof. First, assume $\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta} \geq 1$. We want to prove that there exists some $R_r \in [R_c, +\infty)$ such that $\overline{\pi}_0\left(R/R_c\right) \geq \overline{\pi}_1\left(R/R_c\right)$ for all $R \leq R_r$, $\overline{\pi}_0\left(R_r/R_c\right) = \overline{\pi}_1\left(R_r/R_c\right)$, and $\overline{\pi}_0\left(R/R_c\right) < \overline{\pi}_1\left(R/R_c\right)$ for all $R > R_r$.

Let us start at $R = R_c$ (i.e. $R/R_c = 1$). By Assumption 1:

$$\overline{\pi}_0\left(1\right) \geq \overline{\pi}_1\left(1\right) \tag{B.10}$$

Taking the limit as $R/R_c \to +\infty$, and using $\beta + \theta > 0$ (recall that we asssume $\theta > -\beta$), we also find:

$$\lim_{R/R_c \to +\infty} \overline{\pi}_0(R/R_c) - \lim_{R/R_c \to +\infty} \overline{\pi}_1(R/R_c) = -\infty < 0 \tag{B.11}$$

Since $\overline{\pi}_0$ and $\overline{\pi}_1$ are continuously differentiable functions, if we prove that $\partial\left(\overline{\pi}_0 - \overline{\pi}_1\right)/\partial(R/R_c) < 0\ \forall R/R_c > 1$, then we obtain the desired result.

Taking the derivative of $\overline{\pi}_0$ and $\overline{\pi}_1$ with respect to R/R_c, we have:

$$\frac{\partial \overline{\pi}_0}{\partial(R/R_c)} = \left[1 + \frac{\theta}{1+\beta}\left(\frac{R}{R_c}\right)^{-\frac{\theta+\beta}{\beta}}\right] R_c E_c \text{ for all } R/R_c \geq 1 \tag{B.12}$$

which is strictly decreasing in R/R_c and equal to $\frac{1+\beta+\theta}{1+\beta} R_c E_c$ evaluated at $R/R_c = 1$. Turning to the derivative of $\overline{\pi}_1(R/R_c)$, we have:

$$\frac{\partial \overline{\pi}_1}{\partial(R/R_c)} = \frac{1+\beta+\theta}{1+\beta} \kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta}\left(\frac{R}{R_c}\right)^{\beta+\theta} R_c E_c \tag{B.13}$$

which is strictly increasing in R/R_c and equal to $\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta}\frac{1+\beta+\theta}{1+\beta} R_c E_c$ evaluated at $R/R_c = 1$. The assumption that $\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta} \geq 1$ guarantees that $\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta}\frac{1+\beta+\theta}{1+\beta} R_c E_c \geq \frac{1+\beta+\theta}{1+\beta} R_c E_c$, thus:

$$\frac{\partial \overline{\pi}_1}{\partial(R/R_c)}(1) \geq \frac{\partial \overline{\pi}_0}{\partial(R/R_c)}(1) \tag{B.14}$$

and

$$\frac{\partial \overline{\pi}_1}{\partial(R/R_c)}(R/R_c) > \frac{\partial \overline{\pi}_0}{\partial(R/R_c)}(R/R_c)\ \forall R/R_c > 1 \tag{B.15}$$

since the LHS of (B.15) is strictly increasing in R/R_c while the RHS is strictly decreasing. Collecting previous results, we have:

- $\overline{\pi}_0(1) - \overline{\pi}_1(1) \geq 0$

- $\lim_{R/R_c \to +\infty} \{\overline{\pi}_0(R/R_c) - \overline{\pi}_1(R/R_c)\} < 0$

- $\partial \left(\overline{\pi}_0 - \overline{\pi}_1\right) / \partial (R/R_c) < 0 \;\; \forall R/R_c > 1$

Thus, given continous differentiability of $\overline{\pi}_0 - \overline{\pi}_1$, there exist a unique $R_r \in [R_c, +\infty)$ such that $\overline{\pi}_0\left(R/R_c\right) \geq \overline{\pi}_1\left(R/R_c\right)$ for all $R \leq R_r$, $\overline{\pi}_0\left(R_r/R_c\right) = \overline{\pi}_1\left(R_r/R_c\right)$, and $\overline{\pi}_0\left(R/R_c\right) < \overline{\pi}_1\left(R/R_c\right)$ for all $R > R_r$.

Now, let us consider the case where $\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta} < 1$. We use a proof by contradiction to show that, in that case, there cannot exist a $R' > R_c$ such that $\overline{\pi}_1\left(R'/R_c\right) > \overline{\pi}_0\left(R'/R_c\right)$.

Indeed, assume (contrary to what we want to prove) that there does exist $R' > R_c$ such that $\overline{\pi}_1\left(R'/R_c\right) > \overline{\pi}_0\left(R'/R_c\right)$. Because $R' > R_c$ it must be that, whenever employment is restricted to be below E_c, a firm with productivity R' would choose to be at a corner solution. Thus, it must be that:

$$\overline{\pi}_0\left(R'/R_c\right) = \left[R'/R_c - \frac{\beta}{1+\beta}\left(R'/R_c\right)^{-\frac{\theta}{\beta}}\right] R_c E_c > \frac{1}{1+\beta}(R'/R_c)^{1+\beta+\theta} R_c E_c$$

Using $\overline{\pi}_1\left(R'/R_c\right) > \overline{\pi}_0\left(R'/R_c\right)$, then:

$$\overline{\pi}_1(R'/R_c) = \kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta}\frac{1}{1+\beta}(R'/R_c)^{1+\beta+\theta} R_c E_c - F > \frac{1}{1+\beta}(R'/R_c)^{1+\beta+\theta} R_c E_c$$

Re-arranging, we have:

$$\frac{F}{R_c E_c} < \underbrace{\left(\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta} - 1\right)\frac{1}{1+\beta}\left(\frac{R'}{R_c}\right)^{1+\beta+\theta}}_{<0 \text{ by assumption that } \kappa\delta(\delta/\tau)^{\beta}<1}$$

Given that the RHS is negative and $R'/R_c > 1$, we have:

$$\frac{F}{R_c E_c} < \left(\kappa\delta\left(\frac{\delta}{\tau}\right)^{\beta} - 1\right)\frac{1}{1+\beta}$$

which directly contradicts Assumption 1. Thus, we conclude that whenever $\kappa\delta\left(\delta/\tau\right)^{\beta} < 1$, there does not exist any $R' > R_c$ such that $\overline{\pi}_1(R'/R_c) > \overline{\pi}_0(R'/R_c)$. Thus, whenever $\kappa\delta\left(\delta/\tau\right)^{\beta} < 1$, no firm ever chooses employment above E_c.

$\square$

B.2 Decomposition of the Mover Estimate

Formally, we are interested in the difference in average log wage between firms with a worker representative and firms without:

$$\mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=1\right]-\mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=0\right]$$

First, using the law of total probability, we write, $\forall i \in \{0,1\}$:

$$\begin{aligned}\mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=i\right] &= \mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=i, E_{jt}>E_c\right] \times Pr\left(E_{jt}>E_c \mid B_{jt}=i\right) \\ &\quad + \mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=i, E_{jt}\leq E_c\right] \times Pr\left(E_{jt}\leq E_c \mid B_{jt}=i\right) \qquad \text{(B.16)}\end{aligned}$$

Note that, under the assumption that the policy directly affects all firms above the regulation threshold (regardless of whether they actually have a worker representative on the board), and only firms above the threshold:

$$\mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=0, E_{jt}>E_c\right] = \mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=1, E_{jt}>E_c\right] = \mathbb{E}\left[\ln\left(w_{jt}\right) \mid E_{jt}>E_c\right] \qquad \text{(B.17)}$$

$$\mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=0, E_{jt}\leq E_c\right] = \mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=1, E_{jt}\leq E_c\right] = \mathbb{E}\left[\ln\left(w_{jt}\right) \mid E_{jt}\leq E_c\right] \qquad \text{(B.18)}$$

Also note that:

$$Pr\left(E_{jt}>E_c \mid B_{jt}=i\right) = 1 - Pr\left(E_{jt}\leq E_c \mid B_{jt}=i\right) \qquad \text{(B.19)}$$

Plugging (B.16), (B.17), (B.18), and (B.19) into the difference in average log wage between firms with a worker representative and firms without, we obtain:

$$\begin{aligned}\mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=1\right] - \mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=0\right] &= \underbrace{\left\{Pr\left(E_{jt}>E_c \mid B_{jt}=1\right) - Pr\left(E_{jt}>E_c \mid B_{jt}=0\right)\right\}}_{(*1)} \\ &\quad \times \underbrace{\left\{\mathbb{E}\left[\ln\left(w_{jt}\right) \mid E_{jt}>E_c\right] - \mathbb{E}\left[\ln\left(w_{jt}\right) \mid E_{jt}\leq E_c\right]\right\}}_{(2*)} \qquad \text{(B.20)}\end{aligned}$$

Term $(*1)$ is positive and captures the fact that under $P_1 > P_0$, conditional on having a worker representative, firms are more likely to be large (i.e. to have size above the regulation threshold). Formally, using Bayes rule, we have:

$$\underbrace{Pr\left(E_{jt} > E_c \mid B_{jt} = 1\right) - Pr\left(E_{jt} > E_c \mid B_{jt} = 0\right)}_{(*1)} =$$

$$\frac{Pr\left(B_{jt} = 1 \mid E_{jt} > E_c\right) \times Pr\left(E_{jt} > E_c\right)}{Pr\left(B_{jt} = 1 \mid E_{jt} > E_c\right) \times Pr\left(E_{jt} > E_c\right) + Pr\left(B_{jt} = 1 \mid E_{jt} \leq E_c\right) \times \left[1 - Pr\left(E_{jt} > E_c\right)\right]}$$
$$- \frac{Pr\left(B_{jt} = 0 \mid E_{jt} > E_c\right) \times Pr\left(E_{jt} > E_c\right)}{Pr\left(B_{jt} = 0 \mid E_{jt} > E_c\right) \times Pr\left(E_{jt} > E_c\right) + Pr\left(B_{jt} = 0 \mid E_{jt} \leq E_c\right) \times \left[1 - Pr\left(E_{jt} > E_c\right)\right]}$$

Recall the definition of model parameters $P_1 \equiv Pr\left(B_{jt} = 1 \mid E_{jt} > E_c\right)$ and $P_0 \equiv Pr\left(B_{jt} = 1 \mid E_{jt} \leq E_c\right)$, as well as the model-implied CDF of employment $Pr\left(E_{jt} \leq E_c\right) = 1 - TXE_r^{-\bar{\beta}}$ with $\overline{\beta} \equiv \frac{\gamma_R}{\beta+\theta}$, $X \equiv R_{min}^{\gamma_R}\left(\frac{\beta}{1+\beta}\right)^{\frac{\gamma_R\beta}{\beta+\theta}}\left(\frac{E_0}{exp(\tilde{W})}\right)^{\frac{\gamma_R}{\beta+\theta}}$, and $T \equiv \kappa^{\frac{\gamma_R}{\beta+\theta}}\left[\frac{\delta}{\tau}\right]^{\frac{\beta\gamma_R}{\theta+\beta}}$. We obtain:

$$\underbrace{Pr\left(E_{jt} > E_c \mid B_{jt} = 1\right) - Pr\left(E_{jt} > E_c \mid B_{jt} = 0\right)}_{(*1)} =$$

$$\frac{P_1 \times TXE_r^{-\overline{\beta}}}{P_1 \times TXE_r^{-\overline{\beta}} + P_0 \times \left[1 - TXE_r^{-\overline{\beta}}\right]} - \frac{(1-P_1) \times TXE_r^{-\overline{\beta}}}{(1-P_1) \times TXE_r^{-\overline{\beta}} + (1-P_0) \times \left[1 - TXE_r^{-\overline{\beta}}\right]}$$

Term $(*2)$ can be expressed as:

$$\underbrace{\mathbb{E}\left[\ln\left(w_{jt}\right) \mid E_{jt} > E_c\right] - \mathbb{E}\left[\ln\left(w_{jt}\right) \mid E_{jt} \leq E_c\right]}_{(2*)} = \int_{R_r}^{R_{max}} \ln\left(w_1\left(R\right)\right) f_{R|R\geq R_r}\left(R \mid R \geq R_r\right) dR$$
$$- \int_{R_{min}}^{R_c} \ln\left(w_0^u\left(R\right)\right) f_{R|R<R_r}\left(R \mid R < R_r\right) dR$$
$$- \int_{R_c}^{R_r} \ln\left(w_0^c\left(R\right)\right) f_{R|R<R_r}\left(R \mid R < R_r\right) dR$$

After adding and substracting $\int_{R_r}^{R_{max}} w_0^u\left(R\right) f_{R|R\leq R_r}\left(R|R \leq R_r\right) dR$ and $\int_{R_c}^{R_r} w_0^u\left(R\right) f_{R|R<R_r}\left(R|R < R_r\right) dR$, we can decompose term $(*2)$ into the following three components:

$$\underbrace{\mathbb{E}\left[\ln\left(w_{jt}\right) \mid E_{jt} \geq E_r\right] - \mathbb{E}\left[\ln\left(w_{jt}\right) \mid E_{jt} \leq E_c\right]}_{(*2)} =$$

$$\underbrace{\int_{R_r}^{R_{max}} \left[\ln\left(w_1\left(R\right)\right) - \ln\left(w_0^u\left(R\right)\right)\right] f_{R|R\geq R_r}\left(R \mid R \geq R_r\right) dR}_{\text{Direct Effect of the Regulation}}$$

$$\underbrace{\int_{R_r}^{R_{max}} \ln\left(w_0^u\left(R\right)\right) f_{R|R\geq R_r}\left(R \mid R \geq R_r\right) dR - \int_{R_{min}}^{R_r} \ln\left(w_0^u\left(R\right)\right) f_{R|R<R_r}\left(R \mid R < R_r\right) dR}_{\text{Composition (i.e. firm-size premium)}}$$

$$\underbrace{\int_{R_c}^{R_r} \left[\ln\left(w_0^u\left(R\right)\right) - \ln\left(w_0^c\left(R\right)\right)\right] f_{R|R<R_r}\left(R \mid R < R_r\right) dR}_{\text{Effect of avoidance/bunching on avg wage below threshold}} \tag{B.21}$$

Using the definitions of $w_1\left(R\right) \equiv \frac{\delta}{\tau}\frac{\beta}{1+\beta}R$ and $w_0^u\left(R\right) \equiv \frac{\beta}{1+\beta}R$, we have:

$$\underbrace{\int_{R_r}^{R_{max}} \left[\ln\left(w_1\left(R\right)\right) - \ln\left(w_0^u\left(R\right)\right)\right] f_{R|R\geq R_r}\left(R \mid R \geq R_r\right) dR}_{\text{Direct Effect of the Regulation}} = \ln\left(\frac{\delta}{\tau}\right)$$

and

$$\underbrace{\int_{R_r}^{R_{max}} \ln\left(w_0^u\left(R\right)\right) f_{R|R\geq R_r}\left(R \mid R \geq R_r\right) dR - \int_{R_{min}}^{R_r} \ln\left(w_0^u\left(R\right)\right) f_{R|R<R_r}\left(R \mid R < R_r\right) dR}_{\text{Composition (i.e. firm-size premium)}}$$
$$= \mathbb{E}\left[\ln\left(R\right) \mid R \geq R_r\right] - \mathbb{E}\left[\ln\left(R\right) \mid R < R_r\right]$$

which is the difference in average log productivity between firms above and below the regulation threshold.

We can also express the effect of avoidance (i.e. bunching) on the average wage below the threshold as a function of model parameters. First, note that:

$$\begin{aligned}
\ln\left(w_0^u(R)\right) - \ln\left(w_0^c(R)\right) &= \ln\left(\frac{w_0^u(R)}{w_0^u(R_c)}\right) - \ln\left(\frac{w_0^c(R)}{w_0^u(R_c)}\right) \\
&= \ln\left(\frac{w_0^u(R)}{w_0^u(R_c)}\right) - \ln\left(\frac{w_0^c(R)}{w_0^c(R_c)}\right) \\
&= \frac{\beta+\theta}{\beta}\ln\left(\frac{R}{R_c}\right)
\end{aligned}$$

where the second line uses the definition of R_c which implies $w_0^u\left(R_c\right) = w_0^c\left(R\right)$, and the third line uses the definitions of $w_0^u(R)$ and $w_0^c(R)$. Plugging into the definition of the effect of avoidance, we have:

$$\begin{aligned}
&\underbrace{\int_{R_c}^{R_r}\left[\ln\left(w_0^u\left(R\right)\right) - \ln\left(w_0^c\left(R\right)\right)\right] f_{R|R<R_r}\left(R \mid R<R_r\right) dR}_{\text{Effect of avoidance/bunching on avg wage below threshold}} = \\
&\qquad \frac{\beta+\theta}{\beta}\int_{R_c}^{R_r}\ln\left(\frac{R}{R_c}\right) f_{R|R<R_r}\left(R \mid R<R_r\right) dR \\
&\qquad \frac{\beta+\theta}{\beta}\int_{R_c}^{R_r}\ln\left(\frac{R}{R_c}\right) \frac{f_R\left(R\right)}{F_R\left(R_r\right)} dR \\
&\qquad \frac{\beta+\theta}{\beta}\int_{R_c}^{R_r}\ln\left(\frac{R}{R_c}\right) \frac{f_R\left(R\right)}{F_R\left(R_r\right) - F_R\left(R_c\right)} dR \times \frac{F_R\left(R_r\right) - F_R\left(R_c\right)}{F_R\left(R_r\right)} \\
&\qquad \frac{\beta+\theta}{\beta}\mathbb{E}\left[\ln\left(\frac{R}{R_c}\right) \mid R_c \leq R < R_r\right] \times \frac{Pr\left(R_c \leq R < R_r\right)}{Pr\left(R < R_r\right)}
\end{aligned}$$

Collecting all previous results, we can write:

$$\mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=1\right]-\mathbb{E}\left[\ln\left(w_{jt}\right) \mid B_{jt}=0\right]=$$

$$\underbrace{\left[\frac{P_1 \times TXE_r^{-\overline{\beta}}}{P_1 \times TXE_r^{-\overline{\beta}}+P_0 \times\left[1-TXE_r^{\overline{\beta}}\right]}-\frac{\left(1-P_1\right) \times TXE_r^{-\overline{\beta}}}{\left(1-P_1\right) \times TXE_r^{-\overline{\beta}}+\left(1-P_0\right) \times\left[1-TXE_r^{\overline{\beta}}\right]}\right]}_{\text{Gap in share of large firms } Pr(E_{jt}>E_c|B_{jt}=1)-Pr(E_{jt}>E_c|B_{jt}=0)}$$

$$\times \left[\underbrace{\ln\left(\frac{\delta}{\tau}\right)}_{\text{Direct Effect of Regulation}} + \underbrace{\mathbb{E}\left[\ln(R) \mid R \geq R_r\right]-\mathbb{E}\left[\ln(R) \mid R<R_r\right]}_{\text{Composition (i.e. firm-size premium)}}\right.$$

$$\left. + \underbrace{\frac{\beta+\theta}{\beta} \times \mathbb{E}\left[\ln\left(\frac{R}{R_c}\right) \mid R_c \leq R<R_r\right] \times \frac{Pr\left(R_c \leq R<R_r\right)}{Pr\left(R<R_r\right)}}_{\text{Effect of avoidance/bunching on avg wage below size threshold } (=0 \text{ under no bunching})}\right] \tag{B.22}$$

B.3 Joint distribution of firm size and presence of a worker representative on the board

In this section, we derive the joint distribution of firm size and the presence of a worker representative on the board, for two versions of the model: the baseline model and the model with measurement error.

B.3.1 Baseline Model

Let us start from the equilibrium density of firm size defined in (3) in the main text:

$$\chi^*(E)=\begin{cases}\overline{\beta} XE^{-\left(1+\overline{\beta}\right)} & \text{if } E<E_c \\ X\left(E_c^{-\overline{\beta}}-TE_r^{-\overline{\beta}}\right) & \text{if } E=E_c \\ 0 & \text{if } E_c<E<E_r \\ \overline{\beta} TXE^{-\left(1+\overline{\beta}\right)} & \text{if } E \geq E_r\end{cases} \tag{3}$$

Recall that the (size-dependent) probability of having a worker representative on board is given by:

$$Pr(B=1 \mid E)=\begin{cases}P_0 & \text{if } E \leq E_c \\ P_1 & \text{if } E>E_c\end{cases}$$

Using the law of total probability, we can derive the joint distribution:

$$\chi^*(E, B = 1) = Pr(B = 1 \mid E) \times \chi^*(E)$$
$$= \begin{cases} P_0 \times \chi^*(E) & \text{if } E < E_c \\ P_1 \times \chi^*(E) & \text{if } E \geq E_c \end{cases}$$
$$= \begin{cases} P_0 \times \overline{\beta} X E^{\left(1+\overline{\beta}\right)} & \text{if } E < E_c \\ P_0 \times X \left(E_c^{-\overline{\beta}} - T E_r^{-\overline{\beta}}\right) & \text{if } E = E_c \\ 0 & \text{if } E_c < E < E_r \\ P_1 \times \overline{\beta} T X E^{-\left(\overline{\beta}+1\right)} & \text{if } E \geq E_r \end{cases} \tag{B.23}$$

and

$$\chi^*(E, B = 0) = Pr(B = 0 \mid E) \times \chi^*(E)$$
$$= \begin{cases} (1 - P_0) \times \overline{\beta} X E^{\left(1+\overline{\beta}\right)} & \text{if } E < E_c \\ (1 - P_0) \times X \left(E_c^{-\overline{\beta}} - T E_r^{-\overline{\beta}}\right) & \text{if } E = E_c \\ 0 & \text{if } E_c < E < E_r \\ (1 - P_1) \times \overline{\beta} T X E^{-\left(\overline{\beta}+1\right)} & \text{if } E \geq E_r \end{cases} \tag{B.24}$$

Note that, indexing observations (i.e. firms) by i, the contribution of observation i to the log likelihood can be written as:

$$\mathcal{L}_i = (1 - B_i) \times \ln\left[\chi(E_i, B_i = 0)\right] + B_i \times \ln\left[\chi(E_i, B_i = 1)\right]$$

B.3.2 Model with Measurement Error

We assume that observed employment $E = E^* \exp(\epsilon)$ where E^* is the firm's true size, and measurement error $\varepsilon \sim \mathcal{N}(0, \sigma)$. We start from the (joint) distribution of observed firm size E and presence of worker representative B conditioning on the measurement error ε:

$$
Pr\left(x<E,B=1\mid\varepsilon\right)=\begin{cases}0, & \text{if } \varepsilon\geq\ln(E)-\ln(E_{\min}),\\ P_0\left[1-X\left(E\exp(-\varepsilon)\right)^{-\overline{\beta}}\right], & \text{if } \ln(E)-\ln(E_c)<\varepsilon\leq\ln(E)-\ln(E_{\min}),\\ P_0\left[1-XTE_r^{-\overline{\beta}}\right], & \text{if } \ln(E)-\ln(E_r)<\varepsilon\leq\ln(E)-\ln(E_c),\\ P_1\left[1-XT(E\exp(-\varepsilon))^{-\overline{\beta}}\right] & \\ \quad+(P_0-P_1)\left[1-XTE_r^{-\overline{\beta}}\right], & \text{if } \varepsilon\leq\ln(E)-\ln(E_r).\end{cases}
\tag{B.25}
$$

and

$$
Pr\left(x<E,B=0\mid\varepsilon\right)=\begin{cases}0, & \text{if } \varepsilon\geq\ln(E)-\ln(E_{\min}),\\ (1-P_0)\left[1-X\left(E\exp(-\varepsilon)\right)^{-\overline{\beta}}\right], & \text{if } \ln(E)-\ln(E_c)<\varepsilon\leq\ln(E)-\ln(E_{\min}),\\ (1-P_0)\left[1-XTE_r^{-\overline{\beta}}\right], & \text{if } \ln(E)-\ln(E_r)<\varepsilon\leq\ln(E)-\ln(E_c),\\ (1-P_1)\left[1-XT(E\exp(-\varepsilon))^{-\overline{\beta}}\right] & \\ \quad+(P_1-P_0)\left[1-XTE_r^{-\overline{\beta}}\right], & \text{if } \varepsilon\leq\ln(E)-\ln(E_r).\end{cases}
\tag{B.26}
$$

Under the assumption that $\varepsilon\sim\mathcal{N}\left(0,\sigma\right)$, we can derive the unconditional (joint) probability $Pr\left(x<E,B\right)$, for $B=0,1$,

$$
Pr\left(x<E,B\right)=\int_{-\infty}^{+\infty}Pr\left(x<E,B\mid\varepsilon\right)\frac{1}{\sigma}\phi\left(\frac{\varepsilon}{\sigma}\right)d\varepsilon
\tag{B.27}
$$

In order to solve for the integral in (B.27), the following is useful:

$$
\begin{aligned}
\int_a^b XE^{-\overline{\beta}}\exp\left(\varepsilon\overline{\beta}\right)\frac{1}{\sigma}\phi\left(\frac{\varepsilon}{\sigma}\right)d\varepsilon &= XE^{-\overline{\beta}}\int_a^b\exp\left(\varepsilon\overline{\beta}\right)\frac{1}{\sigma}\phi\left(\frac{1}{\sigma}\right)d\varepsilon\\
&= XE^{-\overline{\beta}}\int_a^b\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\varepsilon}{\sigma}\right)^2+\varepsilon\overline{\beta}\right)d\varepsilon\\
&= XE^{-\overline{\beta}}\int_a^b\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\left(\frac{\varepsilon}{\sigma}\right)^2-2\varepsilon\overline{\beta}\right]\right)d\varepsilon
\end{aligned}
\tag{B.28}
$$

Also note that:

$$\left(\frac{\varepsilon}{\sigma}\right)^2 - 2\overline{\beta}\varepsilon = \left(\frac{1}{\sigma}\varepsilon - \overline{\beta}\sigma\right)^2 - \left(\overline{\beta}\sigma\right)^2$$

therefore:

$$\begin{aligned}
\int_a^b XE^{-\overline{\beta}} \exp\left(\varepsilon\overline{\beta}\right) \frac{1}{\sigma}\phi\left(\frac{\varepsilon}{\sigma}\right) d\varepsilon &= XE^{-\overline{\beta}} \int_a^b \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\left(\frac{\varepsilon}{\sigma}\right)^2 - 2\varepsilon\overline{\beta}\right]\right) d\varepsilon \\
&= XE^{-\overline{\beta}} \exp\left(\frac{1}{2}\left(\overline{\beta}\sigma\right)^2\right) \int_a^b \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\frac{\varepsilon}{\sigma} - \overline{\beta}\sigma\right]^2\right) d\varepsilon \\
&= XE^{-\overline{\beta}} \exp\left(\frac{1}{2}\left(\overline{\beta}\sigma\right)^2\right) \int_a^b \underbrace{\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\frac{\varepsilon - \left(\overline{\beta}\sigma^2\right)}{\sigma}\right]^2\right)}_{\text{Normal density with mean } \overline{\beta}\sigma^2 \text{ and variance } \sigma^2} d\varepsilon \\
&= XE^{-\overline{\beta}} \exp\left(\frac{1}{2}\left(\overline{\beta}\sigma\right)^2\right) \left[\Phi\left(\frac{b - \overline{\beta}\sigma^2}{\sigma}\right) - \Phi\left(\frac{a - \overline{\beta}\sigma^2}{\sigma}\right)\right]
\end{aligned} \tag{B.29}$$

Using (B.28) and (B.29), we can solve the integral in (B.27):

$$\begin{aligned}
Pr\left(x < E, B = 1\right) = \int_{-\infty}^{+\infty} & Pr\left(x < E, B = 1 \mid \varepsilon\right) \frac{1}{\sigma}\phi\left(\frac{\varepsilon}{\sigma}\right) d\varepsilon \\
= & P_0\Phi\left(\frac{\ln E - \ln E_{\min}}{\sigma}\right) \\
& - P_0 XE^{-\overline{\beta}} \exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right) \\
& \times \left[\Phi\left(\frac{\ln E - \ln E_{\min} - \overline{\beta}\sigma^2}{\sigma}\right) - \Phi\left(\frac{\ln E - \ln E_c - \overline{\beta}\sigma^2}{\sigma}\right)\right] \\
& - P_1 XTE^{-\overline{\beta}} \exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right) \Phi\left(\frac{\ln E - \ln E_r - \overline{\beta}\sigma^2}{\sigma}\right) \\
& + XTE_r^{-\overline{\beta}}\left[P_1\Phi\left(\frac{\ln E - \ln E_r}{\sigma}\right) - P_0\Phi\left(\frac{\ln E - \ln E_c}{\sigma}\right)\right]
\end{aligned} \tag{B.30}$$

and

$$
\begin{aligned}
Pr\left(x<E,B=0\right) &= \int_{-\infty}^{+\infty} Pr\left(x<E,B=0 \mid \varepsilon\right)\frac{1}{\sigma}\phi\left(\frac{\varepsilon}{\sigma}\right)d\varepsilon \\
&= (1-P_0)\Phi\left(\frac{\ln E-\ln E_{\min}}{\sigma}\right) \\
&\quad -(1-P_0)XE^{-\overline{\beta}}\exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right) \\
&\quad \times\left[\Phi\left(\frac{\ln E-\ln E_{\min}-\overline{\beta}\sigma^2}{\sigma}\right)-\Phi\left(\frac{\ln E-\ln E_c-\overline{\beta}\sigma^2}{\sigma}\right)\right] \\
&\quad -(1-P_1)XTE^{-\overline{\beta}}\exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right)\Phi\left(\frac{\ln E-\ln E_r-\overline{\beta}\sigma^2}{\sigma}\right) \\
&\quad +XTE_r^{-\bar{\beta}}\left[(1-P_1)\Phi\left(\frac{\ln E-\ln E_r}{\sigma}\right)-(1-P_0)\Phi\left(\frac{\ln E-\ln E_c}{\sigma}\right)\right]
\end{aligned}
$$

To obtain the density, we take the derivative of (B.30) with respect to E and obtain:

$$
\begin{aligned}
\chi\left(E,B=1\right) &= P_0\phi\left(\frac{\ln E-\ln E_{\min}}{\sigma}\right)\frac{1}{\sigma E} \\
&\quad +P_0\overline{\beta}XE^{-\left(\overline{\beta}+1\right)}\exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right) \\
&\quad \times\left[\Phi\left(\frac{\ln E-\ln E_{\min}-\overline{\beta}\sigma^2}{\sigma}\right)-\Phi\left(\frac{\ln E-\ln E_c-\overline{\beta}\sigma^2}{\sigma}\right)\right] \\
&\quad -P_0XE^{-\overline{\beta}}\exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right) \\
&\quad \times\left[\phi\left(\frac{\ln E-\ln E_{\min}-\overline{\beta}\sigma^2}{\sigma}\right)-\phi\left(\frac{\ln E-\ln E_c-\overline{\beta}\sigma^2}{\sigma}\right)\right]\frac{1}{\sigma E} \\
&\quad +P_1\overline{\beta}XTE^{-\left(\overline{\beta}+1\right)}\exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right)\Phi\left(\frac{\ln E-\ln E_r-\overline{\beta}\sigma^2}{\sigma}\right) \\
&\quad -P_1XTE^{-\overline{\beta}}\exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right) \\
&\quad \times\phi\left(\frac{\ln E-\ln E_r-\overline{\beta}\sigma^2}{\sigma}\right)\frac{1}{\sigma E} \\
&\quad +XTE_r^{\bar{\beta}}\left[P_1\phi\left(\frac{\ln E-\ln E_r}{\sigma}\right)-P_0\phi\left(\frac{\ln E-\ln E_c}{\sigma}\right)\right]\frac{1}{\sigma E}.
\end{aligned}
$$

We use the fact that

$$\exp\left(\frac{1}{2}\left(\overline{\beta}\sigma\right)^2\right)E^{-\overline{\beta}}\phi\left(\frac{\ln(E)-\ln(Z)-\overline{\beta}\sigma^2}{\sigma}\right)=Z^{-\overline{\beta}}\phi\left(\frac{\ln(E)-\ln(Z)}{\sigma}\right)$$

to simplify (and re-arrange) the density. We obtain:

$$\begin{aligned}
\chi(E,B=1)=P_0&\left(1-XE_{\min}^{-\overline{\beta}}\right)\frac{1}{\sigma E}\phi\left(\frac{\ln E-\ln E_{\min}}{\sigma}\right)\\
&+P_0\left(XE_c^{-\overline{\beta}}-XTE_r^{-\overline{\beta}}\right)\frac{1}{\sigma E}\phi\left(\frac{\ln E-\ln E_c}{\sigma}\right)\\
&+P_0\overline{\beta}XE^{-(\overline{\beta}+1)}\exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right)\\
&\times\left[\Phi\left(\frac{\ln E-\ln E_{\min}-\overline{\beta}\sigma^2}{\sigma}\right)-\Phi\left(\frac{\ln E-\ln E_c-\overline{\beta}\sigma^2}{\sigma}\right)\right]\\
&+P_1\overline{\beta}XTE^{-(\overline{\beta}+1)}\exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right)\Phi\left(\frac{\ln E-\ln E_r-\overline{\beta}\sigma^2}{\sigma}\right).
\end{aligned} \tag{B.31}$$

Similarly:

$$\begin{aligned}
\chi(E,B=0)=(1-P_0)&\left(1-XE_{\min}^{-\overline{\beta}}\right)\frac{1}{\sigma E}\phi\left(\frac{\ln E-\ln E_{\min}}{\sigma}\right)\\
&+(1-P_0)\left(XE_c^{-\overline{\beta}}-XTE_r^{-\overline{\beta}}\right)\frac{1}{\sigma E}\phi\left(\frac{\ln E-\ln E_c}{\sigma}\right)\\
&+(1-P_0)\overline{\beta}XE^{-(\overline{\beta}+1)}\exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right)\\
&\times\left[\Phi\left(\frac{\ln E-\ln E_{\min}-\overline{\beta}\sigma^2}{\sigma}\right)-\Phi\left(\frac{\ln E-\ln E_c-\overline{\beta}\sigma^2}{\sigma}\right)\right]\\
&+(1-P_1)\overline{\beta}XTE^{-(\overline{\beta}+1)}\exp\left(\frac{1}{2}(\overline{\beta}\sigma)^2\right)\Phi\left(\frac{\ln E-\ln E_r-\overline{\beta}\sigma^2}{\sigma}\right).
\end{aligned} \tag{B.32}$$

Indexing observations (i.e. firms) by i, the contribution of observation i to the log likelihood can be written as:

$$\mathcal{L}_i=(1-B_i)\times\ln\left[\chi(E_i,B_i=0)\right]+B_i\times\ln\left[\chi(E_i,B_i=1)\right]$$

B.4 Normalization of X in the Firm Size Distribution

Because of model restrictions, X is not a free parameter; it is determined as a function of other parameters in the firm size distribution. To see this, consider the equilibrium density of firm size in equation (3). By definition, the density must integrate to one:

$$
\begin{aligned}
1 &= \int_{E_{\min}}^{\infty} \chi^*(E)\, dE \\
&= \int_{E_{\min}}^{E_c} \bar{\beta} X E^{-(\bar{\beta}+1)}\, dE + X\left(E_c^{-\bar{\beta}} - T E_r^{-\bar{\beta}}\right) + \int_{E_r}^{\infty} \bar{\beta} T X E^{-(\bar{\beta}+1)}\, dE \\
&= X E_{\min}^{-\bar{\beta}}.
\end{aligned}
$$

This implies that:

$$X = E_{\min}^{\bar{\beta}}.$$

The same normalization, $X = E_{\min}^{\bar{\beta}}$, also holds in extensions of the model that allow for measurement error and partial compliance.

C Re-considering threat point effects

In Section C.1, we present the results from an event study analysis of firm and worker outcomes around the election of the first worker representative on the corporate board. Comparing outcomes of firms which elect a worker representative compared with firms which adopt the policy in later years, we find no significant change in wages or value added per worker. These results tend to provide support for our assumption that the direct effects of the regulation apply to all firms above the threshold, independent of actual adoption.

Nevertheless, we consider an alternative version of the model in which the direct effects only apply to firms that actually adopt worker representation. In Section C.2, we formally present this *partial exposure* model and derive the distribution of firm size as a function of the parameters. In Section C.3, we estimate the partial exposure model and present the corresponding estimated aggregate effects in Table C.3.

C.1 Event study of adoption of worker representation on the corporate board

We restrict the sample to workers employed by a firm adopting worker representation for the first time between 2006 and 2012.[11] By limiting the time period to these years, we can require all workers to be employed by the adopting firms at least two years before and at

[11] We assume that worker representation is an absorbing state, since we observe only a small number of firms choosing to discontinue worker representation after once having adopted.

Table C.2: Characteristics of the adoption sample

	Full sample	Adoption sample
Worker representative on the corporate board	0.38	0.00
Firm size	1,027.94	309.90
Hourly wage (NOK)	241.75	223.92
Value added pr. worker (NOK)	1,289.69	995.01
Hiring rate	0.21	0.34
Separation rate	0.19	0.17
Firm age (years)	16.77	14.59
Nr of workers	1,561,750	19,239
Nr of firms	127,748	789

Notes: All variables are defined in Appendix Table A.1. The characteristics of the adoption sample are measured two years before the firm elects a worker representative for the first time.

least three years after the year in which a worker representative is elected. This ensures that changes in the composition of workers within adopting firms do not affect our result.

Table C.2 compares the observable characteristics of firms in this *adoption sample*, to the full sample. Firms which adopt worker representation are on average smaller, younger, and growing in the years before adoption. The growth in the number of employees is driven by a higher rate of hiring compared with the full sample.

We compare changes in wages for workers employed in firms adopting worker representation in a given year to changes in wages for workers employed in firms adopting worker representation in later years. More concretely, for any year $c \in \{2006, 2012\}$, let the treatment group consist of all workers employed in firms adopting worker representation in that year. The average change in outcomes between time $c + s$ and the baseline year $c - 2$ in the treatment group is denoted by:

$$\mathbb{E}[Y_{i,c+s} - Y_{i,c-2} \mid C_{j(i)} = c],$$

where $C_{j(i)}$ denotes the year in which the employer j of worker i adopted worker representation. The control group is defined for each cohort c and event time s and consists of all workers in the adoption subsample (as defined in Section 3.2) who are employed in firms that have not adopted worker representation by year $\max(c+s, c)$, but that adopt worker representation in a later year within our sample period. The event study estimator between the treatment and control groups is defined as follows (for $s \geq 0$):

$$\mathbb{E}[Y_{i,c+s} - Y_{i,c-2} \mid C_{j(i)} = c] - \mathbb{E}[Y_{i,c+s} - Y_{i,c-2} \mid C_{j(i)} > c+s]. \tag{C.33}$$

The event study estimator eliminates unobserved time-invariant individual heterogeneity by comparing workers in firms adopting worker representation before and after the year of adoption while accounting for year and event-time effects by using workers in firms adopting worker representation in a later year as a control group before worker representation is adopted in their firms. As long as the workers in the treatment and control groups would have had a common trend in wages between years $(c-2)$ and $(c+s)$ in the absence of adoption of worker representation, the event study estimator in equation (C.33) recovers the average impact of adopting worker representation for cohort c in year $c+s$ for $s \geq 0$.

To implement the event study design, we use the following regression model. For each cohort c and each event time s, we create a subsample consisting of workers in firms adopting worker representation in year c (the treatment group) and a control group of workers in firms that have not adopted worker representation by period $\max(c, c+s)$. Using this subsample, we run the regression:

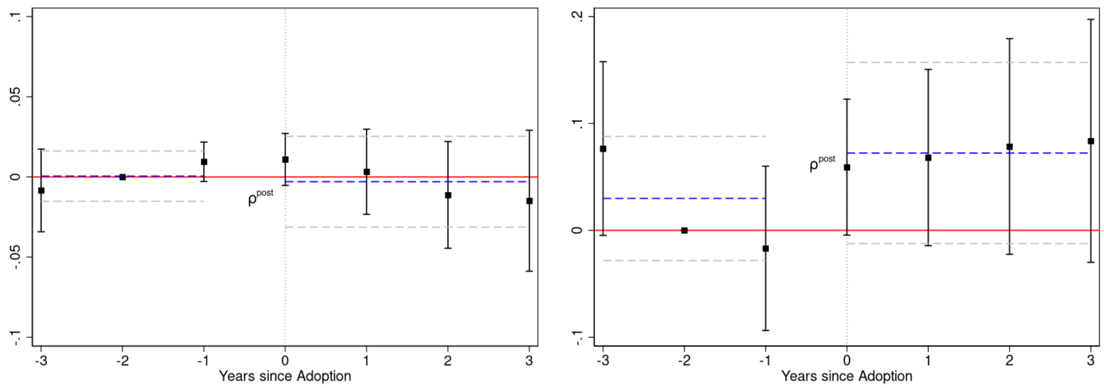
$$Y_{i,t} = \alpha_1^{c,s} + \alpha_2^{c,s}\mathbf{1}[C_{j(i)} = c] + \pi^{c,s}\mathbf{1}[t = c+s] + \rho^{c,s}\mathbf{1}[C_{j(i)} = c, t = c+s] + \epsilon_{i,t}, \tag{C.34}$$

where $\alpha_1^{c,s}$ is the control group mean in the baseline year (i.e., $c-2$), $\alpha_2^{c,s}$ is a fixed effect for the treated workers (cohort c), $\pi^{c,s}$ is a time effect for event time s, and $\rho^{c,s}$ is an interaction effect. Our parameter of interest $\rho^{c,s}$ measures differences in trends in outcomes between years $c+s$ and $c-2$ between the treatment and control groups. For $s \geq 0$, this parameter captures the average impact of adopting worker representation for cohort c under the common trends assumption.

In Figure C.1, we plot the estimated coefficients from equation (C.34) for hourly wages and value added per worker (in logs). For each event time $s \in \{c-3, \ldots, c+3\}$, we report an equally weighted average of the cohort-specific coefficients $\rho^{c,s}$, with the baseline event time

Figure C.1: Firm outcomes after adoption of worker representation

(a) (log) Hourly wages of stayers (b) (log) Value added per worker



Notes: These figures plot the estimated coefficients from equation (C.34) for two different outcomes. For each event time $s \in \{c-3, \dots, c+3\}$, we report an equally weighted average of the cohort-specific coefficients $\rho^{c,s}$, with the baseline event time $c-2$ normalized to zero. Vertical lines represent 95% confidence intervals constructed using standard errors clustered at the firm-level. The dashed horizontal lines indicate the average treatment effects in the pre- and post-treatment periods, where all time-periods and cohorts are weighted equally. The sample used is the adoption sample as described in Table C.2.

$c-2$ normalized to zero.[12] In the post-period, we find no significant change in wages nor value added per worker. These results are consistent with our assumption that the effects of the size-dependent regulation - if any - do not operate through the actual implementation of worker representation.

C.2 Partial Exposure Model

We assume that firms first choose their size E. Then, they have a worker representative on the board with probability

$$Pr(B = 1 \mid E) = \begin{cases} P_0 & \text{if } E \leq E_c \\ P_1 & \text{if } E > E_c \end{cases}$$

[12] By first estimating the parameter $\rho^{c,s}$ separately for each cohort c and then averaging these parameters across cohorts, we avoid the problems pointed out by Callaway and Sant'Anna (2021) and Sun and Abraham (2021), and we ensure that our event study regressions produce positively weighted averages of causal effects under the standard common trends assumption.

Firms with a board representative have productivity δR, labor cost $\tau wE + F$, and non-wage amenities κR^{θ} (regardless of whether they choose employment above or below E_c). Firms without a board representative have productivity R, labor cost wE, and non-wage amenities R^{θ}.

After the uncertainty about B is realized, the firm then pays the required wage to reach size E. The required wage depends on B via the effect of the worker representative on non-wage amenities (κ):

$$w(E,R,B) = \begin{cases} \left(\frac{E}{E_0} \frac{\exp\left(\tilde{W}\right)}{R^{\theta}} \right)^{\frac{1}{\beta}} & \text{if } B = 0 \\ \left(\frac{E}{E_0} \frac{\exp\left(\tilde{W}\right)}{\kappa R^{\theta}} \right)^{\frac{1}{\beta}} & \text{if } B = 1 \end{cases}$$

As a result, the firm maximizes expected profit:

$$\pi(R) \equiv \max_{E} \begin{Bmatrix} (1-P_0)\left[RE - w(E,R,0)E\right] + P_0\left[\delta RE - \tau w(E,R,1)E - F\right] & \text{if } E \leq E_c \\ (1-P_1)\left[RE - w(E,R,0)E\right] + P_1\left[\delta RE - \tau w(E,R,1)E - F\right] & \text{if } E > E_c \end{Bmatrix}$$

Solving the maximization problem gives us the following interior solutions:

$$E(R) = \left[\frac{1+P_i(\delta-1)}{1+P_i(\kappa^{-\frac{1}{\beta}}\tau - 1)} \right]^{\beta} \left(\frac{\beta}{1+\beta} \right)^{\beta} \frac{E_0}{\exp\left(\tilde{W}\right)} R^{\beta+\theta}$$

$$\pi(R) = \frac{1+P_i(\delta-1)}{1+\beta} \left[\frac{1+P_i(\delta-1)}{1+P_i\left(\kappa^{-\frac{1}{\beta}}\tau - 1\right)} \right]^{\beta} \left(\frac{\beta}{1+\beta} \right)^{\beta} \left(\frac{E_0}{\exp\left(\tilde{W}\right)} \right) R^{\beta+\theta+1} - P_i F$$

and

$$w(R,B) = \begin{cases} \frac{1+P_i(\delta-1)}{1+P_i(\kappa^{-\frac{1}{\beta}}\tau-1)} \left(\frac{\beta}{1+\beta} \right) \kappa^{-\frac{1}{\beta}} R & \text{if } B = 1 \\ \frac{1+P_i(\delta-1)}{1+P_i(\kappa^{-\frac{1}{\beta}}\tau-1)} \left(\frac{\beta}{1+\beta} \right) R & \text{if } B = 0 \end{cases}$$

with $i = 0$ if $E \leq E_c$, and $i = 1$ if $E > E_c$.

Firms which decide to bunch at $E = E_c$ must pay the wage:

$$w_c(R,B) = \begin{cases} \left[\frac{E_c}{E_0}\exp\left(\tilde{W}\right)\right]^{\frac{1}{\beta}}\kappa^{-\frac{1}{\beta}}R^{-\frac{\theta}{\beta}} & \text{if } B=1 \\ \left[\frac{E_c}{E_0}\exp\left(\tilde{W}\right)\right]^{\frac{1}{\beta}}R^{-\frac{\theta}{\beta}} & \text{if } B=0 \end{cases}$$

As a result, the expected profit for bunching firms is:

$$\pi_c(R) = \left\{[1+P_0(\delta-1)]\,R - \left[1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)\right]\left[\frac{E_c}{E_0}\exp\left(\tilde{W}\right)\right]^{\frac{1}{\beta}}R^{-\frac{\theta}{\beta}}\right\}E_c - P_0F$$

Assuming the cost of the policy to the marginal firm is positive (i.e. assuming $\pi_0(R_c) \geq \pi_1(R_c)$), there are three types of firms: unregulated firms with productivity $R \leq R_c$, bunching firms with $R_c < R \leq R_r$, and regulated firms with $R > R_r$. Given the distribution of firm productivity $F_R(.)$, we can derive the corresponding equilibrium distribution of firm size, which follows a "broken power law":

$$Pr\left(E_{jt} < E\right) = \begin{cases} 1 - XE^{-\bar{\beta}}, & E < E_c \\ 1 - XTE_r^{-\bar{\beta}}, & E_c \leq E < E_r \\ 1 - XTE^{-\bar{\beta}}, & E \geq E_r \end{cases}$$

where we re-define:

$$\bar{\beta} = \frac{\gamma_R}{\beta+\theta}$$

$$X = R_{min}^{\gamma_R}\left[\frac{1+P_0(\delta-1)}{1+P_0(\kappa^{-\frac{1}{\beta}}\tau-1)}\right]^{\frac{\beta\gamma_R}{\beta+\theta}}\left(\frac{\beta}{1+\beta}\right)^{\frac{\beta\gamma_R}{\beta+\theta}}\left(\frac{E_0}{\exp(\tilde{w})}\right)^{\frac{\gamma_R}{\beta+\theta}}$$

$$T = \left[\frac{1+P_1(\delta-1)}{1+P_1(\kappa^{-\frac{1}{\beta}}\tau-1)}\right]^{\frac{\beta\gamma_R}{\beta+\theta}}\left[\frac{1+P_0(\delta-1)}{1+P_0(\kappa^{-\frac{1}{\beta}}\tau-1)}\right]^{-\frac{\beta\gamma_R}{\beta+\theta}}$$

The "broken power law" distribution of firm size is of the same parametric form as the full exposure model in the main text but the mapping between $\overline{\beta}$, X, T, and the model parameters differ. We can add measurement error in firm size to the model and derive the distribution of observed employment using the same derivations as in Section B.3.2.

C.3 Results

We estimate the partial exposure model and present the results in Table C.3 below. Results are virtually identical to those under full exposure.

Table C.3: Aggregate effects of the policy under partial compliance

		Full exposure		Partial exposure	
		(1)	(2)	(3)	(4)
(A) Firm effects		Baseline	Measurement error	Baseline	Measurement error
Aggregate annual firm profits					
	Direct policy effect	-0.66%	1.53%	-0.64%	1.79%
		(1.19%)	(0.85%)	(1.24%)	(0.86%)
	+ Size distortion effect	0.00%	0.00%	0.00%	0.00%
		(0.00%)	(0.00%)	(0.02%)	(0.00%)
	+ Equilibrium effects	0.15%	0.44%	0.15%	0.44%
		(0.06%)	(0.23%)	(0.27%)	(0.23%)
= Total effect		-0.50%	1.97%	-0.49%	2.23%
		(1.20%)	(0.84%)	(1.35%)	(0.82%)
(B) Worker effects					
Wage compensation					
	Direct policy effect	-0.03%	-0.20%	-0.03%	-0.09%
		(0.55%)	(0.39%)	(0.45%)	(0.13%)
	+ Size distortion effect	0.00%	0.00%	0.00%	0.00%
		(0.00%)	(0.00%)	(0.00%)	(0.00%)
	= Total effect	-0.03%	-0.20%	-0.03%	-0.09%
		(0.55%)	(0.39%)	(0.45%)	(0.13%)
Total compensation		-0.04%	-0.11%	-0.04%	-0.11%
		(0.02%)	(0.06%)	(0.07%)	(0.06%)

Notes: This table presents the model-implied aggregate effects of the policy. Columns (1) and (2) repeat the results from the baseline version of the model presented in Table 4. Columns (3) and (4) report the model-implied aggregate effects from an extension of the model where the policy only affects firms which have a worker representative. Standard errors are constructed using 100 block bootstrap samples where each block is a firm. All numbers are in percent.

D Robustness: varying the labor supply elasticity

Table D.4: Robustness of model parameter estimates to the labor supply elasticity

	$\beta = 3$		$\beta = 4$		$\beta = 5$	
Parameter	Baseline	Measurement error	Baseline	Measurement error	Baseline	Measurement error
(A) Shape parameters						
γ_r	13.162	25.134	13.169	25.133	13.174	25.133
	(0.392)	(2.113)	(0.391)	(2.113)	(0.391)	(2.113)
$\frac{\theta}{\gamma_r^2}$	0.055	0.025	0.049	0.023	0.044	0.021
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
(B) Parameters specific to the size-dependent policy						
$\delta - 1$	-0.007	0.029	-0.007	0.029	-0.007	0.029
	(0.017)	(0.013)	(0.017)	(0.013)	(0.017)	(0.013)
$\kappa - 1$	0.000	0.002	0.000	0.005	0.000	0.008
	(0.019)	(0.017)	(0.026)	(0.023)	(0.032)	(0.029)
$\tau - 1$	-0.007	0.032	-0.007	0.032	-0.007	0.032
	(0.014)	(0.011)	(0.014)	(0.011)	(0.014)	(0.011)
F	-3.29	8.15	-2.47	6.11	-1.97	4.89
	(5.95)	(4.56)	(4.46)	(3.42)	(3.57)	(2.73)

Notes*:* This table presents the estimated parameter values, taking various values of the elasticity of residual labor supply (β) as given. The parameters ($\gamma_R, \theta, \delta, \kappa, \tau$) are estimated jointly using GMM. The fixed cost is backed out as described in the text. P_0, P_1 are estimated using MLE. Standard errors are constructed using 100 replications of block bootstrap where each block is a firm.

Table D.5: Aggregate effects of the policy under various labor supply elasticities

		$\beta = 3$		$\beta = 4$		$\beta = 5$	
		(1)	(2)	(3)	(4)	(5)	(6)
(A) Firm effects		Baseline	Measurement error	Baseline	Measurement error	Baseline	Measurement error
Aggregate annual firm profits							
	Direct policy effect	-0.66%	1.53%	-0.66%	1.53%	-0.66%	1.53%
		(1.19%)	(0.85%)	(1.19%)	(0.85%)	(1.19%)	(0.85%)
	+ Size distortion effect	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
		(0.00%)	(0.00%)	(0.00%)	(0.00%)	(0.00%)	(0.00%)
	+ Equilibrium effects	0.15%	0.44%	0.15%	0.44%	0.15%	0.44%
		(0.06%)	(0.23%)	(0.06%)	(0.23%)	(0.06%)	(0.23%)
= Total effect		-0.50%	1.97%	-0.50%	1.97%	-0.50%	1.97%
		(1.20%)	(0.84%)	(1.20%)	(0.84%)	(1.20%)	(0.84%)
(B) Worker effects							
Wage compensation							
	Direct policy effect	-0.05%	-0.20%	-0.03%	-0.20%	-0.03%	-0.20%
		(0.55%)	(0.39%)	(0.55%)	(0.39%)	(0.55%)	(0.39%)
	+ Size distortion effect	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
		(0.00%)	(0.00%)	(0.00%)	(0.00%)	(0.00%)	(0.00%)
	= Total effect	-0.05%	-0.20%	-0.03%	-0.20%	-0.03%	-0.20%
		(0.55%)	(0.39%)	(0.55%)	(0.39%)	(0.55%)	(0.39%)
Total compensation		-0.05%	-0.15%	-0.04%	-0.11%	-0.03%	-0.09%
		(0.02%)	(0.08%)	(0.02%)	(0.06%)	(0.01%)	(0.05%)

Notes: This table presents the model-implied aggregate effects of the policy. Columns (1) and (2) present the results from the baseline version of the model presented in Table 4 with $\beta = 3$. Columns (2) and (3) report the results from estimating the model with $\beta = 4$. Columns (4) and (5) report the results from estimating the model with $\beta = 5$. Standard errors are constructed using 100 block bootstrap samples where each block is a firm. All numbers are in percent.

E Estimation appendix

E.1 Estimating the features of the firm size distribution using MLE

We estimate the following moments of the firm size distribution described in Section 5.1:

- $\bar{\beta}$: The Pareto tail index of the firm size distribution;
- T: The shift in density above the threshold;
- E_r: The minimum firm size above the regulatory threshold;
- P_0: The share of firms *below* the threshold that have a worker board representative;
- P_1: The share of firms *above* the threshold that have a worker board representative; and
- σ^2: The variance of measurement error (for the measurement-error specification only).

Let the number of firms in our dataset be N. Let $\mu = \left[\bar{\beta}, T, E_r, P_0, P_1, \sigma^2\right]$. We estimate μ, using a maximum likelihood procedure,

$$\hat{\mu} = \underset{\mu}{\operatorname{argmax}} \sum_{i=1}^{N} \log\left(\mathcal{L}_i\right), \tag{E.35}$$

where the likelihood function for firm i is given by $\mathcal{L}_i = \chi^*\left(E_i, B_i = 0\right)^{(1-B_i)} \times \chi^*\left(E_i, B_i = 1\right)^{B_i}$. The likelihood function is based on the joint probability distribution function for firm size and a board representative indicator, given by equations (B.23), and (B.24) for the model without measurement error, and by equations (B.31) and (B.32) for the model with measurement error.

E.1.1 Accounting for a truncated firm size distribution

The MLE sample (defined in Section 3.2) is restricted to firms with a minimum size of 10 employees and a maximum size of 50. We explicitly accounting for this *censoring* in the Maximum Likelihood. Formally, observed employment is $E \in [B_L, B_U]$, where B_L is the minimum observed firm size and B_U is the maximum. We adjust the firm size PDF:

$$\begin{aligned}
\mathbb{P}\left(E \mid E \in [B_L, B_U]\right) &= \frac{\chi^*(E)}{\mathbb{P}\left(E \in [B_L, B_U]\right)} \\
&= \frac{\chi^*(E)}{\mathbb{P}\left(E < B_U\right) - \mathbb{P}\left(E < B_L\right)}
\end{aligned}$$

where $\mathbb{P}(E < x)$ denotes the CDF of the firm size distribution, while $\chi^*(E)$ denotes the PDF. We apply this transformation for all models.

E.2 Discretizing the firm size distribution

In our model, we assume that firm size, E, is continuous; however, firm size in the data takes only integer values. For the MLE estimation, instead of relying on the continuous density $\chi^*(E)$ of firm size (and associated cumulative density $\mathbb{P}(.)$), we "discretize" it. Let x be an (integer) value of firm size observed in the data. We calculate the discretized probability of firm size x, $\mathbb{P}^{discrete}(.)$ as:

$$\mathbb{P}^{discrete}\left(E = x \mid E \in [B_L, B_U]\right) = \frac{\mathbb{P}(E < x + 0.5) - \mathbb{P}(E < x - 0.5)}{\mathbb{P}\left(E < B_U + 0.5\right) - \mathbb{P}\left(E < B_L - 0.5\right)}$$

where $\mathbb{P}$ is the (continuous) probability density function derived from the model.

E.3 Estimating model parameters using GMM

We estimate the model parameters $\Theta \equiv (\gamma_R, \theta, \delta, \kappa, \tau, E_r, P_0, P_1)$ using GMM.

Let $\left[\hat{\overline{\beta}}, \hat{T}, \hat{E}_r, \hat{P}_0, \hat{P}_1, \hat{\rho}, \hat{\xi}\ \hat{\mathcal{M}}\right]$ be the empirical moments used for estimation. Recall that $\hat{\overline{\beta}}, \hat{T}, \hat{E}_r, \hat{P}_0, \hat{P}_1$ are recovered by MLE. Let $\hat{\xi}$ denote the estimated discontinuity in wages at the threshold, $\hat{\rho}$ be the estimated discontinuity in value-added at the threshold, and $\hat{\mathcal{M}}$ be the mover estimate of the worker representation premium. Each of these moments can be expressed as a function of the model parameters (e.g. $\overline{\beta} = \gamma_R/(\beta+\theta)$). The GMM estimator is given by:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}}\, g\left(\Theta\right)' \hat{W} g\left(\Theta\right) \tag{E.36}$$

where $g(\Theta) = \left[\bar{\beta} - \hat{\bar{\beta}},\ T - \hat{T},\ \rho - \hat{\rho},\ \xi - \hat{\xi},\ \mathcal{M} - \hat{\mathcal{M}}, E_r - \hat{E}_r, P_0 - \hat{P}_0, \hat{P}_1 - P_1\right]'$ and $\hat{W}$ is a weighting matrix. Since the model is just-identified, the estimates do not depend on the choice of weighting matrix. In practice, we weigh by the inverse variance of each moment.

E.3.1 Moment mappings without measurement error

In the next two sections, we specify how each moment maps to the model. For the of sake generality, we present expressions that nest both the full- and partial exposure models. To understand the expression for each mapping under both versions of the model, note the following:

- In terms of expressing the firm profits, employment and wages, the full exposure model can be thought of as a special case of the partial exposure model with $P_0 = 0$ and $P_1 = 1$.
- However, when expressing the worker representation pay premium $\mathcal{M}$ as a function of model parameters under full exposure, it still depends on the (size-dependent) likelihood of having a board representative because it compares average wages conditional on B (see e.g. (B.22)). In order to write a general expression for the model-implied moments valid under both partial and full exposure, we need to introduce the following distinction. We let P_0^M and P_1^M denote the actual (size-dependent) probabilities of having a worker representative (which is estimated via MLE). We use P_0 and P_1

to denote the probability, resp. below and above the threshold, that a firm is directly affected by the policy.

- Under partial exposure, $P_0^M = P_0$ and $P_1^M = P_1$, which both correspond to their MLE estimate.
- Under full exposure, P_0^M and P_1^M correspond to their MLE estimate of the size-dependent probability of worker representation, while $P_0 = 0$ and $P_1 = 1$.

To simplify notation, we denote variables X normalized by its minimum value by $\tilde{X} \equiv X/X_{min}$. From the model with no measurement error, we derive the following mappings between the moments and the model parameters:

The Pareto tail index of the firm size distribution:

$$\bar{\beta} = \frac{\gamma_R}{\beta + \theta} \tag{E.37}$$

The bunching shifter:

$$T = \left[\frac{1 + P_1(\delta - 1)}{1 + P_1\left(\kappa^{-\frac{1}{\beta}}\tau - 1\right)}\right]^{\frac{\beta\gamma_R}{\beta+\theta}} \left[\frac{1 + P_0(\delta - 1)}{1 + P_0\left(\kappa^{-\frac{1}{\beta}}\tau - 1\right)}\right]^{-\frac{\beta\gamma_R}{\beta+\theta}} \tag{E.38}$$

The effect of the regulation on the marginal firm's value-added (reduced-form RD estimand): [13]

$$\begin{aligned} \rho &= \mathbb{E}\left[\log Y_{jt} \mid E_{jt} = E_r\right] - \mathbb{E}\left[\log Y_{jt} \mid E_{jt} = E_c\right] \\ &= \mathbb{E}\left[\log\left(\tilde{Y}_{jt}\right) \mid E_{jt} = E_r\right] - \mathbb{E}\left[\log\left(\tilde{Y}_{jt}\right) \mid E_{jt} = E_c\right], \end{aligned} \tag{E.39}$$

where:

[13]For estimation, we make the following substitutions wherever applicable:
$\tilde{R}_c = \left(\tilde{E}_c\right)^{\frac{1}{\beta+\theta}}$, and $\tilde{R}_r = \left(\tilde{E}_r\right)^{\frac{1}{\beta+\theta}} \left[\frac{1+P_1(\delta-1)}{1+P_1\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{-\frac{\beta}{\beta+\theta}} \left[\frac{1+P_0(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{\frac{\beta}{\beta+\theta}}$

$$\mathbb{E}\left[\log\left(\tilde{Y}_{jt}\right) \mid E_{jt} = E_r\right] = \log\left(\tilde{E}_r\right) + \log\left(\tilde{R}_r\right) + (P_1 - P_0)\log\delta,$$

$$\mathbb{E}\left[\log\left(\tilde{Y}_{jt}\right) \mid E_{jt} = E_c\right] = \int_{\tilde{R}_c}^{\tilde{R}_r} \frac{\left[\log\left(\tilde{R}\right) + \log\left(\tilde{E}_c\right)\right]\left(\tilde{R}\right)}{\left(\tilde{R}_c\right)^{-\gamma_R} - \left(\tilde{R}_r\right)^{-\gamma_R}} dF_{\tilde{R}}(\tilde{R})$$

$$= \frac{\left(\tilde{R}_c\right)^{-\gamma_R}\left[\log\left(\tilde{R}_c\right) + \frac{1}{\gamma_R}\right] - \left(\tilde{R}_r\right)^{-\gamma_R}\left[\log\left(\tilde{R}_r\right) + \frac{1}{\gamma_R}\right]}{\left(\tilde{R}_c\right)^{-\gamma_R} - \left(\tilde{R}_r\right)^{-\gamma_R}}$$

$$+ \log\left(\tilde{E}_c\right)$$

The effect of the regulation on the marginal firm's wages (reduced-form RD estimand):

$$\begin{aligned}\xi &= \mathbb{E}\left[\log w_{jt} \mid E_{jt} = E_r\right] - \mathbb{E}\left[\log w_{jt} \mid E_{jt} = E_c\right] \\ &= \mathbb{E}\left[\log\left(\tilde{w}_{jt}\right) \mid E_{jt} = E_r\right] - \mathbb{E}\left[\log\left(\tilde{w}_{jt}\right) \mid E_{jt} = E_c\right],\end{aligned} \tag{E.40}$$

where:

$$\mathbb{E}\left[\log\left(\tilde{w}_{jt}\right) \mid E_{jt} = E_r\right] = \frac{1}{\beta\bar{\beta}}\log T + \log\left(\tilde{R}_r\right) + (P_1 - P_0)\log\left(\kappa^{-\frac{1}{\beta}}\right),$$

$$\mathbb{E}\left[\log\left(\tilde{w}_{jt}\right) \mid E_{jt} = E_c\right] = \int_{\tilde{R}_c}^{\tilde{R}_r} \frac{\left[\frac{1}{\beta}\log\left(\tilde{E}_c\right) - \frac{\theta}{\beta}\log\left(\tilde{R}\right)\right)}{\left(\tilde{R}_c\right)^{-\gamma_R} - \left(\tilde{R}_r\right)^{-\gamma_R}} dF_{\tilde{R}}(\tilde{R})$$

$$= \frac{1}{\beta}\log\left(\tilde{E}_c\right)$$

$$- \frac{\theta}{\beta}\frac{\left(\tilde{R}_c\right)^{-\gamma_R}\left[\log\left(\tilde{R}_c\right) + \frac{1}{\gamma_R}\right] - \left(\tilde{R}_r\right)^{-\gamma_R}\left[\log\left(\tilde{R}_r\right) + \frac{1}{\gamma_R}\right]}{\left(\tilde{R}_c\right)^{-\gamma_R} - \left(\tilde{R}_r\right)^{-\gamma_R}}$$

The average effect of board representation on wages:

$$\mathcal{M} = \mathbb{E}\left[\log\left(\tilde{w}\right) \mid B = 1\right] - \mathbb{E}\left[\log\left(\tilde{w}\right) \mid B = 0\right] \tag{E.41}$$

Let us evaluate each expected value individually. The average wage at firms with a board representative is given by:

$$\mathbb{E}\left[\log\left(\tilde{w}\right) \mid B=1\right] = \frac{P_0^M}{P_0^M + \left(P_1^M - P_0^M\right) T\left(\tilde{E}_r\right)^{-\bar{\beta}}} \times A_{B=1}$$
$$+ \frac{P_0^M}{P_0^M + \left(P_1^M - P_0^M\right) T\left(\tilde{E}_r\right)^{-\bar{\beta}}} \times C_{B=1}$$
$$+ \frac{P_1^M}{P_0^M + \left(P_1^M - P_0^M\right) T\left(\tilde{E}_r\right)^{-\bar{\beta}}} \times D_{B=1}$$

where:

$$A_{B=1} = \left[(1-P_0)\log\left(\kappa^{-\frac{1}{\beta}}\right)\right]\int_1^{\tilde{R}_c} dF_{\tilde{R}}(\tilde{R}) + \int_1^{\tilde{R}_c} \log\left(\tilde{R}\right) f_R\left(\tilde{R}\right) d\left(\tilde{R}\right)$$
$$= \left(\frac{1}{\gamma_R}\right)\cdot\left[1-\left(\tilde{E}_c\right)^{-\bar{\beta}}\left(1+\bar{\beta}\log\left(\tilde{E}_c\right)\right)\right] + (1-P_0)\log\left(\kappa^{-\frac{1}{\beta}}\right)\left[1-\left(\tilde{E}_c\right)^{-\bar{\beta}}\right],$$

$$C_{B=1} = \left[\frac{1}{\beta} ln\left(\tilde{E}_c\right) + (1-P_0)\log\left(\kappa^{-\frac{1}{\beta}}\right)\right]\int_{\tilde{R}_c}^{\tilde{R}_r} dF_{\tilde{R}}(\tilde{R})$$
$$- \frac{\theta}{\beta}\int_{\tilde{R}_c}^{\tilde{R}_r} \log\left(\tilde{R}\right) dF_{\tilde{R}}(\tilde{R})$$
$$= \left[\frac{1}{\beta}\log\left(\tilde{E}_c\right) + (1-P_0)\log\left(\kappa^{-\frac{1}{\beta}}\right)\right]\left[\left(\tilde{E}_c\right)^{-\bar{\beta}} - T\left(\tilde{E}_r\right)^{-\frac{\gamma_R}{\beta+\theta}}\right]$$
$$- \frac{\theta}{\beta\gamma_R}\left\{\left(\tilde{E}_c\right)^{-\bar{\beta}}\left(\bar{\beta}\log\left(\tilde{E}_c\right)+1\right) - T\left(\tilde{E}_r\right)^{-\bar{\beta}}\left(\bar{\beta}\log\left(\tilde{E}_r\right) - \log T + 1\right)\right\},$$

$$D_{B=1} = \left[\frac{1}{\beta\bar{\beta}}\log T + (1-P_0)\log\kappa^{-\frac{1}{\beta}}\right]\int_{\tilde{R}_r}^{\infty} dF_{\tilde{R}}(\tilde{R}) + \int_{\tilde{R}_r}^{\infty}\log\left(\tilde{R}\right) dF_{\tilde{R}}(\tilde{R})$$
$$= T\left(\tilde{E}_r\right)^{-\bar{\beta}}\left[\frac{1}{\beta+\theta}\log\left(\tilde{E}_r\right) + \frac{\beta\gamma_R}{\theta}\log T + (1-P_0)\log\left(\kappa^{-\frac{1}{\beta}}\right) + \frac{1}{\gamma_R}\right].$$

The average wage at firms without a board representative is given by:

$$
\begin{aligned}
\mathbb{E}\left[\log\left(\tilde{w}\right) \mid B=0\right] = & \frac{1-P_0^M}{(1-P_0^M)+(P_0^M-P_1^M)\,T\left(\tilde{E}_r\right)^{-\bar{\beta}}} \times A_{B=0} \\
& + \frac{1-P_0^M}{(1-P_0^M)+(P_0^M-P_1^M)\,T\left(\tilde{E}_r\right)^{-\bar{\beta}}} \times C_{B=0} \\
& + \frac{1-P_1^M}{(1-P_0^M)+(P_0^M-P_1^M)\,T\left(\tilde{E}_r\right)^{-\bar{\beta}}} \times D_{B=0}
\end{aligned}
$$

where:

$$
\begin{aligned}
A_{B=0} &= \left[-P_0 \log\left(\kappa^{-\frac{1}{\beta}}\right)\right] \int_1^{\tilde{R}_c} dF_{\tilde{R}}(\tilde{R}) + \int_1^{\tilde{R}_c} \log\left(\tilde{R}\right) dF_{\tilde{R}}(\tilde{R}) \\
&= \left(\frac{1}{\gamma_R}\right) \cdot \left[1-\left(\tilde{E}_c\right)^{-\bar{\beta}}\left(1+\bar{\beta}\log\left(\tilde{E}_c\right)\right)\right] + -P_0 \log\left(\kappa^{-\frac{1}{\beta}}\right)\left[1-\left(\tilde{E}_c\right)^{-\bar{\beta}}\right]
\end{aligned}
$$

$$
\begin{aligned}
C_{B=0} &= \left[\frac{1}{\beta} ln\left(\tilde{E}_c\right) - P_0 \log\left(\kappa^{-\frac{1}{\beta}}\right)\right] \int_{\tilde{R}_c}^{\tilde{R}_r} dF_{\tilde{R}}(\tilde{R}) \\
&- \frac{\theta}{\beta} \int_{\tilde{R}_c}^{\tilde{R}_r} \log\left(\tilde{R}\right) dF_{\tilde{R}}(\tilde{R}) \\
&= \left[\frac{1}{\beta}\log\left(\tilde{E}_c\right) - P_0 \log\left(\kappa^{-\frac{1}{\beta}}\right)\right]\left[\left(\tilde{E}_c\right)^{-\bar{\beta}} - T\left(\tilde{E}_r\right)^{-\frac{\gamma_R}{\beta+\theta}}\right] \\
&- \frac{\theta}{\beta\gamma_R}\left\{\left(\tilde{E}_c\right)^{-\bar{\beta}}\left(\bar{\beta}\log\left(\tilde{E}_c\right)+1\right) - T\left(\tilde{E}_r\right)^{-\bar{\beta}}\left(\bar{\beta}\log\left(\tilde{E}_r\right) - \log T + 1\right)\right\}
\end{aligned}
$$

$$
\begin{aligned}
D_{B=0} &= \left[\frac{1}{\beta\bar{\beta}}\log T - P_0 \log \kappa^{-\frac{1}{\beta}}\right] \int_{\tilde{R}_r}^{\infty} dF_{\tilde{R}}(\tilde{R}) \\
&+ \int_{\tilde{R}_r}^{\infty} \log\left(\tilde{R}\right) dF_{\tilde{R}}(\tilde{R}) \\
&= T\left(\tilde{E}_r\right)^{-\bar{\beta}}\left[\frac{1}{\beta+\theta}\log\left(\tilde{E}_r\right) + \frac{\beta\gamma_R}{\theta}\log T - P_0\log\left(\kappa^{-\frac{1}{\beta}}\right) + \frac{1}{\gamma_R}\right]
\end{aligned}
$$

E.3.2 Moment mappings with measurement error

In models with measurement error, the mappings for $\bar{\beta}$, T, and $\mathcal{M}$ are the same as under no measurement error. However, with measurement error, the expectation of (log) wages and value-added conditional on observed firm size cannot exhibit a discontinuity at $E = E_c$: the measurement error smooths out any potential discontinuity around the threshold. For this reason, when estimating the model with measurement error, we replace the RD with a simple comparison of the average outcome within a bandwidth to the left, and to the right of the threshold. We choose a bandwidth of 10. We map these moments to our model below.

Average difference in log value added around E_c (with bandwidth b):

$$\tilde{\rho} = \mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E_r \leq E^{OBS} \leq E_{r+b}\right] - \mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E_{c-b} \leq E^{OBS} \leq E_c\right], \tag{E.42}$$

where:

$$\mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E_r \leq E^{OBS} \leq E_{r+b}\right] = \int_{E_r}^{E_{r+b}} \mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E^{OBS} = E\right] \frac{\bar{\beta}E^{-\left(\bar{\beta}+1\right)}}{E_r^{-\bar{\beta}} - E_{r+b}^{-\bar{\beta}}} dE$$
$$\mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E_{c-b} \leq E^{OBS} \leq E_c\right] = \int_{E_{c-b}}^{E_c} \mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E^{OBS} = E\right] \frac{\bar{\beta}E^{-\left(\bar{\beta}+1\right)}}{E_{c-b}^{-\bar{\beta}} - E_c^{-\bar{\beta}}} dE$$

The expected value-added for a given observed firm size is integrated over all possible true firm sizes:

$$\begin{aligned}
&\mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E^{OBS} = E\right] \\
&= \int_{E_{min}}^{\infty} \mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E^{OBS} = E, E^* = x\right] f\left(x \mid E^{OBS} = E\right) dx \\
&= \int_{E_{min}}^{\infty} \underbrace{\mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E^* = x\right]}_{\text{Expected value-added for true E}} \underbrace{f\left(x \mid E^{OBS} = E\right)}_{\text{Measurement error weight}} dx
\end{aligned}$$

And the expected value-added at a given true firm size is given by:

$$\mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E^* = x, E^* < E_c\right] = \frac{1+\beta+\theta}{\beta+\theta} \log\left(\frac{x}{E_{min}}\right)$$

$$\mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E^{*}=x, E^{*}>E_{c}\right]=\frac{1+\beta+\theta}{\beta+\theta}\log\left(\frac{x}{E_{min}}\right)-\frac{\beta}{\beta+\theta}\log\left[\frac{1+P_{1}\left(\delta-1\right)}{1+P_{1}\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]$$
$$+\frac{\beta}{\beta+\theta}\log\left[\frac{1+P_{0}\left(\delta-1\right)}{1+P_{0}\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]+\left(P_{1}-P_{0}\right)\log\delta$$

$$\mathbb{E}\left[\log\left(\tilde{Y}\right) \mid E^{*}=E_{c}\right]=\frac{\left(\tilde{R}_{c}\right)^{-\gamma_{R}}\left[\log\left(\tilde{R}_{c}\right)+\frac{1}{\gamma_{R}}\right]-\left(\tilde{R}_{r}\right)^{-\gamma_{R}}\left[\log\left(\tilde{R}_{r}\right)+\frac{1}{\gamma_{R}}\right]}{\left(\tilde{R}_{c}\right)^{-\gamma_{R}}-\left(\tilde{R}_{r}\right)^{-\gamma_{R}}}+\log\left(\tilde{E}_{c}\right)$$

The second part of the integral is the likelihood of a firm having true firm size of x, given that it has an observed firm size of E. Using Bayes rule, we can decompose this likelihood:

$$f\left(x \mid E^{OBS}=E\right)=\frac{f\left(x, E^{OBS}=E\right)}{Pr\left(E^{OBS}=E\right)}=\frac{f\left(E^{OBS}=E \mid x\right)f\left(E^{*}=x\right)}{Pr\left(E^{OBS}=E\right)}$$

where:

- $f\left(E^{OBS}=E \mid x\right)=f\left(E^{*}\exp(\epsilon)=E \mid x\right)=\frac{1}{\sigma E}\phi\left(\frac{\log\left(\frac{E}{x}\right)}{\sigma}\right)$
- $f(E^{*}=x)$: no-error PDF
- $f\left(E^{OBS}=E\right)$: measurement-error PDF

Average difference in log wages around E_c (with bandwidth b): This moment has a similar mapping to the previous one:

$$\tilde{\xi}=\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E_{r}\leq E^{OBS}\leq E_{r+b}\right]-\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E_{c-b}\leq E^{OBS}\leq E_{c}\right] \tag{E.43}$$

where

$$\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E_{r}\leq E^{OBS}\leq E_{r+b}\right]=\int_{E_{r}}^{E_{r+b}}\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E^{OBS}=E\right]\frac{\bar{\beta}E^{-\left(\bar{\beta}+1\right)}}{E_{r}^{-\bar{\beta}}-E_{r+b}^{-\bar{\beta}}}dE$$
$$\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E_{c-b}\leq E^{OBS}\leq E_{c}\right]=\int_{E_{c-b}}^{E_{c}}\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E^{OBS}=E\right]\frac{\bar{\beta}E^{-\left(\bar{\beta}+1\right)}}{E_{c-b}^{-\bar{\beta}}-E_{c}^{-\bar{\beta}}}dE$$

The expected wage for a given observed firm size is integrated over all possible true firm sizes:

$$\begin{aligned}
&\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E^{OBS} = E\right] \\
&= \int_{E_{min}}^{\infty} \mathbb{E}\left[\log\left(\tilde{w}\right) \mid E^{OBS} = E, E^{*} = x\right] f\left(x \mid E^{OBS} = E\right) dx \\
&= \int_{E_{min}}^{\infty} \underbrace{\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E^{*} = x\right]}_{\text{Expected wage for true E}} \underbrace{f\left(x \mid E^{OBS} = E\right)}_{\text{Measurement error weight}} dx
\end{aligned}$$

And the expected wage at a given true firm size is given by:

$$\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E^{*} = x, E^{*} < E_{c}\right] = \frac{1}{\beta+\theta} \log\left(\frac{x}{E_{min}}\right)$$

$$\begin{aligned}
\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E^{*} = x, E^{*} > E_{c}\right] = \frac{1}{\beta+\theta} \log\left(\frac{x}{E_{min}}\right) &+ \frac{\theta}{\beta+\theta} \log\left[\frac{1+P_{1}(\delta-1)}{1+P_{1}\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right] \\
&- \frac{\theta}{\beta+\theta} \log\left[\frac{1+P_{0}(\delta-1)}{1+P_{0}\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right] + (P_{1}-P_{0}) \log\left(\kappa^{-\frac{1}{\beta}}\right)
\end{aligned}$$

$$\mathbb{E}\left[\log\left(\tilde{w}\right) \mid E^{*} = E_{c}\right] = \frac{1}{\beta} \log\left(\tilde{E}_{c}\right) - \frac{\theta}{\beta} \frac{\left(\tilde{R}_{c}\right)^{-\gamma_{R}}\left[\log\left(\tilde{R}_{c}\right)+\frac{1}{\gamma_{R}}\right] - \left(\tilde{R}_{r}\right)^{-\gamma_{R}}\left[\log\left(\tilde{R}_{r}\right)+\frac{1}{\gamma_{R}}\right]}{\left(\tilde{R}_{c}\right)^{-\gamma_{R}} - \left(\tilde{R}_{r}\right)^{-\gamma_{R}}}$$

E.4 Definining a maximum firm size for aggregate effects

For computations, when calculating the aggregate effects of the policy on firms and workers in the economy, it is helpful to assume that there is a maximum productivity level R_{max} corresponding to some (large) maximum firm size E_{max}:

$$R_{max} = \left(\frac{\beta}{1+\beta}\right)^{-\frac{\beta}{\beta+\theta}} \left[\frac{1+P_{1}(\delta-1)}{1+P_{1}\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{-\frac{\beta}{\beta+\theta}} \left(\frac{E_{0}}{\exp\left(\tilde{W}\right)}\right)^{-\frac{1}{\beta+\theta}} E_{max}^{\frac{1}{\beta+\theta}}$$

We set $E_{min} = 10$ and $E_{max} = 7056$. Results are robust to varying E_{min} and E_{max}.

E.5 Backing out the wage indices and E_0

To compute the aggregate effects of the policy, we must first compute the wage index under the policy and under the counterfactual. The wage index under the policy is given by[14]:

$$\begin{aligned}
\exp\left(\tilde{W}_1\right) &= \int_{R_{min}}^{R_c} \exp(a_{jt}) \times [w_0^u(R)]^\beta \, dF_R\left(R \mid R \leq R_{max}\right) \\
&+ \int_{R_c}^{R_r} \exp(a_{jt}) \times [w_0^c(R)]^\beta \, dF_R\left(R \mid R \leq R_{max}\right) \\
&+ \int_{R_r}^{R_{max}} \exp(a_{jt}) \times [w_1(R)]^\beta \, dF_R\left(R \mid R \leq R_{max}\right) \\
&= \left[\frac{\gamma_r R_{min}^{\gamma_r}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_r}}\right]\left(\frac{\beta}{1+\beta}\right)^\beta \\
&\times \left\{1+\frac{E_c}{E_0}\left[1+P_0(\kappa-1)\right]\left[1+P_0(\kappa^{-1}-1)\right]\left[\frac{\gamma_r R_{min}^{\gamma_r}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma}_r}\right]\left[\frac{R_r^{-\gamma_r}-R_c^{-\gamma_r}}{\gamma_r}\right]\right\}^{-1} \\
&\times \left\{\left[1+P_0(\kappa-1)\right]\left[1+P_0\left(\kappa^{-1}-1\right)\right]\left[\frac{1+P_0(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^\beta\left[\frac{R_c^{\beta+\theta-\gamma_r}-R_{min}^{\beta+\theta-\gamma_r}}{\beta+\theta-\gamma_r}\right]\right. \\
&\left.+\left[1+P_1(\kappa-1)\right]\left[1+P_1\left(\kappa^{-1}-1\right)\right]\left[\frac{1+P_1(\delta-1)}{1+P_1\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^\beta\left[\frac{R_{max}^{\beta+\theta-\gamma_r}-R_r^{\beta+\theta-\gamma_r}}{\beta+\theta-\gamma_r}\right]\right\}
\end{aligned} \tag{E.44}$$

The wage index under the counterfactual is given by[15]:

$$\begin{aligned}
\exp\left(\tilde{W}_0\right) &= \int_{R_{min}}^{R_{max}} \exp\left(a\right) \times [w_0^u(R)]^\beta \, dF_R\left(R \mid R \leq R_{max}\right) \\
&= \left[1+P_0\left(\kappa-1\right)\right]\left[1+P_0\left(\kappa^{-1}-1\right)\right]\left[\frac{1+P_0\left(\delta-1\right)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^\beta\left(\frac{\beta}{1+\beta}\right)^\beta \\
&\times \frac{\gamma_R R_{min}^{\gamma_R}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \times \frac{R_{max}^{\beta+\theta-\gamma_R}-R_{min}^{\beta+\theta-\gamma_R}}{\beta+\theta-\gamma_R}
\end{aligned}$$

[14] As $R_{max} \to \infty$, $\tilde{W}_1$ is finite if and only if $\gamma_R > \beta+\theta$. This condition holds across all of our specifications.

[15] As $R_{max} \to \infty$, $\tilde{W}_0$ is finite if and only if $\gamma_R > \beta+\theta$.

To compute $\exp\left(\tilde{W}_1\right)$ and the aggregate effects, we also need E_0. To retrieve an implied estimate for E_0, we first back out $\frac{E_0}{\exp\left(\tilde{W}_1\right)}$ using w_0^c from the data to get R_c, then using E_c and R_c:

$$\begin{aligned}
R_c &= \frac{1+\beta}{\beta}\left[1+P_0\left(\kappa^{-\frac{1}{\beta}}-1\right)\right]^{-1}\left[\frac{1+P_0(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{-1} w_0^c \\
\frac{E_0}{\exp\left(\tilde{W}_1\right)} &= \left(\frac{1+\beta}{\beta}\right)^{\beta}\left[\frac{1+P_0(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{-\beta} E_c R_c^{-(\beta+\theta)}
\end{aligned} \tag{E.45}$$

We then recover a value E_0 that satisfies both expressions for $\frac{E_0}{\exp\left(\tilde{W}_1\right)}$:

$$\frac{E_0}{\exp\left(\tilde{W}_1\right)} = \left(\frac{1+\beta}{\beta}\right)^{\beta}\left[\frac{1+P_0(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{-\beta} E_c R_c^{-(\beta+\theta)},$$

where $\exp\left(\tilde{W}_1\right)$ on the left hand side of the equation is given by Equation (E.44) and the right hand side comes from the relationship between E_c and R_c in Equation (E.45).

E.6 Computing the aggregate effects of the policy

Once we have estimated the underlying parameters of the model, we can use them to compute the model-implied aggregate effects of the policy. We compute the aggregate effects on firm profits and on workers' wages in the economy.

Aggregate effects on firm profits: We start by computing the total effect of the policy on firm profits. The average profits under the policy are given by:

$$\Pi_1 \equiv \omega_0^u \times \underbrace{\mathbb{E}\left[\pi_0^u(R, \tilde{W}_1) \mid R_{min} \leq R < R_c\right]}_{\text{Unregulated firms}} + \omega_0^c \times \underbrace{\mathbb{E}\left[\pi_0^c(R, \tilde{W}_1) \mid R_c \leq R < R_r\right]}_{\text{Bunching firms}}$$

$$+ \omega_1 \times \underbrace{\mathbb{E}\left[\pi_1(R, \tilde{W}_1) \mid R \geq R_r\right]}_{\text{Regulated firms}}$$

$$= \frac{1 + P_0 (\delta - 1)}{1 + \beta} \left[\frac{1 + P_0 (\delta - 1)}{1 + P_0 \left(\kappa^{-\frac{1}{\beta}} \tau - 1\right)}\right]^{\beta} \left(\frac{\beta}{1 + \beta}\right)^{\beta} \frac{E_0}{\exp\left(\tilde{W}\right)}$$

$$\times \frac{\gamma_R R_{min}^{\gamma_R}}{1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[\frac{R_c^{\beta+\theta-\gamma_R+1} - R_{min}^{\beta+\theta-\gamma_R+1}}{\beta + \theta - \gamma_R + 1}\right]$$

$$- \frac{P_0 F R_{min}^{\gamma_R}}{1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[R_{min}^{-\gamma_R} - R_c^{-\gamma_R}\right]$$

$$+ E_c \left[1 + P_0 (\delta - 1)\right] \frac{\gamma_R R_{min}^{\gamma_R}}{1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[\frac{R_r^{1-\gamma_R} - R_c^{1-\gamma_R}}{1 - \gamma_R}\right]$$

$$- E_c \left[1 + P_0 \left(\kappa^{-\frac{1}{\beta}} \tau - 1\right)\right] \left[\frac{E_c}{E_0} \exp\left(\tilde{W}\right)\right]^{\frac{1}{\beta}} \frac{\gamma_R R_{min}^{\gamma_R}}{1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[\frac{R_c^{-\left(\frac{\theta}{\beta}+\gamma_R\right)} - R_r^{-\left(\frac{\theta}{\beta}+\gamma_R\right)}}{\frac{\theta}{\beta} + \gamma_R}\right]$$

$$- \frac{P_0 F R_{min}^{\gamma_R}}{1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[R_c^{-\gamma_R} - R_r^{-\gamma_R}\right]$$

$$+ \frac{1 + P_1 (\delta - 1)}{1 + \beta} \left[\frac{1 + P_1 (\delta - 1)}{1 + P_1 \left(\kappa^{-\frac{1}{\beta}} \tau - 1\right)}\right]^{\beta} \left(\frac{\beta}{1 + \beta}\right)^{\beta} \frac{E_0}{\exp\left(\tilde{W}\right)}$$

$$\times \frac{\gamma_R R_{min}^{\gamma_R}}{1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[\frac{R_{max}^{\beta+\theta-\gamma_R+1} - R_r^{\beta+\theta-\gamma_R+1}}{\beta + \theta - \gamma_R + 1}\right]$$

$$- \frac{P_1 F R_{min}^{\gamma_R}}{1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[R_r^{-\gamma_R} - R_{max}^{-\gamma_R}\right]$$

The average profits under the counterfactual are:

$$
\begin{aligned}
\Pi_0 \equiv\ & \mathbb{E}\left[\pi_0^u(R,\tilde{W}_0)\right] \\
=\ & \frac{1+P_0\,(\delta-1)}{1+\beta}\left[\frac{1+P_0\,(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{\beta}\left(\frac{\beta}{1+\beta}\right)^{\beta}\frac{E_0}{\exp\left(\tilde{W}\right)} \\
& \times \frac{\gamma_R R_{min}^{\gamma_R}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}}\left[\frac{R_{max}^{\beta+\theta-\gamma_R+1}-R_{min}^{\beta+\theta-\gamma_R+1}}{\beta+\theta-\gamma_R+1}\right] \\
& - P_0 F\frac{R_{min}^{\gamma_R}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}}\left[R_{min}^{-\gamma_R}-R_{max}^{-\gamma_R}\right]
\end{aligned}
$$

We decompose the total effect of the policy on firm profits into three components:

$$
\begin{aligned}
\Pi_1-\Pi_0 =\ & \underbrace{\omega_1\times\mathbb{E}\left[\pi_1(R,\tilde{W}_1)-\pi_0^u(R,\tilde{W}_1)\mid R\geq R_r\right]}_{\text{Direct policy effect }\equiv\Delta\Pi^{DE}} \\
& +\underbrace{\omega_0^c\times\mathbb{E}\left[\pi_0^c(R,\tilde{W}_1)-\pi_0^u(R,\tilde{W}_1)\mid R_c\leq R<R_r\right]}_{\text{Size Distortion Effect}\equiv\Delta\Pi^{SD}}+\underbrace{\mathbb{E}\left[\pi_0^u(R,\tilde{W}_1)-\pi_0^u(R,\tilde{W}_0)\right]}_{\text{Equilibrium Effects }\equiv\Delta\Pi^{EE}}
\end{aligned}
$$

The direct policy effect is given by:

$$
\begin{aligned}
\Delta\Pi^{DE} =\ & \int_{R_r}^{R_{max}}\left[\pi_1\left(R,\tilde{W}_1\right)-\pi_0^u\left(R,\tilde{W}_1\right)\right]\frac{f_R(R)}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}}dR \\
=\ & \left[\frac{\gamma_R R_{min}^{\gamma_R}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}}\right]\left(\frac{\beta}{1+\beta}\right)^{\beta}\left(\frac{E_0}{\exp\left(\tilde{W}_1\right)}\right) \\
& \times\left\{\frac{1+P_1\,(\delta-1)}{1+\beta}\left[\frac{1+P_1\,(\delta-1)}{1+P_1\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{\beta}-\frac{1+P_0\,(\delta-1)}{1+\beta}\left[\frac{1+P_0\,(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{\beta}\right\} \\
& \times\left[\frac{R_{max}^{\beta+\theta-\gamma_R+1}-R_r^{\beta+\theta-\gamma_R+1}}{\beta+\theta-\gamma_R+1}\right] \\
& -(P_1-P_0)\frac{R_{min}^{\gamma_R}F}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}}\left[R_r^{-\gamma_R}-R_{max}^{-\gamma_R}\right]
\end{aligned}
$$

The size distortion effect is given by:

$$\begin{aligned}
\Delta\Pi^{SD} &= \int_{R_c}^{R_r} \left[\pi_0^c\left(R,\tilde{W}_1\right) - \pi_0^u\left(R,\tilde{W}_1\right)\right] \frac{f_R(R)}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} dR \\
&= E_c\left[1+P_0\left(\delta-1\right)\right] \frac{\gamma_R R_{min}^{\gamma_R}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[\frac{R_r^{1-\gamma_R}-R_c^{1-\gamma_R}}{1-\gamma_R}\right] \\
&- E_c\left[1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)\right]\left[\frac{E_c}{E_0}\exp\left(\tilde{W}_1\right)\right]^{\frac{1}{\beta}} \frac{\gamma_R R_{min}^{\gamma_R}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[\frac{R_c^{-\left(\frac{\theta}{\beta}+\gamma_R\right)}-R_r^{-\left(\frac{\theta}{\beta}+\gamma_R\right)}}{\frac{\theta}{\beta}+\gamma_R}\right] \\
&- \frac{1+P_0\left(\delta-1\right)}{1+\beta}\left[\frac{1+P_0\left(\delta-1\right)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{\beta}\left(\frac{\beta}{1+\beta}\right)^{\beta}\frac{E_0}{\exp\left(\tilde{W}_1\right)} \\
&\times \frac{\gamma_R R_{min}^{\gamma_R}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[\frac{R_r^{\beta+\theta-\gamma_R+1}-R_c^{\beta+\theta-\gamma_R+1}}{\beta+\theta-\gamma_R+1}\right]
\end{aligned}$$

The equilibrium effects are given by:

$$\begin{aligned}
\Delta\Pi^{GE} &= \int_{R_{min}}^{R_{max}} \left[\pi_0^u\left(R,\tilde{W}_1\right) - \pi_0^u\left(R,\tilde{W}_0\right)\right] \frac{f_R(R)}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} dR \\
&= E_0\left[\frac{\exp\left(\tilde{W}_0\right)-\exp\left(\tilde{W}_1\right)}{\exp\left(\tilde{W}_0\right)\times\exp\left(\tilde{W}_1\right)}\right]\frac{1+P_0\left(\delta-1\right)}{1+\beta}\left[\frac{1+P_0\left(\delta-1\right)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{\beta}\left(\frac{\beta}{1+\beta}\right)^{\beta} \\
&\times \frac{\gamma_R R_{min}^{\gamma_R}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[\frac{R_{max}^{\beta+\theta-\gamma_R+1}-R_{min}^{\beta+\theta-\gamma_R+1}}{\beta+\theta-\gamma_R+1}\right]
\end{aligned}$$

Aggregate effects on worker wages: We start by computing the total effect of the policy on average worker wages. The average wage under the policy is given by:

$$
\begin{aligned}
W_1 &\equiv \omega_0^u \times \mathbb{E}\left[w_0^u(R) \mid R_{min} \leq R < R_c\right] \\
&\quad + \omega_0^c \times \mathbb{E}\left[w_0^c(R) \mid R_c \leq R < R_r\right] \\
&\quad + \omega_1 \times \mathbb{E}\left[w_1(R) \mid R \geq R_r\right] \\
&= \frac{\gamma_R R_{min}^{\gamma_R} \left(\gamma_R - \beta - \theta - 1\right)^{-1}}{E_T\left[1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}\right]} \left(\frac{\beta}{1+\beta}\right)^{\beta+1} \left[\frac{1 + P_0(\delta - 1)}{1 + P_0\left(\kappa^{-\frac{1}{\beta}}\tau - 1\right)}\right]^{\beta+1} \\
&\quad \times \frac{E_0}{\exp\left(\tilde{W}_1\right)} \left[1 + P_0\left(\kappa^{-\frac{1}{\beta}} - 1\right)\right] \left[R_{min}^{-(\gamma_R - \beta - \theta - 1)} - R_c^{-(\gamma_R - \beta - \theta - 1)}\right] \\
&\quad - \frac{E_c}{E_T}\left[\frac{E_c}{E_0}\exp\left(\tilde{W}_1\right)\right]^{\frac{1}{\beta}} \frac{\gamma_R R_{min}^{\gamma_R}}{1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[1 + P_0\left(\kappa^{-\frac{1}{\beta}} - 1\right)\right] \left[\frac{R_r^{-\left(\frac{\theta}{\beta} + \gamma_R\right)} - R_c^{-\left(\frac{\theta}{\beta} + \gamma_R\right)}}{\frac{\theta}{\beta} + \gamma_R}\right] \\
&\quad + \frac{\gamma_R R_{min}^{\gamma_R} \left(\gamma_R - \beta - \theta - 1\right)^{-1}}{E_T\left[1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}\right]} \left(\frac{\beta}{1+\beta}\right)^{\beta+1} \left[\frac{1 + P_1(\delta - 1)}{1 + P_1\left(\kappa^{-\frac{1}{\beta}}\tau - 1\right)}\right]^{\beta+1} \\
&\quad \times \frac{E_0}{\exp\left(\tilde{W}_1\right)} \left[1 + P_1\left(\kappa^{-\frac{1}{\beta}} - 1\right)\right] \left[R_r^{-(\gamma_R - \beta - \theta - 1)} - R_{max}^{-(\gamma_R - \beta - \theta - 1)}\right]
\end{aligned}
$$

The average wage under the counterfactual is:

$$
\begin{aligned}
W_0 &\equiv \mathbb{E}\left[w_0^u(R)\right] \\
&= \frac{\gamma_R R_{min}^{\gamma_R}\left(\gamma_R-\beta-\theta-1\right)^{-1}}{E_T\left[1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}\right]}\left[\frac{1+P_0(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{\beta+1} \\
&\times \frac{E_0}{\exp\left(\tilde{W}_1\right)}\left(\frac{\beta}{1+\beta}\right)^{\beta+1}\left[1+P_0\left(\kappa^{-\frac{1}{\beta}}-1\right)\right]\left[R_{min}^{-(\gamma_R-\beta-\theta-1)}-R_c^{-(\gamma_R-\beta-\theta-1)}\right] \\
&+\frac{E_c}{E_T}\frac{\gamma_R R_{min}^{\gamma_R}}{1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}}\left[\frac{1+P_0(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right] \\
&\times\left[1+P_0\left(\kappa^{-\frac{1}{\beta}}-1\right)\right]\left(\frac{\beta}{1+\beta}\right)\left[\frac{R_c^{1-\gamma_R}-R_r^{1-\gamma_R}}{\gamma_R-1}\right] \\
&+\frac{\gamma_R R_{min}^{\gamma_R}\left(\gamma_R-\beta-\theta-1\right)^{-1}}{E_T\left[1-\left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}\right]}\left[\frac{1+P_0(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]\left[\frac{1+P_1\left(\delta-1\right)}{1+P_1\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{\beta} \\
&\times\left(\frac{\beta}{1+\beta}\right)^{\beta+1}\frac{E_0}{\exp\left(\tilde{W}_1\right)}\left[1+P_0\left(\kappa^{-\frac{1}{\beta}}-1\right)\right]\left[R_r^{-(\gamma_R-\beta-\theta-1)}-R_{max}^{-(\gamma_R-\beta-\theta-1)}\right]
\end{aligned}
$$

Since we are weighting the above averages by firm size, we must also compute E_T, the total number of workers in the economy in the same units[16] as E:

$$
\begin{aligned}
E_T &= \int_{R_{min}}^{\infty} E\left(R,\tilde{W}_1\right) f_R(R)dR \\
&= \left(\frac{\beta}{1+\beta}\right)^{\beta}\left[\frac{1+P_0(\delta-1)}{1+P_0\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{\beta+1}\left[1+P_0\left(\kappa^{-\frac{1}{\beta}}-1\right)\right]\frac{E_0}{\exp\left(\tilde{W}_1\right)} \\
&\times \frac{R_{min}^{\gamma_R}\gamma_R}{\gamma_R-\beta-\theta}\left[R_{min}^{\beta+\theta-\gamma_R}-R_c^{\beta+\theta-\gamma_R}\right] \\
&+ E_c R_{min}^{\gamma_R}\left[R_c^{-\gamma_R}-R_r^{-\gamma_R}\right] \\
&+\left[\frac{1+P_1(\delta-1)}{1+P_1\left(\kappa^{-\frac{1}{\beta}}\tau-1\right)}\right]^{\beta+1}\left[1+P_1\left(\kappa^{-\frac{1}{\beta}}-1\right)\right]\left(\frac{\beta}{1+\beta}\right)^{\beta}\frac{E_0}{\exp\left(\tilde{W}_1\right)}\frac{\gamma_R R_{min}^{\gamma_R}}{\gamma_R-\beta-\theta}R_r^{\beta+\theta-\gamma_R}
\end{aligned}
$$

[16]Since we take E_0 as given, we do not know its units. Hence, we cannot use E_0 to weight by workers because we do not know if it is in the same units as E.

We can similarly decompose the total effect of the policy on average wages into two components:

$$W_1 - W_0 = \underbrace{\omega_0^c \times \mathbb{E}\left[w_0^c(R, \tilde{W}_1) - w_0^u(R) \mid R_c \leq R < R_r\right]}_{\text{Size Distortion Effect} \equiv \Delta W^{SD}} + \underbrace{\omega_1 \times \mathbb{E}\left[w_1(R) - w_0^u(R) \mid R \geq R_r\right]}_{\text{Direct Policy Effect} \equiv \Delta W^{DE}}$$

The size distortion effect is given by:

$$\begin{aligned}
\Delta W^{SD} = {} & \frac{E_c}{E_T} \frac{\gamma_R R_{min}^{\gamma_R}}{1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}} \left[1 + P_0\left(\kappa^{-\frac{1}{\beta}} - 1\right)\right] \\
& \times \left\{ \left[\frac{E_c}{E_0} \exp\left(\tilde{W}_1\right)\right]^{\frac{1}{\beta}} \left[\frac{R_c^{-\left(\frac{\theta}{\beta}+\gamma_R\right)} - R_r^{-\left(\frac{\theta}{\beta}+\gamma_R\right)}}{\frac{\theta}{\beta} + \gamma_R}\right] \right. \\
& \left. - \left(\frac{\beta}{1+\beta}\right) \left[\frac{1 + P_0(\delta - 1)}{1 + P_0\left(\kappa^{-\frac{1}{\beta}}\tau - 1\right)}\right] \left[\frac{R_c^{1-\gamma_R} - R_r^{1-\gamma_R}}{\gamma_R - 1}\right] \right\}
\end{aligned}$$

The direct effect is given by:

$$\begin{aligned}
\Delta W^{DE} = {} & \frac{\gamma_R R_{min}^{\gamma_R} \left(\gamma_R - \beta - \theta - 1\right)^{-1}}{E_T \left[1 - \left(\frac{R_{min}}{R_{max}}\right)^{\gamma_R}\right]} \left(\frac{\beta}{1+\beta}\right)^{\beta+1} \left[\frac{1 + P_1(\delta - 1)}{1 + P_1\left(\kappa^{-\frac{1}{\beta}}\tau - 1\right)}\right]^{\beta} \frac{E_0}{\exp\left(\tilde{W}_1\right)} \\
& \times \left\{ \left[1 + P_1\left(\kappa^{-\frac{1}{\beta}} - 1\right)\right] \left[\frac{1 + P_1\left(\delta - 1\right)}{1 + P_1\left(\kappa^{-\frac{1}{\beta}}\tau - 1\right)}\right] \right. \\
& \left. - \left[1 + P_0\left(\kappa^{-\frac{1}{\beta}} - 1\right)\right] \left[\frac{1 + P_0\left(\delta - 1\right)}{1 + P_0\left(\kappa^{-\frac{1}{\beta}}\tau - 1\right)}\right] \right\} \\
& \times \left[R_r^{-(\gamma_R - \beta - \theta - 1)} - R_{max}^{-(\gamma_R - \beta - \theta - 1)}\right]
\end{aligned}$$

E.7 The bootstrapping procedure

We bootstrap to get standard errors for all moments estimated by the MLE, all parameters estimated by the GMM, and all additional parameters presented in this paper (F, the aggregate effects, and so on). We use the following procedure to do so:

1. From the full worker-level sample, draw a random sample of n workers using block sampling with replacement, where each block is a firm.

2. Restrict to the mover sub-sample and estimate the worker representation pay premium, $\mathcal{M}$.

3. Aggregate the worker-level bootstrap sample to the firm level.

4. Given the firm-level bootstrap data, estimate all firm size moments using MLE $(\bar{\beta}, T, E_r, ...)$, ρ, and the RD of wages ξ.

5. Given the above moments, estimate the parameters using GMM.

6. Compute the other implied parameters (F, the aggregate effects, and so on).

We repeat this process until we have 100 draws of each moment and parameter for all models. We then take the standard deviation of the estimated parameters and the aggregate effects over the draws for each model.