The Aggregate Effects of Enforcing Labor Regulations

Eshaan Patel*
Columbia University

Dafne Murillo†
Columbia University

Job Market Paper

November 1, 2026

Click here for the latest version.

Abstract

Labor regulations and their enforcement are traditionally viewed as contributing to lower employment in the formal sector. However, when labor markets are imperfectly competitive, enforcing regulations can lead to increased hiring. This paper builds a quantitative general equilibrium model to analyze how the endogenous enforcement of labor regulations interacts with markdowns in a monopsonistic labor market. We estimate the model using a novel dataset from Peru's labor inspection agency along with labor market microdata. We then quantify the effect of enforcement on aggregate outcomes. Relative to a world without enforcement, status quo enforcement slightly worsens misallocation, while perfect enforcement sharply reduces formal employment and lowers welfare by 15% of wages. Welfare has an inverted U-shape in enforcement intensity and peaks when the cost to workers of filing a complaint is a third of its level in the status quo. Ultimately, the aggregate effects of labor regulations depend on the enforcement regime, not just on the regulations themselves.

*I am deeply grateful to Suresh Naidu and Conor Walsh for invaluable guidance and support. I also thank Michael Best, Hugo Lhuillier, Eric Verhoogen, and participants of the Columbia University Development Colloquium. Finally, I would like to thank my parents and my wife for their unwavering support. Email: eshaan.patel@columbia.edu.

†Email: dm3160@columbia.edu.

1 Introduction

Governments regulate labor markets by mandating that firms provide their workers with additional compensation and amenities. Regulations are almost always paired with an enforcement regime, which induces firms to comply and provides workers what they are owed. However, enforcement is often imperfect, especially in developing countries, where the probability of inspection rises with firm size (Almeida and Ronconi, 2016). This inspection-size relationship means that firms face higher expected penalties as they grow, which holds back productive firms and creates misallocation of labor (Ulyssea, 2018). Thus, labor regulations and their enforcement are traditionally viewed as contributing to lower employment in the formal sector (Botero et al., 2004; World Bank, 2004; International Monetary Fund, 2024).

However, imperfect competition in the labor market complicates this view. A large literature finds that labor markets across a variety of contexts are imperfectly competitive (Manning, 2011; Azar and Marinescu, 2024; Caldwell et al., 2026). This endows firms with labor market power: firms mark down compensation below a worker's marginal product and hire fewer workers than is efficient. Moreover, when markdowns increase with firm size, labor is misallocated toward smaller firms (Berger et al., 2022). In this setting, enforcing labor regulations can increase hiring and improve efficiency. Enforcement increases the expected compensation of workers at larger firms, which have higher inspection rates and larger markdowns, and this lowers those firms' cost of attracting workers. At the same time, expected penalties raise the cost of hiring. The combination of enforcement and labor market power is therefore crucial to understanding the real-world effects of labor regulations, but their net effect on aggregate outcomes is ambiguous.

This paper builds a quantitative general equilibrium model of the endogenous enforcement of labor regulations in a monopsonistic labor market with an informal sector. We focus on the Peruvian labor market, where labor regulations mandate additional compensation and amenities. However, enforcement is imperfect: it occurs through inspections triggered by worker complaints. We obtain a novel dataset from the Peruvian labor inspection agency covering the universe of inspected firms and their inspection outcomes, along with detailed violation-level data. We build our quantitative model around the Peruvian inspection process and estimate it using legal documents, inspection data, and labor market microdata. We use the estimated model to ask: how does enforcement affect firm size and compliance, misallocation, and welfare?

In our model, workers have idiosyncratic preferences over firms and the informal sector outside option, which results in an upward-sloping labor supply curve facing each firm. Firms offer value to workers to move up the labor supply curve through a wage and a compliance level, which is the probability that they honor labor regulations on compensation and amenities. Workers value compliance because it increases delivered utility, and they observe each firm's wage and expected compliance before choosing where to work.

Once workers start working, the compliance probability is realized and workers may not receive the compensation or amenities they are owed. In this case, workers can complain to trigger an inspection if the value of the regulation exceeds their complaint cost. This creates a sliding scale of enforcement intensity – from no enforcement (infinite complaint cost) to perfect enforcement (zero complaint cost). Upon inspection, firms must pay back workers their owed compensation. Workers value this expected backpay, and a job with a higher inspection rate is worth more for a given wage and compliance rate. The inspection and violations data show that the number of violations per worker (i.e., noncompliance) falls with firm size, but the *total* number of violations in the workforce rises with firm size. Consistent with this evidence, we observe that the probability of a complaint-driven inspection rises with firm size.

Increasing compliance allows firms to lower the wage and avoid penalties. However, honoring an obligation has an opportunity cost, such as raising cash to pay mandated compensation or pausing work to provide a mandated amenity, and these costs vary across states. This makes compliance probabilistic: firms honor obligations in the cheapest states first, so raising compliance promises payment in increasingly expensive states and is convexly costly. Estimated compliance rises from 39% at the smallest firms to 93% at the largest but never reaches 100%. Compliance delivers a large portion of the value that workers receive, and they pay for it through lower wages. Firms comply at similar rates without enforcement, so attracting workers, not deterrence, accounts for the compliance observed in the data.

We find that enforcement need not increase compliance: backpay from enforcement can crowd out a firm's voluntary compliance. Enforcement raises the probability of inspection and the expected penalty, increasing the incentive to comply. However, workers value the backpay they expect from inspections, so firms that are more likely to be inspected can comply less voluntarily and still attract workers. Complying also lowers the backpay workers expect, which further lowers the return to compliance. On net, status quo enforcement lowers employment-weighted compliance from 84% to 83%. With high enough enforcement rates,

though, the deterrence effect can dominate. Perfect enforcement raises employment-weighted compliance to 92%, and small firms' wages fall to the minimum wage.

The intensity of enforcement determines how labor regulations affect firm size. We can see this most clearly by looking at firm choices around a size threshold of 20 workers, above which a firm begins to owe profit sharing and must create a safety and health committee. Empirically, we observe only mild bunching at this threshold in Peru. Our model reproduces this observation and offers an explanation: crossing the threshold raises the expected penalty by only 0.3% of profit because few firms are inspected. However, under perfect enforcement, the expected penalty jump is 26% of profit, and the model produces a spike below 20 workers and a hole above it, as seen in other contexts (Garicano et al., 2016; Almunia and Lopez-Rodriguez, 2018). The mild bunching observed in Peru is therefore a sign of weak enforcement, not necessarily of a low cost of the regulation. The same pattern holds in the aggregate: status quo enforcement barely moves formal employment, while perfect enforcement cuts it from 25% to 13% of the workforce.

We analyze how enforcement contributes to misallocation by deriving the wedge between marginal product and the cost of delivering value to a worker. When markets are perfectly competitive, firms set marginal product equal to the cost of delivering value to the marginal worker. This is the efficient benchmark, where the wedge is zero. With imperfectly competitive labor markets (and no enforcement), the wedge is the monopsony markdown: firms must offer raises to inframarginal workers as they move up the labor supply curve. Markdowns leave the formal sector inefficiently small, and under-hiring is worst at the largest firms, which have the largest markdowns. Employment in the formal sector is 46% of the workforce in the efficient allocation, against 25% with markdowns and no enforcement.

Enforcement changes this wedge through additional forces that increase and decrease the cost of hiring a marginal worker, reallocating labor across firms. Enforcement raises the cost of hiring because penalties grow with employment. The largest such force is the opportunity cost of paying backpay and sanctions after an inspection, which is similar to the cost of voluntary compliance except that the government requires the firm to pay even in expensive states. Enforcement also lowers the cost of hiring, because workers value the expected backpay at larger, more inspected firms. In the status quo, these forces nearly offset below a few hundred workers, and the wedge rises only at the largest firms, where markdowns are already largest. Under perfect enforcement, inspections are guaranteed at every firm, so the hiring advantage that imperfect enforcement gave large firms disappears.

Small firms that comply at high rates also hit the minimum wage and turn workers away into the informal sector. Together, these raise the employment-weighted wedge by 106% relative to no enforcement.

Finally, we analyze the welfare change across enforcement regimes. We measure welfare as total surplus: firm profits plus worker utility converted to money, less the net cost of enforcement. We decompose the welfare change into three sources: the allocation of labor discussed above, the change in the wages and compliance that firms post, and the net cost of enforcement. Welfare in the no-enforcement economy is below the efficient benchmark by 12% of wages due to markdowns. All of this loss comes from the allocation of labor, and 77% of it comes from firms above 3,000 workers.

In the status quo, welfare rises by 0.3% of wages. Enforcement's effect on the wedge misallocates labor and lowers welfare by 0.44% of wages. Tilting the mix of wages and compliance to reduce penalties lowers it by 0.94%, and the net cost of enforcement lowers it by 0.54%, mostly through the penalty opportunity cost. On the other hand, workers at large firms are paid partly in backpay, so those firms deliver less utility voluntarily. At the mean wage, a util costs the firm more to deliver than it is worth to the worker, so delivering less of it raises welfare by 2.2%.

Perfect enforcement lowers welfare by 15% of wages. Most of this loss is misallocation: enforcement's effect on the wedge lowers welfare by 9.7%, largely because formal firms at the minimum wage turn away 7.9 million workers into the informal sector. Workers gain slightly, since higher compliance and backpay offset lost jobs and lower wages, while profits fall. Perfect enforcement raises welfare only if workers' welfare is weighted 60 times as much as that of firm owners.

Importantly, welfare has an inverted-U shape in enforcement intensity: some enforcement improves welfare, but too much enforcement of these regulations can be costly. Welfare peaks at 0.4% of wages above no enforcement when complaining is 3 times cheaper than in the status quo. At the peak, firm choices barely differ from the status quo: formal employment and employment-weighted compliance move by less than half a percentage point, and no firm turns workers away. Relative to the status quo, the gain comes from what firms post and improved allocation of labor, despite a higher net cost of enforcement. Welfare falls below its no-enforcement level once complaining is 20 times cheaper than in the status quo. Ultimately, the aggregate effects of labor regulations depend on the enforcement regime, not just on the regulations themselves.

Contributions to the Literature

We contribute to several strands of the literature by estimating a model of labor regulations that incorporates two important features of labor markets in developing countries: imperfect enforcement and labor market power.

This paper is related to work on mandated benefits and how workers value them. When workers value labor regulations, wages absorb part of the cost of mandates (Summers, 1989; Gruber, 1994). Where inspections are frequent, workers pay for mandated benefits through lower wages (Almeida and Carneiro, 2012; Pignatti, 2026). We treat compliance as a firm choice and find that firms use it as compensation: workers pay for it through lower wages, and firms comply at similar rates even without enforcement. The value of non-wage job attributes is rarely observed directly, so the literature infers it from worker flows and earnings (Sorkin, 2018; Lamadon et al., 2022). In our data, the violations data record each regulated amenity separately, allowing us to estimate each amenity's value from how often it is reported. We find that each amenity is worth between 19% and 80% of the wage.

We also contribute to the literature on enforcement, firm size, and compliance. Several papers find that stricter enforcement of labor regulation constrains firm size (Samaniego de la Parra and Fernández Bujanda, 2024; Brotherhood et al., 2024; Bosch et al., 2026; Almeida and Carneiro, 2009), leads to exit (do Prado et al., 2025), and reduces worker wages (Gusti and Carvalho, 2024). Our paper provides a model of enforcement that explains the inspection-size relationship and how it affects firm choices. We find that enforcement can both increase and decrease the cost of hiring a worker, and that it can crowd out voluntary compliance. Kleven et al. (2016) microfound compliance and enforcement when employees collude with the firm. Feinmann et al. (2026) document such collusion in Brazil and find that a labor lawsuit by a former employee raises the reported wages of the firm's remaining workers by only 1%, well below full compliance. We complement this theoretical work and reduced-form evidence by modeling the complaint decision and the firm's compliance when workers are victims of noncompliance, not colluders. We find similarly weak deterrence at current enforcement levels, and we quantify how stronger enforcement would change firm size, misallocation, and welfare.

Several papers consider models of size-dependent regulations or enforcement in perfectly competitive markets and find that it can reduce aggregate productivity and increase informality (Guner et al., 2008; Garicano et al., 2016; Almunia and Lopez-Rodriguez, 2018;

Ulyssea, 2018; Bachas et al., 2019). Levy and Cruces (2021) argue that labor regulations hurt aggregate productivity. This paper extends this work to imperfectly competitive markets, where market power interacts with enforcement and makes aggregate outcomes theoretically ambiguous. We find that some enforcement raises welfare though too much enforcement of the current regulations does indeed lower welfare, so the aggregate cost of size-dependent enforcement depends on its intensity.

Finally, we contribute to the literature on labor market power and the misallocation of labor across firms. Dispersion in marginal products across firms, such as that coming from markdowns, lowers aggregate productivity (Hsieh and Klenow, 2009). Many papers find evidence of market power in labor markets across the developed and developing world (Berger et al., 2022; Brooks et al., 2021; Amodio and de Roux, 2024; Armangué-Jubert et al., 2025; Caldwell et al., 2026). The main policy this literature evaluates is the minimum wage, which recovers little of the efficiency lost to monopsony (Berger et al., 2025). We contribute to this literature by characterizing markdowns when firms offer a broad set of labor regulations covering compensation and amenities in a context with imperfect enforcement. We find that larger firms raise compliance along with wages, so employment rises much faster with the wage than the wage elasticity of labor supply alone implies. Attributing all of this response to wages would therefore understate the markdown. We also model the informal sector as workers' outside option, and markdowns contribute to an oversized informal sector (Maloney, 2004; Amodio et al., 2025; Felix, 2026). In addition, distortions from enforcement and markdowns interact, and in this second-best world ambiguous welfare effects (Lipsey and Lancaster, 1956; Bergquist et al., 2026). This paper quantifies the net effect of enforcing labor regulations in general equilibrium.

Paper Roadmap

The paper proceeds as follows. Section 2 describes Peru's labor regulations and the inspection process, and documents the two forces behind the inspection curve. Section 3 builds a general equilibrium model of monopsonistic firms where worker complaints drive enforcement. Section 4 enriches the enforcement process to map it to the inspection data and estimates the model in three blocks. Section 5 compares firm choices across enforcement regimes and reports empirical counterfactuals about compliance as compensation and bunching at the size threshold. Section 6 decomposes the welfare change from no enforcement, and reports its incidence on workers and firms. Section 7 concludes.

2 Labor Regulations and Inspections

We focus on Peru, where an extensive set of labor regulations mandates compensation and amenities for nearly all workers in the formal sector, but enforcement is imperfect. Limited capacity and information mean that enforcement depends largely on worker complaints. This section describes the regulations, the enforcement funnel that generates our data, and two facts about firm size and compliance that the model is built to explain.

2.1 *Labor Regulations and Enforcement in Peru*

Peru's labor regulations fall into two categories. Regulated compensation is made up of monetary benefits that firms must pay workers in addition to the wage, such as severance deposits and gratuities. Regulated amenities are non-monetary benefits, such as paid leave and workplace safety. Table 2 lists the eleven regulations we model. At full compliance the four compensation rates raise compensation 56% above the base wage. There is also a minimum wage, though it is low: 0.97 of the average earnings in the informal sector.

The regulations apply across the formal sector, with partial or full exemptions at revenue thresholds. Above 20 workers, firms also owe profit sharing and must meet additional health and safety requirements.[1] Firms in other contexts bunch below such thresholds, which is informative about the costs of compliance (Garicano et al., 2016; Harju et al., 2019). In our context, however, what a firm saves by staying small depends on how likely it is to be inspected, which dilutes the incentive to bunch. We return to this in section 5.3.

SUNAFIL, the national labor inspection agency, enforces these regulations. We obtain a novel dataset that contains the universe of inspected firms starting from the inception of the agency, their outcomes, and violation-level data for firms that are both inspected and sanctioned. Two-thirds of its inspections originate with a worker complaint, which are the focus of our analysis.[2] From 2014 to 2019 there are 22,087 first inspections, an unconditional inspection rate of 1.1% and a probability of a sanction upon inspection of 27%. However, these rates vary greatly with firm size, as we discuss below.

We study the legal documents and procedures behind SUNAFIL's records to build the

[1] Legislative Decree No. 892; Law No. 29783, Supreme Decree No. 005-2012-TR.

[2] The remaining third are planned inspections that mostly target informal employment rather than non-compliance at formal firms.

enforcement process we estimate in Block 2 (section 4.2.1). First, a worker complaint triggers an inspection. Workers can complain confidentially through an online portal, and SUNAFIL is required by law to respond to each complaint, or set of complaints, with an in-person inspection.[3] Second, the inspector investigates the regulations complained about and can expand the inspection to related violations. Third, the firm is given a short window to resolve the violations. Every violation must be addressed by restoring the owed amount to the worker (i.e., backpay), but a firm that resolves a violation within the window is not sanctioned. Finally, a violation that remains unresolved is sanctioned according to a legal schedule and recorded in the violations data, with the regulation violated, the number of workers shorted and the proposed sanction. The schedule depends on the severity of the violation, the size of the firm and the number of shorted workers, and realized sanctions follow it closely (section 4.1.2).

2.2 *Compliance, Complaints, and Size*

Figure 1 shows the pattern behind the inspection curve. The share of workers shorted on a recorded regulation, our measure of noncompliance, falls with firm size, so larger firms comply more. Kumler et al. (2020) also find that, in Mexico, compliance for payroll taxes increases with firm size. However, compliance does not rise fast enough to offset the larger number of workers, so the total number of violations in the workforce rises with size. Any shorted worker can complain, and more violations raise the chance that at least one of them does. The probability of a complaint-driven inspection therefore rises with firm size, a pattern that holds across developing countries (Almeida and Ronconi, 2016).[4] Our model builds the inspection probability from these two forces, so whether it rises or falls with size depends on which one dominates.

Another fact that motivates our aggregate analysis is the pervasiveness of noncompliance. Approximately 45% of formal workers are employed at firms that have been complained about, and the violations data alone cover 4.7% of the formal workforce. This is a lower bound, since uninspected firms may also be noncompliant. If we naively assume that the share of workers shorted in the violations data holds at every firm, 14% of formal workers are

[3]Beginning in 2020, SUNAFIL responded to some complaints with letters instead, which breaks the link between inspections and complaints. We therefore use data from 2014 to 2019.

[4]The probability of a sanction given an inspection also reflects two forces: more shorted workers make resolution within the window harder, while larger firms have more administrative capacity to resolve violations. On net it is non-monotonic in firm size (fig. 15 in appendix C).

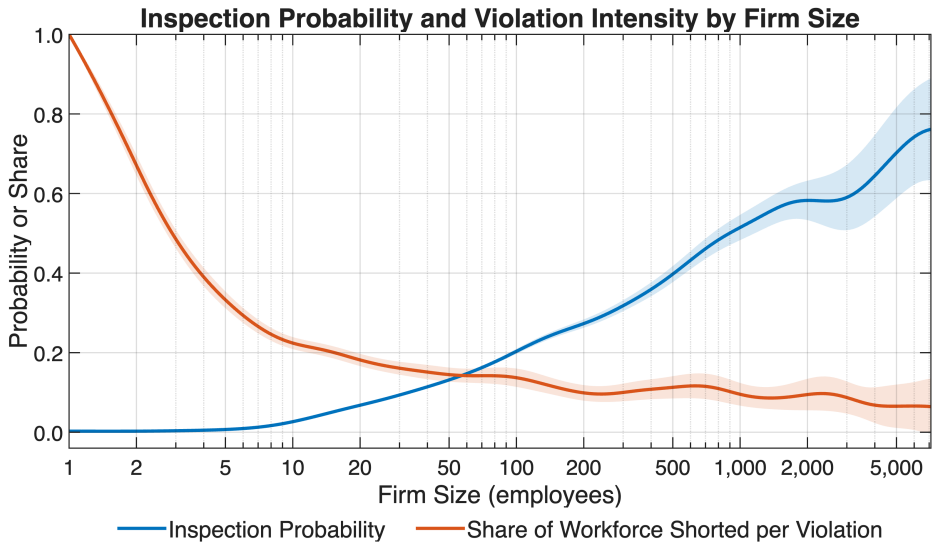


Figure 1: The two forces behind the inspection curve. Both series are local-linear fits over log head count on the first-inspection panel, with 95% confidence bands. The share of workers shorted per recorded regulation falls with firm size, while the probability of a first inspection rises.

shorted on a given regulation. Compliance choices are therefore welfare relevant, especially to the extent that they reflect market power rather than a transfer from workers to firms.

However, the violations data are a selected sample. A firm enters them only if a worker complains and the firm then fails to resolve what the inspection surfaces. Figure 2 shows the resulting tilt in the estimated model (section 4): most formal firms have fewer than ten workers, but the inspected firms spread across the whole size distribution. Because compliance rises with size, compliance measured in the violations data overstates compliance in the population, and the Block 2 likelihood (section 4.3) corrects for this selection.

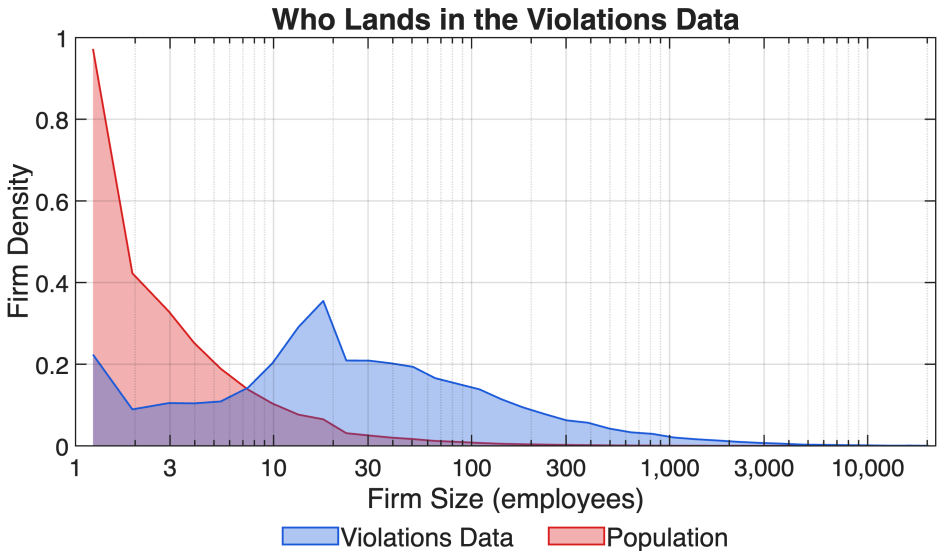


Figure 2: Who lands in the violations data, in the estimated model. The red area is the density of formal firms over log firm size; the blue area weights each firm by its probability of an inspection, the first gate into the violations data. Both are densities per unit of log size.

3 Model

We build a standard model of monopsony in general equilibrium and incorporate endogenous enforcement of labor regulations. First, we model worker utility received in the labor market, which includes the utility delivered by the firm and the backpay secured through complaint-driven enforcement. Then, we write the firm problem, where the firm chooses its compliance and the wage it offers to workers, anticipating its expected penalties and backpay. We then derive the firm's optimality conditions under each enforcement regime.

3.1 *Worker Preferences*

Workers value both the compensation and amenities they receive from their employer, and they also value the backpay they can secure through complaint-driven enforcement. Additionally, workers have idiosyncratic preferences over firms and the informal sector, which generates upward-sloping labor supply curves to each firm.

3.1.1 *Labor Market*

Firms compete for workers in a labor market with imperfect competition driven by idiosyncratic preferences over firms. We assume a standard logit structure for these preferences, which gives a closed form for the labor supply to each firm (Card et al., 2018). The labor market is populated by L workers, each of whom chooses to work for one of the formal firms $\mathcal{J}$ or in the informal sector.

The utility of working for firm j is given by:

$$u_{ij} = V_j + \varepsilon_{ij} \tag{1}$$

where V_j is the expected value of firm j and ε_{ij} is the idiosyncratic component to worker i's utility from working at firm j. These preferences assume a logit structure with scale normalized to one, in which workers choose among the formal firms and the informal sector is the outside option. This gives us a standard logit model of labor supply derived in appendix A.8 where

$$\ln N_j = V_j - IV + \ln L, \qquad IV \;=\; \ln\Big(\exp(V_0) + \sum_{j\in\mathcal{J}} \exp(V_j)\Big) \tag{2}$$

where IV is the aggregate market index of the entire labor market.[5]

3.1.2 *Regulatory Structure and Complaints*

There are a set of labor regulations $\mathcal{K}$ that govern the compensation and amenities that firms must provide to their workers. Regulations are split into two categories: regulated compensation (e.g. severance deposits, gratuities, pensions) and regulated amenities (e.g. rest and leave, workplace safety, labor formalities). The firm honors its obligations with probability $\lambda \in [0, 1]$, which is the firm's compliance choice. We discuss why compliance is probabilistic and its microfoundation in section 3.2.2.

The set of regulated compensation is given by $\mathcal{K}^w$, each providing a multiplicative bonus of δ_k^w on the base wage, so that the promised compensation is $w\prod_k(1+\delta_k^w)$. The set of regulated amenities is given by $\mathcal{K}^a$, each modeled as discrete benefit unit offering a fixed utility to the worker. The full set of regulations is $\mathcal{K} := \mathcal{K}^w \cup \mathcal{K}^a$, and the utility of each

[5]Firms choose the value V_j in order to attract workers through the labor supply response captured by the elasticity $\eta_j := \frac{\partial \ln N_j}{\partial V_j} = 1 - \frac{N_j}{L}$, where N_j/L is the share of workers who choose firm j. Even the largest firms in the economy are small relative to the labor market, so $N_j/L \approx 0$ and $\eta_j \approx 1$.

regulation is given by Δ_k.[6]

We say that a worker is shorted on regulation k if the firm fails to honor it, given by the indicator $S_{ik} \in \{0, 1\}$, which occurs with probability:

$$\mathbb{P}(\text{shorted}) = 1 - \lambda \qquad \text{i.i.d. across workers and regulations}$$

If workers are shorted on regulation k, they have the ability to complain to the labor inspection agency to receive backpay. A complaint results in an inspection that launches a firm-wide investigation of violations of regulation k and uncovers related violations. The inspection requires firms to restore the shorted workers, which delivers backpay for investigated violation k: workers receive the *backpay on* k, the utility Δ_k they were owed on regulation k.

The worker shorted on regulation k at a firm of size ℓ draws an idiosyncratic cost of complaining

$$\omega_{ik} \overset{\text{iid}}{\sim} \text{Exp}\big(\rho(\ell)\big), \qquad \rho \;=\; \rho_0\, \ell^{\rho_1} \tag{3}$$

When $\rho_1 > 0$, the cost of complaining is lower at larger firms, which captures a variety of reasons why workers may be more willing to complain at larger firms (e.g. lower fear of retaliation, lower dynamic concerns). We drop the argument ℓ in $\rho(\ell)$ for notational simplicity.

A worker complains only when the utility on k exceeds their idiosyncratic cost of complaining,

$$\Delta_k > \omega_{ik}.$$

Workers do not complain strategically or cooperatively – they do not consider their coworker's complaint behavior when deciding whether to complain. This is a behavioral-primitive assumption that buys tractability and is common in models of expressive choice (e.g., Schuessler, 2000). Additionally, workers do not discount backpay utility when deciding whether to complain, only when choosing a firm.

3.1.3 *Worker Utility and Backpay Value*

A firm determines V_j by offering a base wage w_j and a compliance probability λ_j. A worker's utility combines wages, the utility from regulated compensation and amenities, and the ex-

[6]The set $\mathcal{K}^w = \{1, \ldots, K^w\}$ contains the K^w regulations that govern compensation, and $\mathcal{K}^a = \{K^w + 1, \ldots, K^w + K^a\}$ contains the K^a regulations that govern amenities.

pected utility they receive from complaining (and the subsequent inspection), which we call backpay.

$$V(w_j, \lambda_j, \ell_j) = \underbrace{\beta \ln(w_j) + \lambda \sum_{k \in \mathcal{K}} \Delta_k}_{W_j} + \vartheta\, B(\lambda, \ell), \tag{4}$$

The first two terms are the expected *delivered utility*, offered voluntarily by the firm

$$W_j := \beta \ln w + \lambda \sum_{k \in \mathcal{K}} \Delta_k \tag{5}$$

where Δ_k is the utility of regulation k and λ is the probability that the firm honors it. For a regulated amenity $\Delta_k = \beta_k^a$, and for regulated compensation $\Delta_k = \beta \ln(1 + \delta_k^w)$, because the premium multiplies the base wage and utility is log in compensation (appendix A.7).

While workers complain non-strategically, we allow them to be *sophisticated* about their complaint behavior when choosing firms. Though they have realized neither their shocks S_{ik} nor their complaint costs $\{\omega_{ik}\}$, they know the distributions and anticipate the consequences of their behavioral complaint rule.[7]

The backpay value $B(\lambda, \ell)$ is the expected restored utility upon complaint and inspection net of the expected complaint cost. A shorted worker is made whole only if an inspection investigates the regulation, which happens when the worker, a coworker, or a complaint on a related regulation brings it to the inspector. Against this the worker pays the complaint cost whenever they complain. We give the closed form of B in appendix A.3.

The backpay value falls with compliance λ for two reasons. A worker is shorted less often, so there is less to restore, and a shortfall that does occur is less likely to be investigated, because fewer coworkers are shorted and complain. On the other hand, the backpay value increases with firm size ℓ because a larger firm has more coworkers who might complain, so a shortfall is more likely to be investigated.[8] The value V therefore depends on ℓ, and a firm that hires one more worker changes the backpay it is expected to deliver.

The backpay weight ϑ governs the extent to which workers discount the backpay value

[7] This is the sophisticated-agent device of O'Donoghue and Rabin (1999): behavior is governed by a present-focused rule (the non-strategic impulse to complain about being shorted), while the forward-looking self foresees that rule and prices its consequences correctly rather than imagining it will act differently.

[8] The expected complaint cost moves with size in either direction, since cheaper complaints mean that more workers complain but that each pays less.

when choosing firms: $\vartheta = 1$ means that workers fully capitalize it, while $\vartheta = 0$ means that they ignore it entirely. Note that workers do not discount backpay when deciding whether to complain, only when choosing a firm.

3.2 *The Production Environment*

The formal firm internalizes how its compliance and size move worker complaints, and hence its expected enforcement. These choices also move the value workers receive and so the labor supply the firm faces. Enforcement therefore both shrinks and expands employment, unlike standard models in which enforcement is an exogenous size-dependent tax (e.g. Ulyssea, 2018).

3.2.1 *Formal Firm Problem*

Firm j has production with heterogeneous productivity $z_j \sim$ Pareto and diminishing returns to labor:

$$y_j \;=\; z_j\, \ell_j^{\alpha}, \qquad 0 < \alpha < 1. \tag{6}$$

We treat y_j as value added, so intermediate inputs are already netted out, and we do not model capital separately. The profit $\pi(z_j)$ in the firm's problem below is therefore operating surplus, which includes the return to capital. The product market is perfectly competitive, with price normalized to 1.

The firm chooses a base wage w and a compliance rate λ, which together determine the expected value V_j and hence the labor supply N_j it receives. The firm either locates itself on the labor supply curve ($\ell_j = N_j$) or it rations hiring ($\ell < N_j$). The firm faces a per-worker cost function $c(w, \lambda)$ and an expected penalty $\mathcal{P}(w, \lambda, \ell)$ for being inspected and sanctioned. The firm's problem is therefore:

$$\pi(z_j) = \max_{w,\lambda,\ell}\; z_j \ell^{\alpha} - c(w,\lambda)\ell - \mathcal{P}(w,\lambda,\ell), \qquad \text{s.t.} \quad \begin{array}{l} \ell \le N_j(w,\lambda,\ell) \\ w \ge w^r \end{array} \tag{7}$$

where $w^r > 0$ is the minimum wage. The labor supply constraint is given by eq. (2) and is a fixed point because a larger firm is inspected more often, which raises the backpay value, which raises worker value, which raises employment again.

Firm costs are split into two components. The first is the cost of delivered compensation

and amenities: the compensation and amenity bill multiplied by the probability that the firm honors it. The second component is the cost of compliance: the owed compensation and amenity bill $\mathsf{b}(w)$ multiplied by the cost of moving up the opportunity cost curve $\mathcal{X}(\lambda)$. The per-worker cost function is therefore

$$c(w,\lambda) \;:=\; \underbrace{\Lambda(\lambda)\,w \;+\; c^a K^a \lambda}_{\text{delivery}} \;+\; \underbrace{\mathsf{b}(w)\,\mathcal{X}(\lambda)}_{\text{compliance cost}} \tag{8}$$

The expected penalty has two pieces: expected worker backpay and expected sanctions. The first is the cost to make every shorted worker whole on every regulation an inspection investigates, $\mathcal{P}^{\text{bp}}$. The second is the sanction the firm pays on the violations it fails to resolve in time, $\mathcal{P}^{\text{s}}$. The firm pays both at the inspection date, so both carry the opportunity cost of not being able to delay the payment, $\mathcal{X}(1)$. The expected penalty is therefore

$$\mathcal{P}(w,\lambda,\ell) \;=\; \big(1+\mathcal{X}(1)\big)\Big[\mathcal{P}^{\text{bp}}(w,\lambda,\ell) \;+\; \mathcal{P}^{\text{s}}(\lambda,\ell)\Big] \tag{9}$$

We discuss these in further detail below.

3.2.2 *Cost of Delivery*

The delivered compensation is the base wage multiplied by the expected multiplicative premium a worker receives on top of the base wage:

$$\Lambda(\lambda) = \prod_{k\in\mathcal{K}^w} \big[1+\delta_k^w \lambda\big] \quad \geq 1. \tag{10}$$

The factor is a *product*, not a sum: regulated compensation compounds on the base wage exactly as eq. (81). With complete noncompliance, $\Lambda(0) = 1$, and the firm pays only the base wage. The delivered amenity cost is a product of three terms: the number of amenity units to deliver $\mathcal{K}^a$, the probability of delivering each unit λ, and the per-unit cost c^a.

3.2.3 *Cost of Compliance*

A firm choice of compliance comes with an opportunity cost. Obligations come due at i.i.d. draws of a *shadow opportunity cost* $q \sim F$ on $[1,\infty)$: paying one dollar in state q costs the transfer plus $q-1$ of lost production. This can represent a variety of frictions, such as the cost

of raising cash to pay regulated compensation or the cost of delaying production to deliver regulated amenities.

The firm chooses to honor its obligations only in the cheapest states: it pays if and only if $q \leq q^*$, which gives a compliance $\lambda = F(q^*)$ (see appendix A.6). Thus, compliance is a probability rather than an intensity. We assume that the base wage is paid costlessly, so the cost falls only on the promised obligations, and the *compliance cost* is

$$\underbrace{w\big(\Lambda(1)-1\big)+c^aK^a}_{:=\mathsf{b}(w)} \times \underbrace{\int_0^\lambda \big(F^{-1}(u)-1\big)\,du}_{:=\mathcal{X}(\lambda)}. \tag{11}$$

The first term is the total cost of the promised obligations, where $w(\Lambda(1)-1)$ is the total promised compensation (minus the base wage) and c^aK^a is the total promised amenities.

The second term adds up the opportunity cost of compliance across the states in which the firm honors its obligations. The marginal cost of compliance is $\mathcal{X}'(\lambda) = F^{-1}(\lambda) - 1$, which rises with λ because the firm must pay in more expensive states. Thus, $\mathcal{X}$ is convex for any F. We assume that $q-1$ is exponential with mean ϕ, which gives

$$\mathcal{X}'(\lambda) \;=\; -\phi\ln(1-\lambda), \qquad \mathcal{X}(\lambda) \;=\; \phi\big[(1-\lambda)\ln(1-\lambda)+\lambda\big], \qquad \mathcal{X}_1 := \mathcal{X}(1) \;=\; \phi, \tag{12}$$

The marginal cost rises without bound as $\lambda \to 1$, so an interior compliance rate always exists.[9]

3.2.4 *Expected Backpay Cost*

Once an inspection investigates regulation k, the government makes the firm restore *every* worker shorted on k, not just those who complained. The backpay transfer $\mathcal{P}^{\text{bp}}(w,\lambda,\ell)$ is the expected value of this restoration: across its ℓ workers, the regulated compensation and amenities the firm owed but did not deliver, on the regulations an inspection investigates. We give its closed form in appendix A.5.

The backpay transfer increases in the base wage w, since regulated compensation is a percent bonus of the base wage. Thus, a higher wage means that firms owe more in backpay.

[9] A corner solution at $\lambda = 1$ would reduce the inspection probability of that firm to zero, which would be inconsistent with the observed inspection data. Avoiding this corner solution improves tractability and is consistent with the data, so we impose it as a restriction.

Higher compliance λ lowers the transfer on two margins: the firm shorts fewer workers, and a shortfall that does occur is less likely to be investigated, because fewer coworkers are shorted and complain.[10] A larger firm ℓ pays more, both because it has more workers to restore and because each shortfall is more likely to be investigated.

During ordinary production, the firm chooses to pay obligations in the cheapest states, leading to the opportunity cost of compliance given by $\mathcal{X}(\lambda)$ in eq. (12). However, when a firm is inspected, the government forces it to pay its obligations in whatever state prevails at the inspection date. This results in the firm paying at the *average* opportunity cost $\mathcal{X}_1 := \mathcal{X}(1)$, not $\mathcal{X}(\lambda)$. The gap between the two rates is the option value of timing: voluntary compliance chooses its states, backpay cannot. A sanction falls due on the same date and is paid the same way. The firm's outlay on its penalties is therefore

$$\big(1 + \mathcal{X}_1\big)\big(\mathcal{P}^{\text{bp}} + \mathcal{P}^{\text{s}}\big) \tag{13}$$

with $\mathcal{P}^{\text{bp}}$ the backpay transfer and $\mathcal{P}^{\text{s}}$ the sanction. The backpay transfer moves utility to workers and the sanction is revenue to the government, but the *penalty opportunity cost* $\mathcal{X}_1\big(\mathcal{P}^{\text{bp}} + \mathcal{P}^{\text{s}}\big)$ is a real resource cost that accrues to nobody.

3.2.5 *Expected Sanctions*

Sanctions in Peru follow a published schedule that rises with the number of shorted workers, the severity of the violation and the size of the firm. We parametrize that schedule as

$$\mathsf{S}_k(\lambda, \ell) \;=\; \ell^{\zeta}\, g_k \left(1 - e^{-\nu n_k}\right), \tag{14}$$

with severity $g_k > 0$, intensive-margin curvature $\nu > 0$, and size scaling ζ. The firm-wide count of workers shorted on regulation k is

$$n_k \;=\; \sum_{i=1}^{\ell} S_{ik} \;\sim\; \text{Binomial}(\ell, 1 - \lambda), \tag{15}$$

The expected sanction $\mathcal{P}^{\text{s}}(\lambda, \ell)$ is the expected sum of S_k over the regulations the firm

[10] Regulated compensation compounds on the base wage, so honoring one premium raises the value of restoring the others. Lemma 2 shows that the two margins dominate whenever the compensation rates sum to at most one, as the statutory rates do (0.47).

is sanctioned on (we describe the process leading to sanction in detail in section 4.2.3). Appendix A.2 derives it in closed form.

3.2.6 *The Informal Sector*

The informal sector consists of a continuum of identical, infinitesimally small firms, represented as a single price-taking firm with linear production $y_0 = z_0 \ell_0$. The informal sector lies entirely outside the regulatory system: it makes no compliance investment ($\lambda_0 = 0$), there are no complaints, inspections, or backpay. Worker value in eq. (4) therefore collapses to the utility of the base wage, plus an exogenous taste ξ_0 for the informal sector:

$$V_0 \;=\; \beta \ln w_0 \;+\; \xi_0. \tag{16}$$

The informal firm earns zero profits, so $z_0\,\ell_0 \;=\; w_0\,\ell_0$, and the taste ξ_0 functions as a residual that closes the labor market.[11]

3.3 *Equilibrium*

Equilibrium is defined by firms choosing their optimal choice of wage, compliance, and size; workers maximize utility and sort into firms; the labor market clears. More precisely, an equilibrium is a set of firm choices $\{w_j, \lambda_j, \ell_j\}_{j \in \mathcal{J}}$ and worker sorting $\{j(i)\}$ such that:

Firm j maximizes its objective given the choices of other firms.

$$(w_j, \lambda_j, \ell_j) \in \arg\max_{w,\lambda,\ell} \; z_j \ell^{\alpha} - c(w,\lambda)\ell - \mathcal{P}(w,\lambda,\ell), \quad \text{s.t.} \quad \ell \le N_j(w,\lambda,\ell),\; w \ge w^r.$$

Workers maximize utility and sort into firms.

$$j(i) \in \arg\max_{j \in \mathcal{J}} V_j + \varepsilon_{ij} \Rightarrow \ln N_j = V_j - IV + \ln L.$$

Labor market clears.

$$\sum_{j \in \mathcal{J}} \ell_j + \ell_0 = L$$

[11] We assume for now that formal firms cannot hire informal workers, though this assumption can be relaxed.

3.4 *Counterfactual Enforcement Regimes*

We call the current complaint-based enforcement, in which an inspection happens only when a worker complains, the *status quo*. We compare it with two counterfactual regimes. Under *no enforcement*, no worker ever complains: firms are never inspected and never pay backpay or sanctions, so they comply only to attract workers. Under *perfect enforcement*, every shorted worker complains: any shortfall triggers an inspection, every shorted worker is restored in full, and the firm is sanctioned on the violations it fails to resolve in time. Workers then receive every obligation, from the firm directly or as backpay, so firm size no longer changes the value of a job through enforcement. Both regimes are limits of the status quo in the cost of complaining. No enforcement sends the complaint cost scale ρ_0 of eq. (3) to zero, which makes complaining infinitely costly, and perfect enforcement sends it to infinity, which makes complaining free. We derive the limit of every enforcement object in appendix A.10.

3.5 *The Firm's Optimality Conditions*

The firm problem in eq. (7) has the same form in every regime, and only the penalty $\mathcal{P}$ and the backpay value B change across them. Let $\mu \geq 0$ be the multiplier on the labor supply constraint. When the wage is above the minimum, the first-order conditions reduce to two conditions, which we derive in appendix A.10 together with the case of a firm that rations. The first sets the mix of compliance and the wage,

$$\underbrace{\frac{c_\lambda + \mathcal{P}_\lambda/\ell}{D + \vartheta B_\lambda}}_{\text{cost of a util through compliance}} = \underbrace{\frac{c_w + \mathcal{P}_w/\ell}{\beta/w}}_{\text{cost of a util through the wage}} = \eta\mu, \tag{17}$$

where $D := \sum_{k \in \mathcal{K}} \Delta_k$ is the utility of every regulation honored, so that the delivered utility of eq. (5) rises by D per unit of compliance. The firm equates the cost of a util across its two instruments, and that common cost is $\eta\mu$. The second sets its size,

$$\underbrace{\alpha z_j \ell^{\alpha-1}}_{MRPL} = \underbrace{c(w,\lambda)}_{\text{per-worker cost}} + \underbrace{\mathcal{P}_\ell}_{\text{size penalty}} + \underbrace{\mu}_{\text{recruiting cost}} - \underbrace{\eta\mu\,\vartheta\,\ell\,B_\ell}_{\substack{\text{value delivered} \\ \text{by inspection}}}. \tag{18}$$

To hire one more worker the firm must raise the value of the job for all of its workers, and we call this cost the *recruiting cost* μ. Since $\eta \approx 1$, the cost of a util is close to the recruiting

cost. A larger firm is also inspected more often, so part of that value is delivered by backpay.

Without enforcement $\mathcal{P} = 0$ and $B = 0$. A firm still complies with regulations, but only to attract workers, and the recruiting cost μ, the gap between the marginal product and the per-worker cost, is the monopsony markdown. With enforcement, μ is the markdown plus the wage penalty of section 6.1.2. Enforcement appears in three terms. The penalty saved lowers the cost of compliance, and wage backpay raises the cost of the wage, so both push the firm toward compliance. The backpay value lost lowers the return to compliance and pushes the other way, so the net effect is ambiguous in sign (section 5.1.1). On the size margin, enforcement adds the size penalty and the value delivered by inspection, which also push in opposite directions. Under perfect enforcement every shortfall is restored, so the backpay value $(1-\lambda)D$ no longer depends on size and the last term of eq. (18) drops out. The delivery cost of compliance then cancels against the backpay it saves, because a firm that shorts a worker pays them anyway, and compliance trades its opportunity cost against the penalty opportunity cost and the sanction it saves.

4 Estimation

We quantify the model above by using the structure of the model to match the observed data. Estimation proceeds in three recursive blocks. *Block 1* sets the statutory compensation rates from the law, fits the sanction function in the model to the published schedule and observed sanctions, and sets the economic aggregates from household and administrative data. *Block 2* enriches the model's enforcement process so that it maps to the inspection and violations data, and estimates it by maximum likelihood on the full distribution of inspection outcome and violation data. *Block 3* estimates the production technology and labor market parameters by simulated method of moments. Each block opens with a table of its parameters, and table 12 in appendix C describes how each is set.

4.1 *Block 1: Statutory and Institutional Parameters*

In Block 1, we use statutory values to set the number of regulations on compensation and amenities, the compensation rates, and the minimum wage. We then estimate the sanction function from the 2016 sanction records, verifying that the schedule is followed in practice. Finally, we measure the scale of the economy in 2016.

Parameter	Model object	How it is set
$K^w,\ K^a$	Regulation counts by type	Violations data classification
$\{\delta_k^w\}$	Compensation rate of each regulation	Statutory rates
$\{g_k\},\ \nu,\ \zeta$	Sanction schedule S_k of eq. (14)	NLS on 2016 sanction records
$L,\ J$	Workforce, formal firms	ENAHO 2016; employer–employee panel
N_F	Formal employment	Employer–employee panel
w_0	Informal wage	ENAHO 2016
w^r	Minimum wage	Statutory value

Table 1: Block 1 parameters, set directly from legal texts and administrative data.

4.1.1 *Regulations and Compensation Rates*

Table 1 lists the Block 1 parameters and how each is set.

We model eleven regulations, which together cover 90% of the records in the violations data and 97% of the shorted-worker burden. We set the multiplicative bonus δ_k^w of the regulations on compensation to the fraction of the monthly wage set by law. Table 2 lists the regulations and their rates, and we explain in appendix C which categories we leave out and why. At full compliance the four rates raise compensation 56% above the base wage, so the compensation that these regulations add is large relative to the wage. The minimum wage w^r is the 2016 statutory minimum averaged over the year, S/ 817.

4.1.2 *Sanction Function*

We fit the sanction schedule of eq. (14) by nonlinear least squares on the 2016 sanction records, with a separate level for each severity class, and a regulation's severity g_k interpolates between these levels (table 2). Sanctions rise far less than proportionally with firm size ($\hat{\zeta} = 0.15$) and flatten after the first few shorted workers ($\hat{\nu} = 0.48$), and realized sanctions follow the legal schedule closely, so firms can anticipate the sanction they would pay on a recorded violation (appendix C gives the details). Most importantly, expected sanctions are low. Multiplying the mean fine per sanctioned inspection by the inspection rate and the probability of a sanction given an inspection gives an expected sanction of about 0.8% of one month's payroll a year both at firms with at most ten workers and at firms above one hundred workers.

	Regulation	Set	Group	δ_k^w	$\bar{h}_k$	g_k (S/)
1	Severance deposit (CTS)	$\mathcal{K}^w$	$\mathcal{R}^1$	1/12	2.002	2,445
2	Gratuities	$\mathcal{K}^w$	$\mathcal{R}^1$	2/12	2.001	2,443
3	Health insurance	$\mathcal{K}^w$	$\mathcal{R}^2$	0.09	2.896	3,914
4	Pension	$\mathcal{K}^w$	$\mathcal{R}^2$	0.13	2.779	3,721
5	Wage timeliness	$\mathcal{K}^a$	$\mathcal{R}^1$		2.005	2,449
6	Working time, rest and leave	$\mathcal{K}^a$	$\mathcal{R}^1$		2.778	3,720
7	Fundamental rights	$\mathcal{K}^a$	$\mathcal{R}^4$		2.777	3,719
8	Workplace safety and health	$\mathcal{K}^a$	$\mathcal{R}^3$		2.627	3,472
9	Safety and health management and records	$\mathcal{K}^a$	$\mathcal{R}^3$		2.215	2,795
10	Labor formalities	$\mathcal{K}^a$	$\mathcal{R}^0$		2.115	2,630
11	Social security and indirect benefits	$\mathcal{K}^a$	$\mathcal{R}^5$		1.588	1,601

Table 2: The eleven modeled regulations. δ_k^w is the compensation rate as a fraction of the monthly wage, $\bar{h}_k$ the mean severity class of the regulation's recorded violations (1 Minor, 2 Serious, 3 Very Serious), and g_k its severity in 2016 soles under eq. (120). The group column gives the related group of section 4.2.2, with $\mathcal{R}^0$ the always-checked set, as estimated in Block 2 (section 4.3).

4.1.3 *Scale of the Economy*

We use the 2016 national household survey and the employer-employee panel to set the scale of the economy: total workforce L, formal employment N_F, and the number of formal firms J. The informal wage w_0 is mean monthly informal earnings in the household survey, S/ 843. We use w_0 as the numeraire and divide every monetary object in the model by it. The minimum wage w^r is S/ 817 a month, or 0.97 of the informal wage. The floor sits just below informal earnings, and it binds at the bottom of the formal size distribution, where the smallest firms pay the minimum wage all year.

Aggregate	Symbol	Value	Source
Total employment	L	14,719,456	ENAHO 2016, survey-weighted
Formal firms	J	332,579	Employer–employee panel, 2016
Formal employment	N_F	3,464,773	Employer–employee panel, 2016
Formal share	N_F/L	0.24	Derived

4.2 *Block 2 Theory: The Complaint and Inspection Process*

In Block 2 we estimate the complaint and inspection process by maximum likelihood, holding the Block 1 parameters fixed. First, we enrich the model's enforcement process so that it

maps to the inspection and violations process in Peru, and we derive the likelihood of each inspection outcome. Then, section 4.3 takes this likelihood to the data and reports the estimates.

4.2.1 *Complaints and Inspections*

The model of section 3 summarizes the enforcement process through the escape probability, the expected sanction and the inspection probability. To map the model to the inspection data, we enrich this process with the contextual details of section 2, which lets us recover firms' compliance and workers' complaint behavior from the Peruvian inspection data. Non-compliance leads workers to complain, complaints trigger inspections that surface a set of regulations, and the regulations the firm fails to resolve within a short window are sanctioned and recorded in the violations data. The regulations, compliance and complaint costs are those of section 3.1, and the shorted count n_k is that of eq. (15).

A shorted worker complains when $\Delta_k > \omega_{ik}$, with the complaint cost ω_{ik} of eq. (3). The silence probability given a shortfall, m_k, and the unconditional probability of silence, ψ_k, are therefore

$$m_k = e^{-\rho \Delta_k}, \qquad \psi_k(\lambda) \;:=\; \underbrace{\lambda}_{\text{not shorted}} \;+\; \underbrace{(1-\lambda)\, m_k}_{\text{shorted and silent}} \;. \tag{19}$$

Let $\mathcal{T}$ be the set of regulations that are complained about across the workforce. Regulation k generates no complaint if and only if every one of the ℓ workers is silent on k, which occurs with probability

$$a_k(\lambda, \ell) \;:=\; \mathbb{P}(k \notin \mathcal{T}) \;=\; \psi_k^{\ell}, \tag{20}$$

where ℓ is the number of workers. Inspection probability is therefore the probability that at least one regulation generates a complaint, which is

$$\mathscr{I}(\lambda, \ell) := \mathbb{P}(\mathcal{T} \neq \emptyset) = 1 - \prod_{k \in \mathcal{K}} a_k. \tag{21}$$

Thus, in a similar spirit to Kleven et al. (2016), we build the inspection probability from the complaint behavior of shorted workers, which depends on the compliance λ and the number of workers ℓ.

4.2.2 *The Set of Investigated Violations*

An inspector investigates the regulations complained about and also regulations that get *surfaced* during the inspection. We partition the set of regulations $\mathcal{K}$ into M disjoint *related groups* $\mathcal{R}^1, \ldots, \mathcal{R}^M$ and an *always-checked set* $\mathcal{R}^0$. For $k \in \mathcal{R}^m$, the set of regulations investigated upon a complaint on k is $\mathcal{R}_k = \mathcal{R}^m \cup \mathcal{R}^0$.[12] Let $\mathcal{R}$ be the *surfaced set*, the set of regulations that are investigated in an inspection and on which at least one worker is shorted:

$$\mathcal{R} \;=\; \Big\{k \in \bigcup_{j \in \mathcal{T}} \mathcal{R}_j : n_k \geq 1\Big\} \tag{22}$$

Let $\mathcal{D}_k := \{j \neq k : k \in \mathcal{R}_j\}$ be the set of other regulations whose complaint would pull k into the investigation. It runs in the opposite direction from $\mathcal{R}_k$, the set of regulations investigated following a complaint on k, and the two differ whenever $\mathcal{R}^0$ is non-empty.

A shorted worker-regulation pair is made whole if and only if its regulation is surfaced, which happens when a complaint lands on some regulation whose set pulls k in. Conditional on the pair being shorted, the probability that it is not surfaced is the *escape probability*

$$\mathrm{esc}_k(\lambda, \ell) \;:=\; \mathbb{P}\big(k \notin \mathcal{R} \mid S_{ik} = 1\big) \;=\; m_k\, \psi_k^{\ell-1} \prod_{j \in \mathcal{D}_k} a_j, \tag{23}$$

derived in appendix A.4. A regulation k escapes investigation (conditional on being shorted) if: the shorted worker themself does not complain, no other worker shorted on k complains, no complaint arrives on any other regulation that would pull k into the investigation. Size enters through the exponent $\ell - 1$, each $a_j = \psi_j^{\ell}$, and the size-dependent complaint cost of eq. (3). Escape shrinks with size on all three channels, so a shortfall at a large firm is far more likely to be restored than the same shortfall at a small one.

[12] A complaint on a regulation of the always-checked set opens only that set: $\mathcal{R}_k = \mathcal{R}^0$ for $k \in \mathcal{R}^0$. Every investigation therefore covers $\mathcal{R}^0$.The always-checked set holds the regulations an inspector verifies on every visit, such as the firm's labor registration. Setting $M = 1$ and $\mathcal{R}^0 = \emptyset$ recovers the case in which any complaint audits every regulation.

4.2.3 *Sanction Probability and Recorded Violations*

When the firm is inspected, the government offers a fixed compliance window in which to resolve the surfaced violations and avoid sanctions.[13] Firms resolve regulations in random order, until the cumulative burden of shorted workers exceeds their resolution capacity, the amount of shorted workers they can make whole in the time window. The order of resolution does not touch backpay: a surfaced regulation is restored whether it is resolved in-window or after sanction.

Firms draw their resolution capacity according to

$$\mathcal{C} \;\sim\; \text{Exponential}\big(\gamma(\ell)\big), \qquad \gamma(\ell) \;:=\; \gamma_0\, \ell^{-\gamma_1} \tag{24}$$

where $\gamma_0 > 0$ governs the difficulty of rapidly resolving a given shorted-worker burden and $\gamma_1 \geq 0$ determines the scale effect of firm size on resolution capacity. Firms resolve regulations until the cumulative burden of shorted workers $\mathcal{N} := \sum_{k \in \mathcal{R}} n_k$ exceeds their resolution capacity $\mathcal{C}$. The set of regulations that are unresolved and sanctioned is the *recorded set* $\mathcal{A} \subseteq \mathcal{R}$. Formally, we index the surfaced set $\mathcal{R}$ by a random permutation $r : \mathcal{R} \to \{1, \ldots, |\mathcal{R}|\}$, where $r(k)$ is the position of regulation k in the random order. Define the first position r^* at which the cumulative burden exceeds the resolution capacity:

$$r^* \;=\; \min\Big\{ r(k) : \sum_{\substack{j \in \mathcal{R}: \\ r(j) \leq r(k)}} n_j \;>\; \mathcal{C} \Big\}. \tag{25}$$

The recorded set is then

$$\mathcal{A} \;=\; \big\{ k \in \mathcal{R} :\; r(k) \geq r^* \big\}. \tag{26}$$

The probability that a firm resolves all of its surfaced burden within the capacity window, and hence avoids sanction completely, is

$$\sigma(\mathcal{N}, \ell) \;:=\; \mathbb{P}(\mathcal{C} \geq \mathcal{N}) \;=\; e^{-\gamma(\ell)\,\mathcal{N}}, \qquad \mathcal{N} := \sum_{k \in \mathcal{R}} n_k \tag{27}$$

We now collect the assumptions of the enforcement process and state the likelihood of

[13] This layer is necessary to reflect the empirical fact that many firms let off without a sanction, and is explicitly outlined in the law. Only firms that are sanctioned enter the violation records, and hence the observable data.

a firm's observed outcome: not inspected, inspected with no sanction, or inspected and sanctioned, indexed by $d \in \{1, 2, 3\}$, respectively. The proof of proposition 1 is given in appendix A.1 and uses lemma 1.

Proposition 1 (Likelihood of the Recorded Set)**.** *Suppose that*

(i) each worker is shorted on each regulation independently with probability $1 - \lambda$*, so that the shorted-worker count* n_k *follows eq.* (15)*;*

(ii) a shorted worker complains if and only if $\Delta_k > \omega_{ik}$*, with complaint costs* ω_{ik} *drawn independently across workers and regulations from eq.* (3)*;*

(iii) a complaint on $k \in \mathcal{R}^m$ *triggers an inspection that investigates* $\mathcal{R}_k = \mathcal{R}^m \cup \mathcal{R}^0$*;*

(iv) the firm resolves the surfaced regulations in a random order, with a resolution capacity $\mathcal{C}$ *drawn independently from eq.* (24)*, and the regulations it has not reached when the cumulative burden exceeds* $\mathcal{C}$ *form the recorded set* $\mathcal{A}$ *of eq.* (26)*.*

Then the log-likelihood of the observed outcomes of firms j *with compliance* λ_j *and size* ℓ_j *is*

$$\mathscr{L} \;=\; \sum_{j} \Big[\ln P^{d_j}(\lambda_j, \ell_j) \;+\; \mathbb{1}(d_j = 3)\, \ln f\big(\omega_j \mid d_j = 3; \lambda_j, \ell_j\big) \Big], \tag{28}$$

where $P^d := \mathbb{P}(d)$ *are the outcome probabilities and* $f(\omega \mid d = 3)$ *is the pmf of* $\omega = \{n_k\}_{k \in \mathcal{A}}$*.*

The three probabilities above are the likelihoods of the observed outcome for a firm with characteristics (w, λ, ℓ). Ex ante identical firms are sorted into the three outcomes by realized randomness, and eq. (55)–eq. (57) capture this selection mechanism exactly. These likelihoods depend on the parametrization of worker utility used to quantify Δ_k, defined in section 3.1; apart from that, however, they are independent of the rest of the model.

4.3 *Block 2 Estimation: The Complaint and Inspection Process*

The model's likelihood in proposition 1 forms the basis of the estimation strategy, and Table 3 lists the Block 2 parameters and the features of the data that identify each. We start by estimating the related groups, then hold them fixed. We then make three adjustments to Proposition 1 to bring the data in line with the model. First, we restrict the sample to each firm's years up to and including its first inspection. Second, we construct bins of firm size in

Object	Model object	What identifies it
$\{\mathcal{R}^m\}_{m=0}^{M}$	Related groups	Regulations recorded together
$\tilde{\rho}_0,\ \rho_1$	Distribution of the complaint cost ω_{ik}	Shape of the inspection rate in size
$\gamma_0,\ \gamma_1$	Distribution of the resolution capacity $\mathcal{C}$	Shape of the sanction rate in size
$\{\tilde{\beta}_k^a\}_{k\in\mathcal{K}^a}$	Value of each amenity	How often each amenity is recorded
$\{\lambda_b\}_{b=1}^{15}$	Compliance in each size bin	Shorted workers per recorded regulation

Table 3: Block 2 parameters, estimated by maximum likelihood. The related groups are estimated first and then held fixed. The compliance rates λ_b are firm choices rather than parameters, which we estimate for use as moments in Block 3. The parameters marked with a tilde, $\tilde{\rho}_0 = \rho_0\beta$ and $\tilde{\beta}_k^a = \beta_k^a/\beta$, do not contain the wage-utility weight β, which is estimated in Block 3.

which to estimate a single compliance rate. Third, we restrict compliance estimation to the violations data to create a tighter link between compliance and the recorded set.

4.3.1 *Estimation Strategy*

In the model a complaint opens its related group, and the data support this. Health insurance and pension, for example, are recorded together 14 times as often as independence predicts, and without related groups the model would record about 1 regulation per inspection, against 2.0 to 2.8 in the data. We let the data choose the groups. We place each regulation at a position on a line, let a complaint about one regulation pull in the others with a probability that falls with the distance between them, and estimate the positions by maximum likelihood before the rest of Block 2 (appendix C.2). The positions fall into five clusters, which we take as the related groups $\mathcal{R}^1, \ldots, \mathcal{R}^5$, and labor formalities, which is checked on most visits whatever the complaint, forms the always-checked set $\mathcal{R}^0$. We hold this partition fixed in the rest of Block 2.

We then adjust the likelihood by restricting the sample to ensure that inspection rates are informative about the complaint process. SUNAFIL began operating in 2014 and maintained a consistent inspection policy (a complaint guarantees an inspection) through 2019.[14] We therefore restrict the sample to 2014-2019. Additionally, we observe that worker complaint

[14]Starting in 2020, SUNAFIL began issuing dissuasive letters to firms as a first course of action, breaking the relationship between the inspection rate and the complaint rate.

behavior changes after the first inspection. We therefore use first inspections only: each firm contributes its years in the panel up to and including its first inspection, and nothing after it.[15] This leaves 1,965,447 firm-years, of which 16,067 end in a first inspection without a sanction and 6,020 in a sanction. We observe the recorded set for 3,885 of the sanctioned firm-years (table 15 in appendix C).

We estimate compliance within fifteen size bins, listed in table 4. We estimate one compliance rate λ_b for each size bin b, shared by all eleven regulations. Firm j enters the panel in year t_j^0, the later of 2014 and its year of entry, and is first inspected in year τ_j. Summing over firms splits the log-likelihood into its two parts,

$$\mathscr{L} \;=\; \underbrace{\sum_j \sum_{t=t_j^0}^{\tau_j} \ln P_{jt}^{d_{jt}}}_{\mathscr{L}_{\text{cat}}} \;+\; \underbrace{\sum_j \sum_{t=t_j^0}^{\tau_j} O_{jt}\, \ln f(\omega_{jt} \mid d_{jt}=3)}_{\mathscr{L}_{\text{viol}}}, \tag{29}$$

where $O_{jt} \in \{0,1\}$ indicate that we observe the recorded set ω_{jt} of a sanctioned firm-year.[16] Every term has the closed form of proposition 1, so we evaluate the likelihood exactly.

The log-likelihood depends on the Block 2 parameters

$$\theta_2 \;=\; \big(\; \underbrace{\tilde\rho_0,\ \rho_1}_{\text{complaints}}\;,\; \underbrace{\gamma_0,\ \gamma_1}_{\text{resolution}}\;,\; \underbrace{\{\tilde\beta_k^a\}_{k\in\mathcal{K}^a}}_{\text{amenity values}}\;\big). \tag{30}$$

where worker preferences enter the likelihood only through the silence probability $m_k = e^{-\rho\Delta_k}$ (eq. (19)). The product $\rho\Delta_k$ depends on ρ_0, β and β_k^a only through $\tilde\rho_0 = \rho_0\beta$ and $\tilde\beta_k^a = \beta_k^a/\beta$. We therefore estimate these two objects in Block 2 and recover ρ_0 and β_k^a once Block 3 estimates the wage weight β.

The compliance profile enters the likelihood twice and the two entries carry very different information. In the violations data ($\mathscr{L}_{\text{viol}}$), λ_b can be read directly off the number of shorted workers $n_k \sim \text{Binomial}(\ell, 1-\lambda_b)$ as in eq. (15), correcting for selection into that data. In the categorical data ($\mathscr{L}_{\text{cat}}$), it competes with $\tilde\rho_0$, ρ_1, γ_0, γ_1 and $\tilde\beta_k^a$ to explain who is inspected and who is sanctioned. The two datasets are quite different in size: a few thousand violation records against two million firm-years. We therefore estimate in two stages. The first stage

[15] We observe higher re-complaint rates after a firm's first inspection, whether the inspection originates from a complaint or from SUNAFIL's own initiative.

[16] We assume O_{jt} is uncorrelated with the recorded set, so the observed sets are a random sample of the sanctioned firm-years. appendix C.2 explains why not all sanctioned firm-years appear in the violations data.

restricts the estimation of compliance to the violations log-likelihood, the second fits the remaining parameters to the full log-likelihood at that compliance, and we alternate until convergence.

$$\textbf{Stage A:} \qquad \{\hat{\lambda}_b\} \;=\; \arg\max_{\{\lambda_b\}} \; \mathscr{L}_{\text{viol}}\big(\{\lambda_b\};\; \theta_2\big), \tag{31}$$

$$\textbf{Stage B:} \qquad \hat{\theta}_2 \;=\; \arg\max_{\theta_2} \; \mathscr{L}\big(\theta_2;\; \{\hat{\lambda}_b\}\big). \tag{32}$$

Note that stage A still sees the categorical likelihood through the selection correction applied in the violations likelihood. The selection correction depends on the categorical parameters, so we update them in stage B and feed them back into stage A.

4.3.2 *Estimates and Fit*

Different features of the data identify different parameters. The shape of the inspection curve identifies the two complaint-cost parameters $\tilde{\rho}_0$ and ρ_1. The shape of the sanction curve identifies the two resolution parameters γ_0 and γ_1. The amenity preference weights $\tilde{\beta}^a_k$ are identified from the frequency with which each regulation appears in the violations data, relative to the fixed value of the compensation regulations, $\beta \ln(1+\delta^w_k)$. As mentioned above, the compliance rates λ_b are identified from the number of shorted workers on each recorded regulation.

Table 4 reports the estimates and their standard errors and fig. 3 shows the fit of model to the data.[17] Every parameter is estimated precisely and the model fits the data well. These estimates show that the complaint cost falls with firm size ($\hat{\rho}_1 = 0.2$), so complaints come more readily at larger firms. Resolution capacity also grows with size ($\hat{\gamma}_1 = 0.74$), but so does the burden of shorted workers the firm must resolve, so the net effect of size on the probability of a sanction is ambiguous. Compliance rises from 0.39 at one or two workers to 0.83 at five to ten workers, then more slowly to 0.93 at 1,000 or more workers. Workers value labor formalities, wage timeliness and social security most. Because the amenity values are measured relative to the wage weight, each has a wage equivalent that needs no estimate of β: amenity k is worth the raise x_k that solves $\beta \ln(1+x_k) = \beta^a_k$, so $x_k = e^{\tilde{\beta}^a_k} - 1$. These range from 19% of the wage for working time, rest and leave to 80% for labor formalities.

[17] The standard errors use the sandwich variance in eq. (126) clustered on the firm.

Complaints and resolution			**Compliance**		
$\tilde{\rho}_0$	0.00180	(0.00013)	1–2	0.39	(0.02)
ρ_1	0.20	(0.02)	3	0.69	(0.02)
γ_0	0.24	(0.01)	4	0.75	(0.03)
γ_1	0.74	(0.02)	5–10	0.83	(0.01)
			11–20	0.84	(0.01)
Amenity values $\tilde{\beta}^a_k$			21–29	0.86	(0.02)
Wage timeliness	0.53	(0.04)	30–39	0.85	(0.02)
Working time and leave	0.17	(0.03)	40–49	0.86	(0.02)
Fundamental rights	0.21	(0.02)	50–100	0.86	(0.01)
Workplace safety	0.35	(0.03)	101–149	0.88	(0.02)
Safety management	0.24	(0.02)	150–199	0.90	(0.02)
Labor formalities	0.59	(0.06)	200–299	0.93	(0.02)
Social security	0.54	(0.04)	300–499	0.90	(0.02)
			500–999	0.88	(0.02)
			1,000+	0.93	(0.01)

Table 4: Block 2 estimates, with standard errors in parentheses from the sandwich variance in eq. (126), clustered on the firm. Amenity values are measured relative to the wage weight β. Compliance is reported by the head-count range of each size bin.

The violations data also give us a way to estimate amenity preferences without survey data and without observing firm choices. Existing estimates of how workers value non-wage job attributes come from two sources. Stated-preference studies ask workers to trade wages against those attributes (e.g., Mas and Pallais, 2017). Revealed-preference studies use workers' choices among employers, which requires observing the wages and amenities that each firm offers (e.g., Sorkin, 2018; Lamadon et al., 2022). We instead use observed complaints, under the assumption that a shorted worker complains only when the value of the regulation exceeds their complaint cost and the assumptions of our model.

In fig. 14 in appendix C we show that estimated amenity values respond to perturbations in the slice of the data that we expect to be most informative about them, the frequency with which each regulation is recorded in the violations data. We raise the share that a set of related regulations by 50% and re-estimate Block 2.[18] We should expect to see the preferences for that set rise, while the preferences for the other regulations fall (since the total violation burden remains fixed). Indeed, this is what we see. The last panel is a placebo test: we perturb the frequency of regulated compensation, whose preference is fixed by statute,

[18]We are less able to distinguish between the different amenity preferences within a related group, since a complaint on any regulation within the group triggers the rest to surface and become visible.

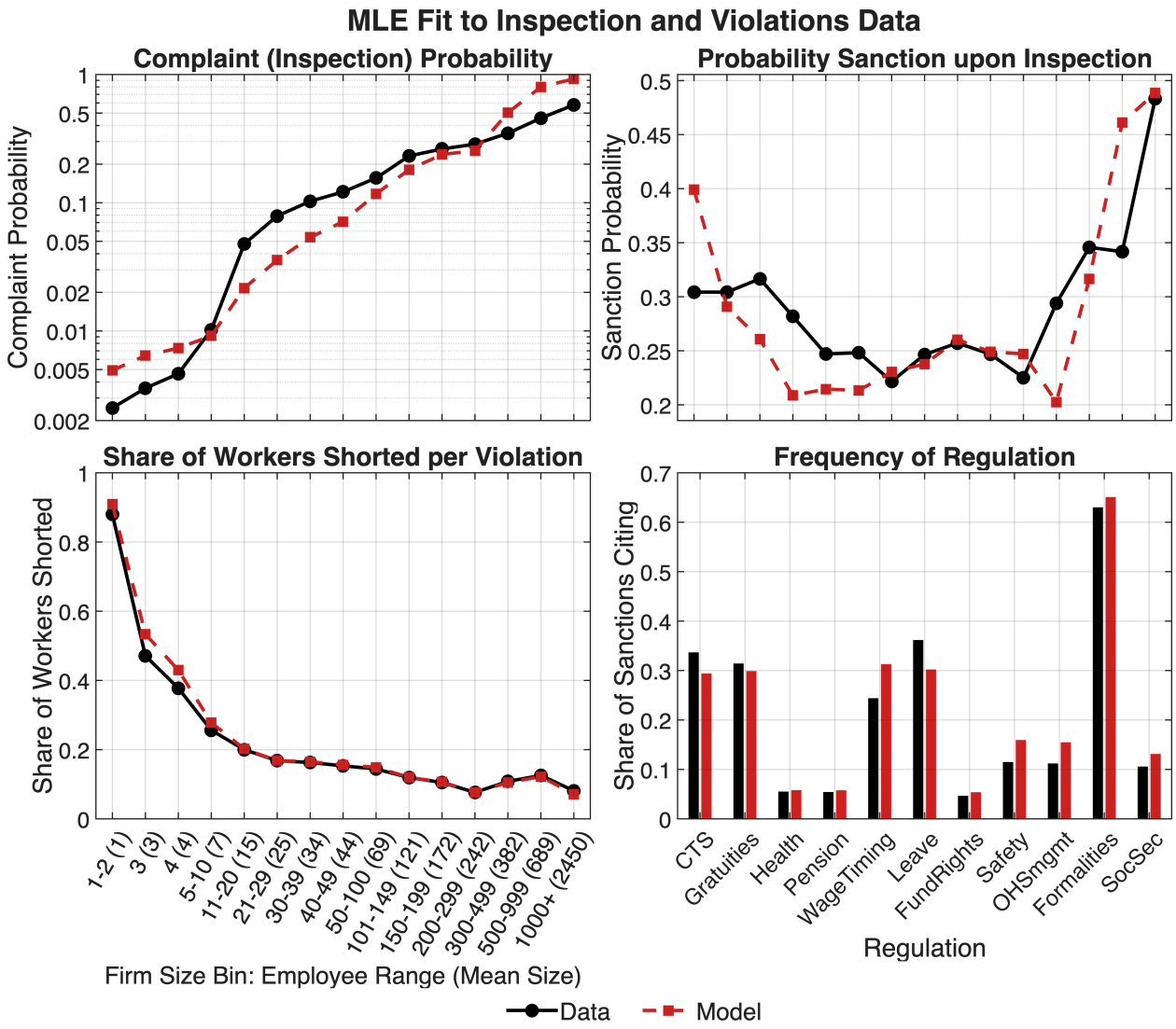


Figure 3: The model's fit to the data. This is the predicted values from the Block 2 estimated parameters against the data used to estimate the model.

and we see an even decline in the preferences for the other sets.

4.4 *Block 3: Costs, Production, and the Labor Market*

In Block 3 we estimate the parameters that govern costs, production, and the labor market by simulated method of moments. These remaining parameters are disciplined by combining the compliance profile estimated in Block 2, the wage profile observed in administrative data, the profit share of value added, and the shape of the firm size distribution. We first

Parameter	Model object	What identifies it
α	Returns to scale, $y = z\ell^{\alpha}$	Profit share of value added
$\mu_z,\ \sigma_z$	Pareto distribution of z	Firms with 10+ workers; tail slope
ξ_0	Taste for the informal sector	Formal employment share
β	Wage weight in utility, $\beta \ln w$	Slope of the wage profile in size
ϕ	Compliance cost, $\mathcal{X}(\lambda)$	Compliance level
c^a	Amenity delivery cost	Compliance slope in size
ϑ	Backpay weight, $V = W + \vartheta B$	Compliance of the largest firms

Table 5: Block 3 parameters, estimated by simulated method of moments, solving the equilibrium at every candidate.

describe the estimator, then the moments, and then show that these moments speak to different parameters: the compliance costs, the labor supply elasticities, the curvature of the production function and the productivity distribution.

4.4.1 *Estimation*

Table 5 lists the Block 3 parameters and the moments that identify each.

The equilibrium of section 3.3 forms the basis of the estimation strategy: solving it at a candidate parameter vector produces a cross-section of firms that we compare with the data. We estimate these eight parameters in three steps. First, we solve the equilibrium at each candidate parameter vector, where each candidate value of β gives us $\rho_0 = \tilde{\rho}_0/\beta$ and $\beta_k^a = \tilde{\beta}_k^a\,\beta$. Second, we score each candidate by a weighted distance between simulated and data moments. Third, we minimize that distance in two phases.

$$\theta_3 = \big(\underbrace{\alpha,\ \mu_z,\ \sigma_z}_{\text{production}},\quad \underbrace{\xi_0}_{\text{labor supply}}\quad,\ \underbrace{\phi,\ c^a}_{\text{costs}},\quad \underbrace{\beta, \vartheta}_{\text{preferences}}\quad\big), \tag{33}$$

The only source of heterogeneity in the model is the productivity distribution, which we take to be Pareto with lower bound μ_z and tail index σ_z. We place S nodes at equal-probability intervals of that distribution, and solve the model at each node and compute the simulated moments as the weighted average across nodes. The grid is deterministic and identical at every candidate, so a difference in the criterion is a difference in the model and never in a draw. We use one node per formal firm, $S = J$, which gives the top of the size

Moment	Definition	Source
Compliance		
λ, 1–3 workers	Mean of $\hat{\lambda}_b$ over bins 1–2	Block 2 estimates
λ, 4–199 workers	Mean of $\hat{\lambda}_b$ over bins 3–11	Block 2 estimates
λ, 200–499 workers	Mean of $\hat{\lambda}_b$ over bins 12–13	Block 2 estimates
λ, 500+ workers	Mean of $\hat{\lambda}_b$ over bins 14–15	Block 2 estimates
Wages		
$\ln w$, all firms	Mean of $\ln(w/w_0)$ over all fifteen bins	Employer–employee panel
$\ln w$, 500+ workers	Mean of $\ln(w/w_0)$ over bins 14–15	Employer–employee panel
Scale		
Formal employment share	N_F/L, formal jobs over total employment	Block 1 aggregates
Firms with 10+ workers	The share of firms at or above ten workers	Employer–employee panel
Tail slope	$\ln s_{100} - \ln s_{10}$, how fast that share thins out	Employer–employee panel
Profit share of value added	Median over firms with 50+ workers	Economic Census

Table 6: The Block 3 moments. The size groups are unions of the fifteen bins of table 4.

distribution the same support in the simulation that it has in the data.

We minimize eq. (131) in two phases. A greedy algorithm first moves a few parameters at a time against the category of moments that fits worst, solving the equilibrium at every trial. A local refinement then minimizes the criterion over all eight parameters jointly, starting from where the first phase stops (appendix C.3).

4.4.2 *Moments*

Table 6 lists the 10 moments one by one. We compute each simulated moment with the same statistic as its data counterpart, so that a gap reflects the model rather than a difference in measurement. Each moment is weighted equally in the objective.

The compliance targets are the Block 2 estimates $\hat{\lambda}_b$ from eq. (31), averaged over the bins of four size groups with each bin weighted by its count of firms. In order to compare model and data moments, we need to perform the same measurement on the simulated economy. Taking the average compliance of the simulated firms in a bin does not recover the mean compliance of the data, because the compliance bins are obtained from a maximum-likelihood procedure. To keep things computationally tractable, instead of re-estimating the compliance profile on the simulated economy, we use the information-weighted mean, which is the first order approximation of the measure of interest. We then verify that the approximation is accurate by re-estimating the compliance profile on the simulated economy in the final

stage.[19]

We obtain wage moments from matched Employer-Employee Panel data for the entire formal sector in 2016. We compute mean log monthly wages by size bin relative to the informal wage w_0, take the average across months, and then average the fifteen bin means with each bin weighted by its count of firms. We take a second wage moment over firms with 500 or more workers alone. The two are not redundant: the all-firm mean is almost entirely a statement about the bottom of the distribution, so the additional moment provides information about the top of the distribution. In fig. 17, we check that adding worker characteristics does not change the wage profile (see appendix C.3 for more details).

We obtain moments from the firm size distribution using the same data source as the wage profile. We take the fraction of firms with at least 10 workers, which fixes the level of the distribution, and the tail slope $\ln s_{100} - \ln s_{10}$, which fixes how fast the share thins out between ten and a hundred workers. We take the formal employment share N_F/L as a moment, which is informative about the taste for the informal sector ξ_0.

Finally, to separate α from the firm size, we compute profit share of value-added from the 2021 Economic Census. Production in the model is value added and profit is operating surplus, which includes the return to capital, so the census share of operating surplus in value added is the direct counterpart of the model's profit share. Unfortunately, the census does not exist in 2016, but we compute the same measure using the 2007 census and find that the profit share is extremely stable over time. Additionally, the Census head count includes working owners and unpaid family members whom the census payroll does not pay, so at small firms operating surplus includes the return to their labor. We therefore take the median profit share of value added among firms with at least 50 workers, in both the census and the simulated economy (appendix C.3).

4.4.3 *Estimates and Identification*

The model fits profit share and firm size distribution almost exactly, and the compliance profile matches the qualitative shape of the data. Table 7 reports the estimates and fig. 4 shows the fit of the model to the moments. We show the particular moments used in the estimation along with the rest of the moment profile, which serves as a validation of the model. The largest misfit is in the wage profile, where the model matches the two targeted

[19] See appendix C.3 for a derivation of the information-weighted mean.

	Parameter	Estimate		Parameter	Estimate
Production	α	0.86	Costs	ϕ	0.64
	μ_z	3.92		c^a	0.36
	σ_z	3.55	Preferences	β	3.21
Labor supply	ξ_0	21.25		ϑ	0.66

Table 7: Block 3 estimates.

moments but does not capture the curvature of the data. No utility function in the CRRA family from log to linear can match the curvature of the data. One common solution in the literature is to incorporate exogenous amenities ξ_j, which we can incorporate as a draw with a parametrized correlation with productivity (and hence size) that closes the residual. For now, we leave this as a potential extension, and accept that the model does not capture the curvature of the wage profile.

With the wage weight $\hat{\beta} = 3.2$, the full regulatory package is worth $D = \sum_{k \in \mathcal{K}} \Delta_k = 9.9$ utils to a worker (eq. (5)), the same as raising her log wage by $D/\beta = 3.1$. There are three key parameter estimates to highlight. First, the opportunity cost of compliance parameter is $\hat{\phi} = 0.64$. Recall that $\phi = \mathcal{X}_1$, the mean shadow opportunity cost a firm pays on its full-compliance bill, such as when a penalty is owed $((1 + \mathcal{X}_1)(\mathcal{P}^{\text{bp}} + \mathcal{P}^{\text{s}}))$. Thus, the penalty opportunity cost adds 64% to the backpay transfer and to the sanction. Second, the amenity delivery cost for each of the seven regulated amenities is $\hat{c}^a = 0.36$ relative to the one-to-one transfer for wages. Finally, workers weight the backpay value by $\hat{\vartheta} = 0.66$ in the firm selection stage. This could be a real cost of discounting, or a lack of foresight on the worker's part. The interpretation of this parameter becomes relevant only when discussing welfare.

Each parameter moves several moments, but the moments that move most differ across parameters (fig. 16 in appendix C.3). The returns to scale α moves the profit share, the Pareto parameters move the level and tail slope of the firm size distribution, the informal taste ξ_0 moves the formal employment share, and the wage weight β moves the slope of the wage profile. Among the cost parameters, ϕ moves the level of compliance, c^a its slope, and ϑ the compliance of the largest firms. The sensitivity matrix has full rank, so the parameters are locally identified (see Andrews et al., 2017).

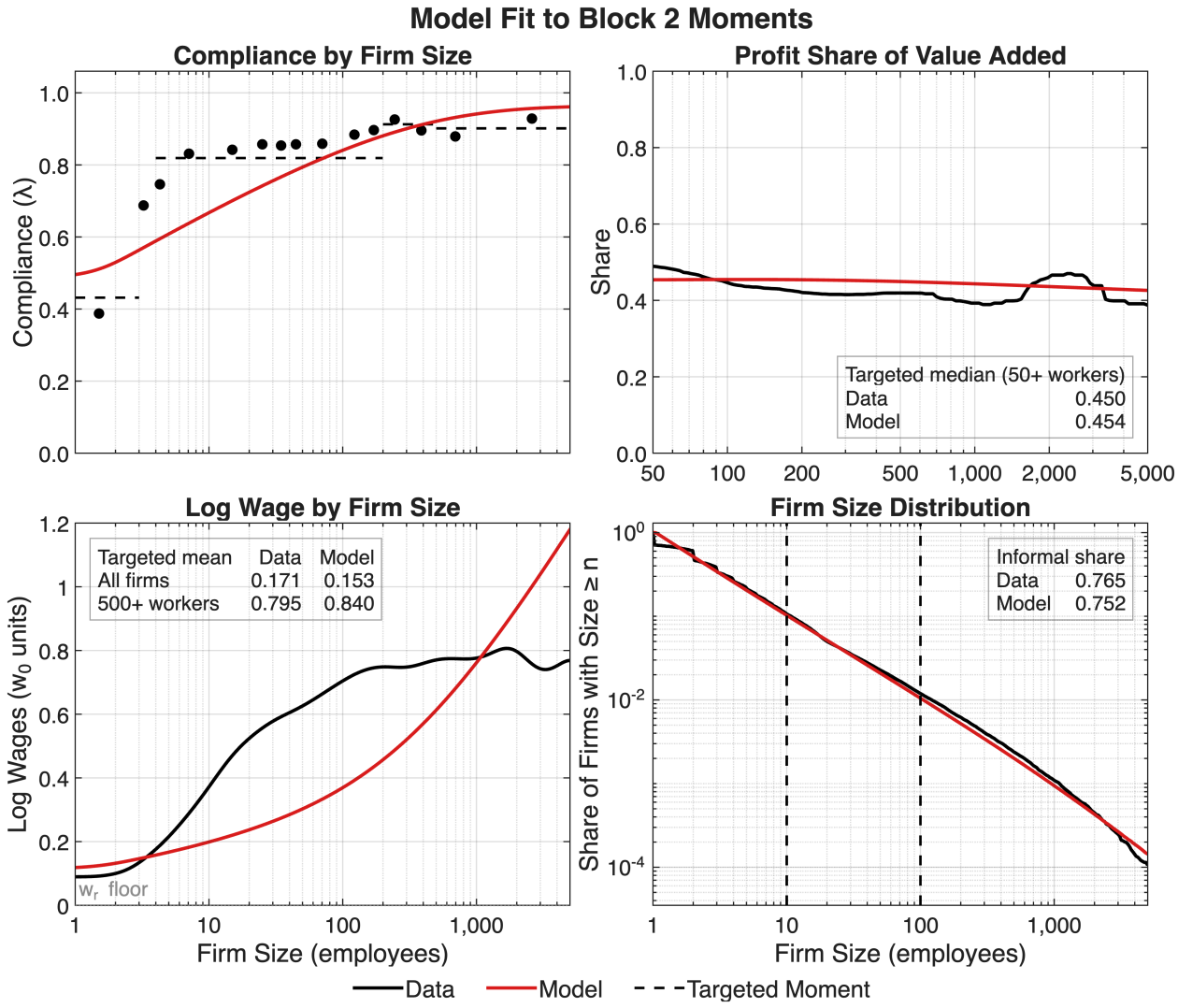


Figure 4: Model against data. Black is the data and red the model. Dashed lines mark the targeted moments; the points and curves behind them are the size-bin profiles, which the estimation does not use.

5 Results and Counterfactuals

We use our estimated model to compare firm choices across the three enforcement regimes and to illustrate two mechanisms in the Peruvian labor market. We first show how enforcement changes compliance, wages and size, and why it can lower compliance. We then show the compensating differential obtained with compliance, and we decompose worker utility. Finally, we ask why the size threshold produces only mild bunching at the enforcement rates we estimate, and how much it would produce under perfect enforcement.

5.1 *Firm Choices across Enforcement Regimes*

At first glance, the regulatory burden that the Peruvian labor code imposes on firms seems like enforcement would cause a large contraction in the formal sector (Botero et al., 2004; International Monetary Fund, 2024). We find that imperfect enforcement in the status quo avoids that outcome, but perfect enforcement does indeed decrease hiring in the formal sector.

Figure 5 plots the three firm choices – compliance, wages, and size – under the no-enforcement, status quo, and perfect enforcement regimes. Status quo enforcement barely moves firm choices relative to no-enforcement, and compliance remains high in both. This is consistent with Viollaz (2018), who finds in household survey data that enforcement in Peru has little effect on compliance or firm size. At small firms, whose expected penalties are small because complaint rates are low, the two lines coincide for compliance, wages, and size. The largest firms comply less and pay higher wages under the status quo enforcement regime, though the differences are small. Enforcement raises the expected penalty for noncompliance, which encourages compliance, but it also creates a channel in which complying reduces the backpay value a worker receives, which discourages compliance. The second effect dominates at every size, so enforcement lowers compliance across the whole size distribution. We discuss this further in section 5.1.1.

Under perfect enforcement firm choices change at every size. Small firms raise compliance dramatically, pay the minimum wage, and ration workers. Above a few hundred workers, firms comply less than without enforcement, pay higher wages, and hire fewer workers. Formal employment falls from 25% of the workforce under no enforcement to 13% under perfect enforcement, against 25% in the status quo. The black line is the planner benchmark of section 6.1, which moves workers out of the informal sector and toward the largest firms.

5.1.1 *The Compliance Response to Enforcement*

Surprisingly, enforcement does not necessarily raise compliance at every firm. Moving from a no-enforcement world to the status quo enforcement economy affects the firm's compliance choices through terms of opposite sign. We can decompose the change in the marginal cost

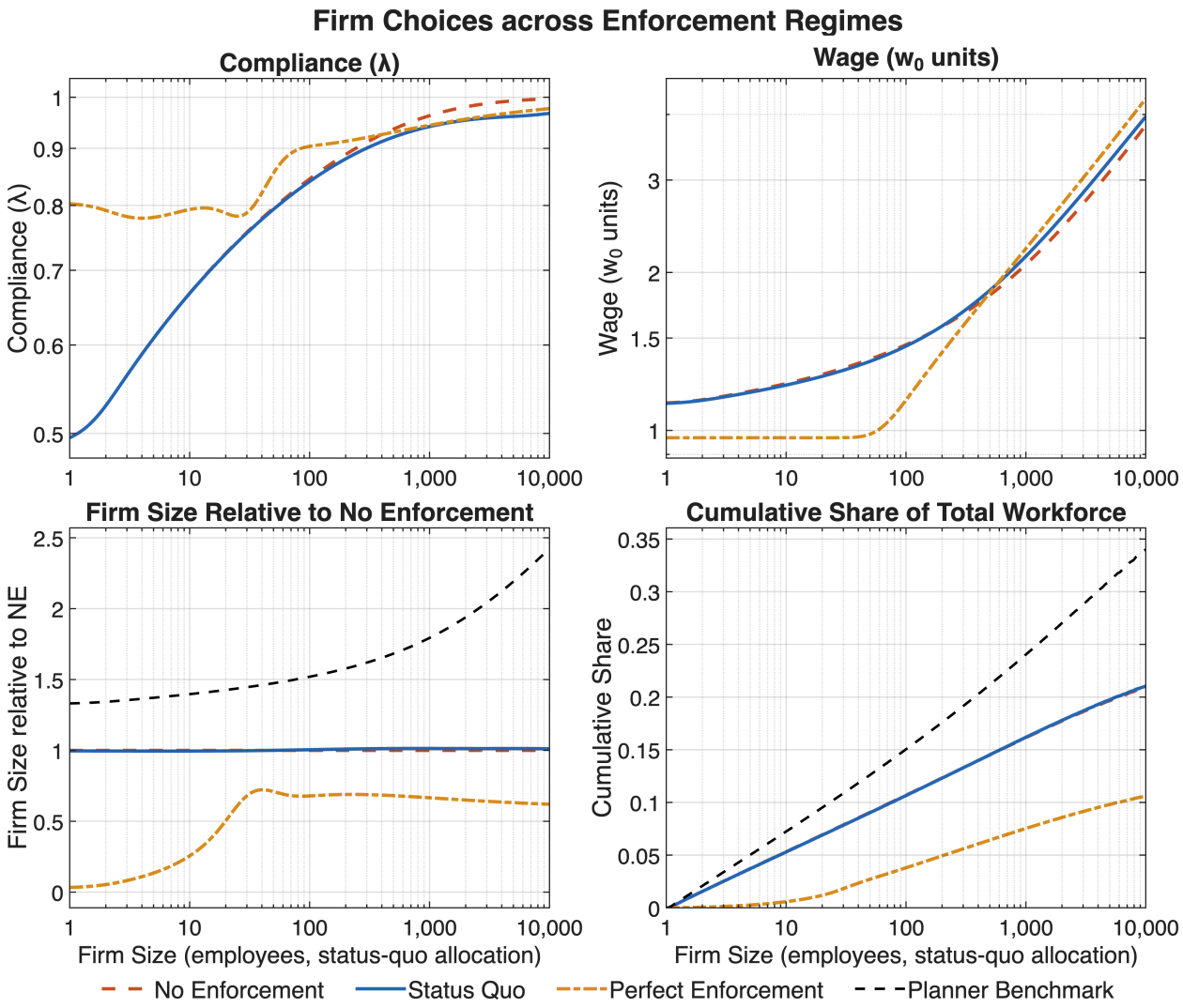


Figure 5: Firm choices across enforcement regimes, against firm size in the status quo. The planner benchmark is defined in appendix D.2.

of compliance at a firm into

$$c_\lambda(w^E,\lambda^E) - c_\lambda(w^E,\lambda^{NE}) \;=\; \underbrace{-\frac{\mathcal{P}_\lambda}{\ell}}_{\substack{\text{penalty}\\ \text{saving}}} + \underbrace{\eta\mu^E\vartheta B_\lambda}_{\text{backpay loss}} + \underbrace{\eta D(\mu^E-\mu^{NE})}_{\substack{\text{change in}\\ \text{recruiting cost}}} - \underbrace{\left[c_\lambda(w^E,\lambda^{NE}) - c_\lambda(w^{NE},\lambda^{NE})\right]}_{\text{wage effect}}$$

If the left hand side is positive, then compliance rises with enforcement. The first term on the right hand side is the penalty savings: enforcement raises the expected penalty for non-compliance, which encourages compliance. The second term centers around backpay. With

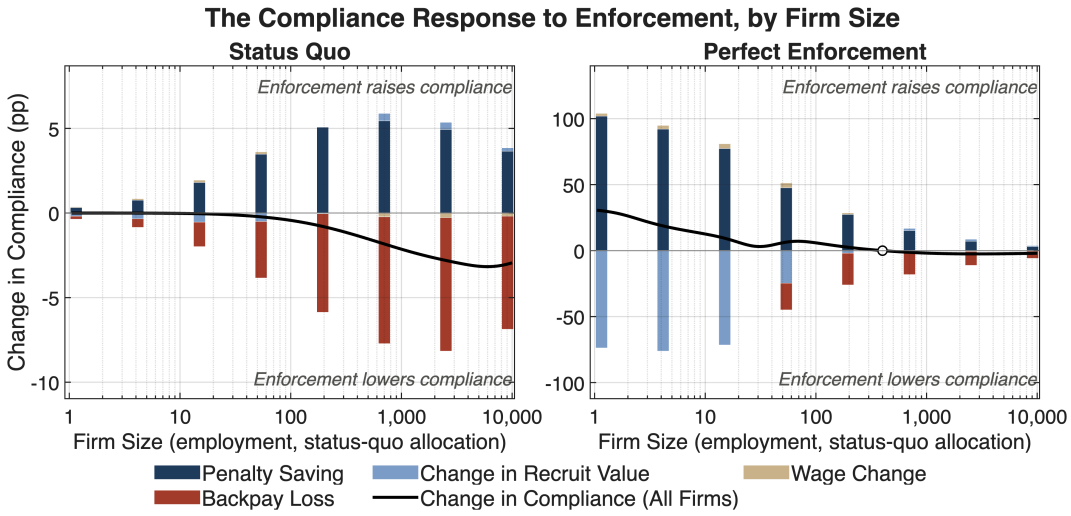


Figure 6: The compliance response to enforcement. The left panel is the status quo and the right panel perfect enforcement, each against no enforcement. Bars stack the four terms of the change in the marginal cost of compliance, in percentage points of compliance, and the line is the net change.

enforcement, workers expect to receive backpay, which attracts workers to the firm. When a firm complies, it reduces the expected backpay that workers receive, which discourages compliance. Increasing compliance delivers utility to workers in both no enforcement and enforcement regimes, so it disappears from the difference. The third term is the change in the recruiting cost μ, which can either encourage or discourage compliance depending on whether μ rises or falls with enforcement (we discuss this in section 6.1). The fourth term is the wage effect: when a firm pays a lower wage, the marginal cost of compliance is lower, which encourages compliance.

Figure 6 shows the net effects across the two regimes: status quo and perfect enforcement. The left panel shows that the backpay loss dominates the penalty savings across the firm size distribution, so that compliance falls everywhere. The right panel shows that perfect enforcement paints a different picture. The smallest firms are rationing labor, which means that supply of workers is greater than demand at the minimum wage. The cost of recruiting a worker goes to zero, which reduces the incentive to comply. However, the penalty savings are large and pushes compliance up. As firm size increases, firms stop rationing labor, so the backpay loss eventually grows for them until compliance falls at the largest firms.

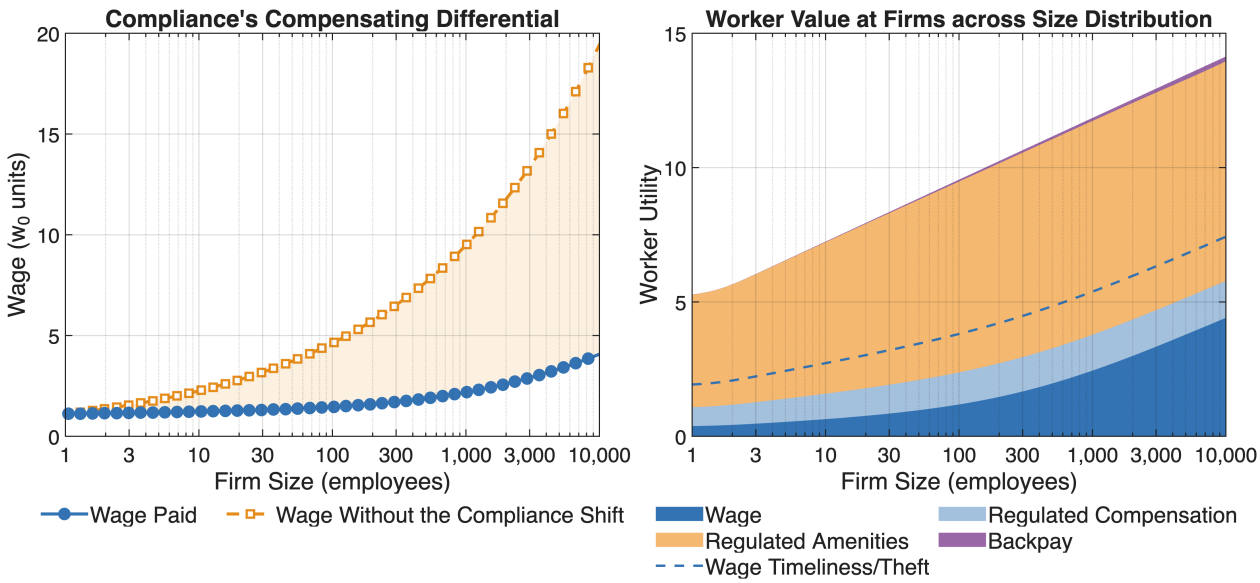


Figure 7: Compliance as compensation. Panel 1 is the simulated wage profile against the profile that would obtain with compliance held fixed. Panel 2 decomposes the utility a worker receives at a firm into its sources.

5.2 *Compliance as Compensation*

Workers value compliance and pay for it by accepting lower wages. We can see in panel 1 of fig. 7 that the wage profile is flatter than it would be if compliance were held fixed. The figure plots the estimated wage profile versus a simulated profile where compliance is held fixed. When compliance is held fixed, firms must move up the labor supply curve using only wages to attract workers. However, in practice, larger firms increase compliance along with wages which shifts the labor supply curve out and thus flattens the wage profile. Studies often attribute the residual between the measured labor supply elasticity and the observed wage profile to exogenous amenities (e.g., Lamadon et al., 2022). We partially endogenize this residual exogenous amenity as endogenous firm compliance with regulated compensation and regulated amenities.

Workers place a large value on amenities relative to wages, so compliance is a powerful way to deliver utility. Panel 2 of fig. 7 decomposes the value workers receive at each firm size into wage utility, regulated compensation, regulated amenities (with wage timeliness, which could also be read as a wage component, marked separately) and backpay. Backpay matters only at the largest firms, which are much more likely to be inspected. The slope of the total shows how much utility a firm must offer to attract additional workers.

The same logic matters for how the wage elasticity of labor supply is measured. A more productive firm grows by raising its compliance along with its wage, so the employment change that comes with a wage change reflects both. The total change in labor supply associated with a change in wage is

$$\frac{d\ln N_j}{d\ln w_j} = \underbrace{\frac{1}{1-\eta_j\vartheta\,\ell_j B_\ell}}_{\text{size feedback}}\Bigg[\underbrace{\eta_j\beta}_{\text{compliance fixed}} + \underbrace{\eta_j\big(D+\vartheta B_\lambda\big)\frac{d\lambda_j}{d\ln w_j}}_{\text{compliance moves with the wage}}\Bigg], \tag{34}$$

where $d\lambda_j/d\ln w_j$ is how much compliance rises with the wage across firms (see appendix A.9 for a derivation). The first term is the wage elasticity of labor supply that determines the markdown, and the second is the labor supply that higher compliance attracts. Figure 8 plots each term by firm size. The wage elasticity is $\hat{\beta} = 3.2$ at every firm, which is close to the estimates in the literature surveyed by Caldwell et al. (2026). The total change in labor supply is much larger: it averages 26 across firms, of which the compliance term is 22. Assuming that the entire observed employment change is due to wages would therefore over-estimate the wage elasticity of labor supply and under-estimate markdowns. The markdown averages 43% of the wage, against 5.7% if compliance is assumed fixed (fig. 12 in appendix A.9). The gap declines for larger firms, where compliance moves less with the wage: the total elasticity falls from 28 at firms with one or two workers to 3.8 above a thousand.

5.3 *Bunching at the Size Threshold*

At twenty workers a firm begins to owe profit sharing and must create a safety and health committee (see Appendix B.1 for details). This shifts the expected penalty, while what firms deliver and what workers value are continuous in ℓ (i.e. firms can always choose to provide these benefits). What jumps is the set that the firm can be cited on: below twenty a firm is citable on 88% of the sanction base it would face above.

Figure 9 shows the empirical and model simulated size distribution around this threshold. The left panel shows mild bunching at twenty in the empirical distribution: the excess mass of firms just below twenty is $b = 0.64$ (standard error of 0.25), about two-thirds of a counterfactual size bin's worth of extra firms (appendix B.2). The right panels show that the model reproduces it under the status-quo enforcement regime. This is despite the fact that the change in expected penalty does indeed jump by 0.14. However, the jump is not large enough to induce a firm to stay small, since the enforcement rate at twenty is low and

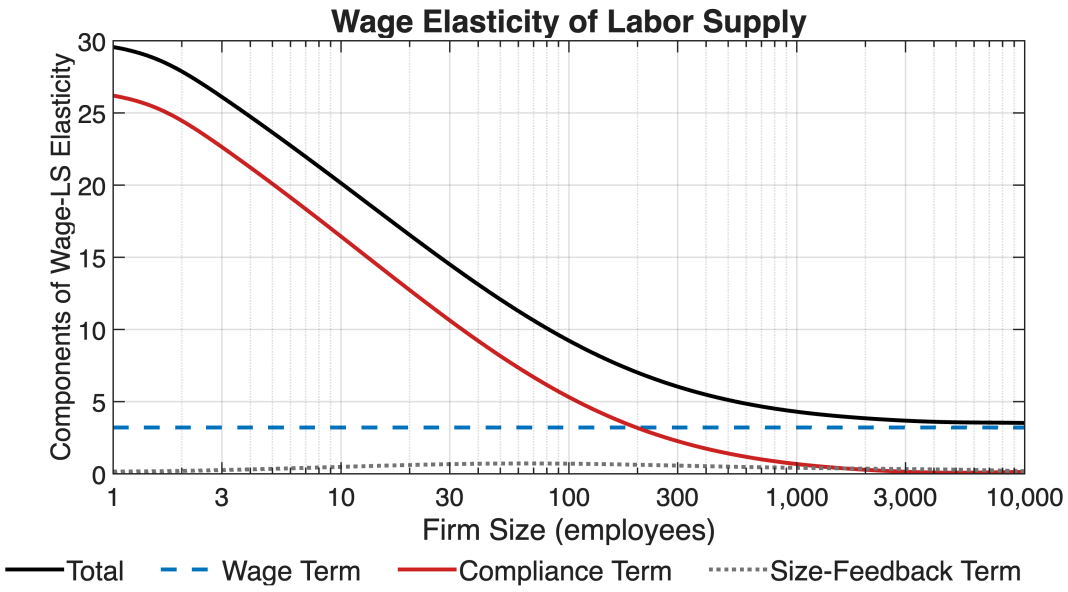


Figure 8: The wage elasticity of labor supply by firm size in the status quo. The total is the sum of the wage term, the compliance term and the size feedback of eq. (34).

smooth around the threshold. The bunching the model produces under the size-dependent schedule is almost indistinguishable from what a uniform schedule would produce.

Under perfect enforcement the same schedule produces a spike below twenty and a hole above it. Because rationing pushes most small firms below one worker under perfect enforcement, fig. 9 shows that regime's size distribution among firms with at least one worker, which shifts its mass toward larger sizes. Crossing the threshold raises the expected penalty on an added worker by 6.9 under perfect enforcement, which is 26% of the profit of a firm at nineteen workers. Thus, unlike the literature looking at bunching in contexts with strong enforcement (e.g. Garicano et al., 2016), bunching in contexts with weak enforcement depends more on the enforcement regime than on the statutory schedule.

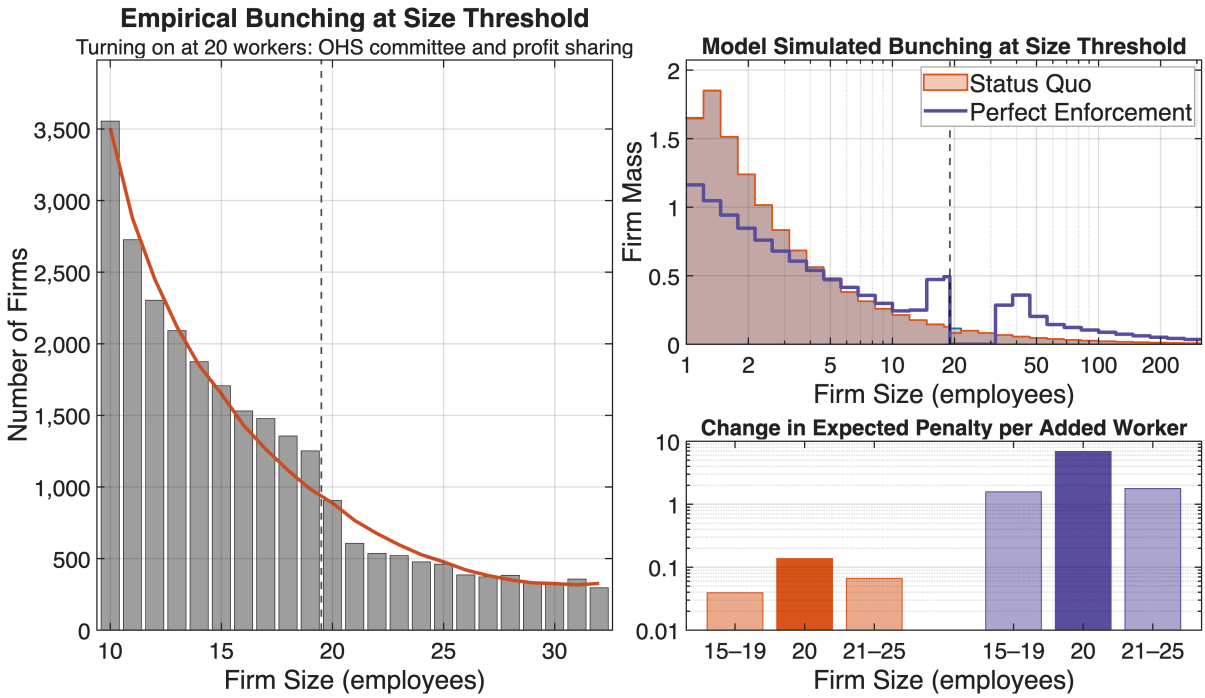


Figure 9: Bunching at twenty workers. The left panel is the 2016 size distribution against a smooth counterfactual. The right panels give the simulated distribution under the three regimes and the step in the expected penalty that creates the notch. The perfect-enforcement distribution is conditional on firms with at least one worker: rationing shrinks most small firms below one worker, so only 19% of firms remain and the curve is rescaled to them.

6 Welfare

In this section, we ask how enforcement changes welfare. We derive a planner benchmark and decompose the welfare change from no enforcement into three sources: where workers work, what firms post, and the net cost of enforcement. Finally, we report how workers and firms fare under each regime.

6.1 *Welfare Accounting*

To compare regimes we need a measure of welfare and a benchmark to measure each regime against. We measure welfare as profits plus worker utility, and take as the benchmark the allocation that maximizes it at the no-enforcement postings. The welfare change from no enforcement then splits into three sources: where workers work, what firms post, and the net cost of enforcement. We derive every step in appendix D.

6.1.1 *Welfare function*

To add profits and utility, we convert utility into money at the marginal utility of money at the mean wage, $\mathsf{m} := \beta/\bar{w}$, where $\bar{w} = 1.7$ is money income per worker in the status quo. We call the wage and compliance of a firm its *posting*, $\mathbf{p}_j := (w_j, \lambda_j)$, and write $\mathbf{p}$ for the postings of all firms. Every worker who chooses a firm works there, so employment $\ell = \{\ell_j\}_{j \in \mathcal{J}_0}$ is the logit share of eq. (2), with the informal sector indexed by $j = 0$ and $\mathcal{J}_0 := \mathcal{J} \cup \{0\}$. Welfare is profits, plus worker utility converted to money, less the net enforcement cost (appendix D.1),

$$\mathcal{W} \;=\; \mathcal{F}(\ell;\, \mathbf{p}) \;-\; \text{NEC}, \tag{35}$$

where

$$\mathcal{F}(\ell;\, \mathbf{p}) \;:=\; \sum_{j \in \mathcal{J}} \Big[\, y_j - c_j \ell_j + \frac{W_j\, \ell_j}{\mathsf{m}} \,\Big] \;+\; \frac{U^{\text{taste}}(\ell)}{\mathsf{m}}, \tag{36}$$

$$\text{NEC} \;:=\; \sum_{j \in \mathcal{J}} \Big[\, \underbrace{\mathcal{X}_1\big(\mathcal{P}_j^{\text{bp}} + \mathcal{P}_j^{\text{s}}\big)}_{\substack{\text{penalty} \\ \text{opportunity cost}}} \;+\; \underbrace{\mathcal{P}_j^{\text{bp}} - \frac{\ell_j\, \vartheta B_j}{\mathsf{m}}}_{\substack{\text{backpay paid less} \\ \text{its value to workers}}} \;+\; \underbrace{c^{\text{insp}}\, \mathscr{I}_j}_{\text{agency spending}} \,\Big]. \tag{37}$$

The surplus $\mathcal{F}$ is output less the per-worker cost of each posting, plus the money value of what the postings deliver, plus $U^{\text{taste}}(\ell) := \ell_0\, \xi_0 - \sum_{j \in \mathcal{J}_0} \ell_j \ln(\ell_j/L) + L\gamma_E$, the value of the informal option and of the taste draws, with γ_E Euler's constant. The net enforcement cost is the opportunity cost of paying backpay at the inspection date, the backpay firms pay less its value to workers, and agency spending, whose cost per inspection c^{insp} we anchor to SUNAFIL's budget (appendix D.7). Without enforcement NEC $= 0$ and $\mathcal{W} = \mathcal{F}$.

We report each welfare change $\Delta\mathcal{W}$ as the uniform raise to every worker's wage that changes welfare by the same amount. A raise of x log points raises every worker's utility by βx, so $\Delta\mathcal{W} = L\beta x/\mathsf{m}$ and

$$x \;=\; \frac{\Delta\mathcal{W}}{L\, \bar{w}}. \tag{38}$$

Log points add across the terms of a decomposition, which the exact raise $e^x - 1$ does not. We therefore report the lines of a decomposition in log points and the headline totals in percent.

6.1.2 *The Wedge*

The *marginal surplus* of firm j is the change in surplus from moving one worker from the informal sector to firm j, holding every other firm's employment and every posting fixed,

$$G_j(\ell;\,\mathbf{p}) \;:=\; \frac{\partial \mathcal{F}(\ell;\,\mathbf{p})}{\partial \ell_j} \;=\; MRPL_j \;-\; c_j \;-\; \frac{\mathcal{V}_j(\ell) - W_j}{\mathsf{m}}, \qquad \mathcal{V}_j(\ell) \;:=\; \xi_0 + \ln \frac{\ell_j}{\ell_0}, \quad (39)$$

where $\mathcal{V}_j(\ell)$ is the inverse labor supply: the value workers would have to sort on for firm j to employ ℓ_j workers against ℓ_0 in the informal sector. The cost of the marginal worker at j is the posting's cost c_j plus the money value of the gap between $\mathcal{V}_j(\ell)$ and what the posting delivers, W_j. The planner's optimum ℓ^* maximizes the surplus at the no-enforcement postings, and it is the unique allocation at which every marginal surplus is zero (appendix D.2).

We call the marginal surplus at a regime's allocation, evaluated at the no-enforcement postings $\mathbf{p}^{NE}$, its *wedge*. Every regime's wedge is then the slope of the same function $\mathcal{F}\big(\cdot\,;\,\mathbf{p}^{NE}\big)$ whose peak is ℓ^*, so comparing wedges compares how far each allocation is from ℓ^*, and enforcement's own postings and costs enter the welfare change separately. Write $E \in \{SQ, PE\}$ for an enforcement regime. Under free sorting, and with $MRPL_j$ evaluated at each regime's own allocation,

$$\text{Planner:} \quad G_j\big(\ell^*;\,\mathbf{p}^{NE}\big) \;=\; 0, \quad (40)$$

$$\text{No enforcement:} \quad G_j^{NE} \;:=\; G_j\big(\ell^{NE};\,\mathbf{p}^{NE}\big) \;=\; MRPL_j - c_j^{NE}, \quad (41)$$

$$\text{Enforcement:} \quad G_j^{E} \;:=\; G_j\big(\ell^{E};\,\mathbf{p}^{NE}\big) \;=\; MRPL_j - c_j^{NE} - \frac{W_j^E - W_j^{NE} + \vartheta B_j}{\mathsf{m}}. \quad (42)$$

Workers sort on what firms offer, so the wedge subtracts the gap between what workers sort on and what the no-enforcement posting delivers. Without enforcement that gap is zero, and the firm hires until $MRPL_j = c_j + \mu_j$ (eq. (93)), so the wedge is the recruiting cost, which without enforcement is the monopsony markdown, $G_j^{NE} = \mu_j^{NE} > 0$. We drop the j subscript on η, since $\eta_j \approx 1$. Under enforcement, workers sort on $W_j^E + \vartheta B_j$, which differs from what the no-enforcement posting delivers by the change in the posting and by backpay. Backpay is not a cost to the firm, and its value to workers enters welfare through the net enforcement cost rather than through the posting. A firm whose workers sort on backpay therefore draws more workers than its posting alone would justify, which lowers its wedge.

At the estimate the no-enforcement market falls short of the planner benchmark by 11 log

points of wages. The planner raises formal employment from 25% to 46% of the workforce, and raises employment to 1.3 times its market level at the smallest firms and 2.8 times at the largest (fig. 5).

6.1.3 *Decomposing the Welfare Change*

By eq. (35), welfare in the two regimes is

$$\mathcal{W}^{NE} \;=\; \mathcal{F}\big(\ell^{NE};\, \mathbf{p}^{NE}\big), \qquad\qquad \mathcal{W}^{E} \;=\; \mathcal{F}\big(\ell^{E};\, \mathbf{p}^{E}\big) \;-\; \mathrm{NEC}^{E},$$

where NEC^{E} is eq. (37) at the enforcement allocation. Adding and subtracting the surplus at the enforcement allocation with the no-enforcement postings, $\mathcal{F}\big(\ell^{E};\, \mathbf{p}^{NE}\big)$, splits the welfare change into three steps that each change one thing:

$$\begin{aligned}\Delta\mathcal{W} \;:=\; \mathcal{W}^{E} - \mathcal{W}^{NE} \;=\; &\underbrace{\mathcal{F}\big(\ell^{E};\, \mathbf{p}^{NE}\big) - \mathcal{F}\big(\ell^{NE};\, \mathbf{p}^{NE}\big)}_{\text{reallocation}} \\ &- \underbrace{\Big[\mathcal{F}\big(\ell^{E};\, \mathbf{p}^{NE}\big) - \mathcal{F}\big(\ell^{E};\, \mathbf{p}^{E}\big)\Big]}_{\text{posting}} - \underbrace{\mathrm{NEC}^{E}}_{\text{net enforcement cost}}.\end{aligned} \tag{43}$$

The reallocation moves the allocation from ℓ^{NE} to ℓ^{E} with the postings held at $\mathbf{p}^{NE}$. The posting term switches the postings from $\mathbf{p}^{NE}$ to $\mathbf{p}^{E}$ with the allocation held at ℓ^{E}. The last term subtracts the net enforcement cost at ℓ^{E}.

Reallocation. The marginal surplus is the slope of the surplus, so the reallocation adds up the marginal surplus of every worker moved. We write it in terms of the change in each firm's wedge (proposition 4),

$$\text{Reallocation} \;=\; -\sum_{j\in\mathcal{J}} \big(G_j^{E} - G_j^{NE}\big)\, \bar{\ell}_j^{\,\text{gap}}, \qquad \bar{\ell}_j^{\,\text{gap}} \;:=\; \int_0^1 \big[\mathbf{Q}(s)\, G(s)\big]_j \, ds, \tag{44}$$

where $G(s) := G^{NE} + s\,(G^{E} - G^{NE})$ moves the wedges in a straight line from one regime to the other, and $\mathbf{Q}(s)$ is minus the inverse of the Hessian of $\mathcal{F}\big(\cdot\,;\, \mathbf{p}^{NE}\big)$ along that path. To first order, $\mathbf{Q}\,G$ is the employment the planner would add at each firm, so the *employment gap* $\bar{\ell}_j^{\,\text{gap}}$ is the number of workers the planner would add at firm j, averaged along the path. Raising firm j's wedge by one unit therefore lowers welfare by $\bar{\ell}_j^{\,\text{gap}}$, and it raises welfare at

a firm the planner would shrink ($\bar{\ell}_j^{\text{gap}} < 0$).

Enforcement moves the wedge through the firm's hiring decision. The firm hires at its enforced posting, and its hiring condition in eq. (18) sets $MRPL_j - c_j^E = \mu_j^E + \mathcal{P}_\ell - \eta\mu_j^E \vartheta \ell_j B_\ell$. The wedge instead values each worker at the no-enforcement posting, so it also carries $\Delta c_j^{\text{net}} := (c_j^E - c_j^{NE}) - (W_j^E - W_j^{NE})/\mathsf{m}$, the change in the posting's cost per worker net of the change in what it delivers. At a firm that hires on its labor supply curve and pays above the minimum wage in both regimes, splitting μ_j^E with the wage condition of eq. (17) gives (proposition 5)[20]

$$\begin{aligned} G_j^E - G_j^{NE} \;=\; & \underbrace{\frac{w_j\, c_w}{\beta\eta} - \mu_j^{NE}}_{\text{markdown}} \;+\; \underbrace{\frac{w_j\, \mathcal{P}_w}{\beta\eta\ell_j}}_{\text{wage penalty}} \;+\; \underbrace{\mathcal{P}_\ell^{\text{bp}} - \eta\mu_j\, \vartheta\, \ell_j B_\ell - \frac{\vartheta B_j}{\mathsf{m}}}_{\text{backpay}} \\ & + \underbrace{\mathcal{X}_1\big(\mathcal{P}_\ell^{\text{bp}} + \mathcal{P}_\ell^{\text{s}}\big)}_{\text{penalty OC}} \;+\; \underbrace{\mathcal{P}_\ell^{\text{s}}}_{\text{sanction}} \;+\; \underbrace{\text{Mix}_j}_{\text{mix}} \;+\; \underbrace{\text{Transfer}_j}_{\text{posting transfer}}, \end{aligned} \tag{45}$$

with every term except μ_j^{NE} evaluated at the enforcement allocation and postings. Raising the wage now raises the backpay the firm owes, so recruiting through the wage carries a *wage penalty*, and the markdown and the wage penalty add up to $\mu_j^E - \mu_j^{NE}$. The backpay line collects three effects. One more worker raises the backpay the firm pays. A larger firm is inspected more often, which raises every worker's backpay value and saves the firm recruiting costs. And workers sort partly on backpay, which lowers the wedge. One more worker also raises the penalty opportunity cost and the sanction. Finally, the mix and the posting transfer split Δc_j^{net}, and they act like a per-worker tax: a posting that costs more per util delivered makes the firm stop hiring earlier.

Posting. At the fixed allocation ℓ^E, switching firm j's posting changes the surplus by $-\ell_j^E\, \Delta c_j^{\text{net}}$. Let

$$e(W) \;:=\; \min_{w \geq w^r,\ \lambda \in [0,1]} c(w,\lambda) \qquad \text{s.t.} \qquad \beta \ln w \;+\; \lambda D \;\geq\; W \tag{46}$$

be the least cost of delivering utility W without enforcement, so that $c_j^{NE} = e(W_j^{NE})$. Adding and subtracting $e(W_j^E)$ splits the posting change per worker into the mix and the posting

[20] At a firm that pays the minimum wage, the wage condition does not split μ_j. The compliance condition of eq. (17) instead gives $\eta\mu_j = (c_\lambda + \mathcal{P}_\lambda/\ell_j)/(D + \vartheta B_\lambda)$, the hiring condition holds as stated, and μ_j takes the place of the markdown plus the wage penalty.

transfer,

$$\begin{aligned}
\text{Posting} &= \sum_{j\in\mathcal{J}} \ell_j^E \left(\text{Mix}_j + \text{Transfer}_j\right), \\
\text{Mix}_j &:= c_j^E - e(W_j^E) \;\geq\; 0, \\
\text{Transfer}_j &:= \int_{W_j^{NE}}^{W_j^E} \Big[e'(W) - \frac{1}{\mathsf{m}}\Big] dW \;=\; \big[e(W_j^E) - e(W_j^{NE})\big] - \frac{W_j^E - W_j^{NE}}{\mathsf{m}}.
\end{aligned} \tag{47}$$

The mix is nonnegative, because the enforced firm chooses its wage and compliance partly to reduce its expected penalty rather than to deliver W_j^E at least cost. The *posting transfer* weighs what the firm pays per util, $e'(W)$, against the money value of a util to the workers, $1/\mathsf{m}$. It is distinct from the backpay transfer $\mathcal{P}^{\text{bp}}$, which moves money only when an inspection restores a shortfall.

The posting change enters the welfare change twice: through hiring, as a line of eq. (45) weighted by the employment gap, and on the workers employed, weighted by ℓ_j^E. Grouping the two, the welfare change is

$$\begin{aligned}
\Delta\mathcal{W} \;=\; & -\sum_{j\in\mathcal{J}} \bar{\ell}_j^{\text{gap}} \big[\text{markdown} + \text{wage penalty} + \text{backpay} \\
& \qquad + \text{penalty OC} + \text{sanction } (+\,\text{rationing under PE})\big] && \text{hiring forces} \\
& - \sum_{j\in\mathcal{J}} \left(\bar{\ell}_j^{\text{gap}} + \ell_j^E\right) \text{Mix}_j && \text{posting: mix} \\
& - \sum_{j\in\mathcal{J}} \left(\bar{\ell}_j^{\text{gap}} + \ell_j^E\right) \text{Transfer}_j && \text{posting: transfer} \\
& - \text{NEC}^E && \text{net enforcement cost,}
\end{aligned} \tag{48}$$

where the *hiring forces* are the other lines of eq. (45) weighted by the gap. Table 8 reports this decomposition. Because the posting transfer values every util at the single money metric m, we also report the welfare change without it, $\Delta\mathcal{W}^{\text{no transfer}} := \Delta\mathcal{W} + \sum_{j\in\mathcal{J}} \left(\bar{\ell}_j^{\text{gap}} + \ell_j^E\right) \text{Transfer}_j$.

6.1.4 *Perfect Enforcement*

Under perfect enforcement the backpay value $B_j^{PE} = (1-\lambda_j)D$ does not depend on size, so the backpay slope drops out. Many firms also ration hiring at the minimum wage, hiring fewer

workers than sort to them. We assume that rationing is taste-efficient: among the assignments of workers that produce the realized employment, the one that occurs maximizes total taste utility. The welfare function and the decomposition then hold unchanged (appendix D.5), but hired workers receive V_j while the allocation sorts on the *effective values* $v_j^{\text{eff}} := \mathcal{V}_j(\ell^{PE})$. The gap

$$V_j - v_j^{\text{eff}} \;=\; \underbrace{\ln \frac{\ell_0}{\ell_0^{\text{free}}}}_{\text{common}} \;+\; \underbrace{\ln \frac{N_j}{\ell_j}}_{\text{own}} \;\geq\; 0, \tag{49}$$

where ℓ_0^{free} is the informal employment that free sorting on V would produce, adds a rationing term $(V_j - v_j^{\text{eff}})/\mathsf{m}$ to the wedge. The common part raises every wedge, because turned-away applicants crowd the informal sector, and the own part raises the wedge of each rationing firm by the log of its applicants over its hires. At a firm that rations $\mu_j = 0$, so the markdown, the wage penalty and the backpay slope drop out and are replaced by $-\mu_j^{NE}$ (proposition 6).

6.2 *The Welfare Effects of Enforcement*

We find that looking at the costs alone overstates the welfare loss of the status quo, and that imperfect enforcement avoids most of the loss that perfect enforcement would cause. Both firms and workers slightly benefit from the status quo, while perfect enforcement would hurt firms dramatically with a small gain to workers.

6.2.1 *The Welfare Change from No Enforcement*

Table 8 collects the sources of each welfare change of eq. (48), in w_0 and in log points of wages (eq. (38)). Relative to no enforcement, the status quo changes welfare by 0.27 log points of wages and perfect enforcement by -16, and every figure below is in the same units.

Figure 10 shows the hiring forces by firm size: the lines of eq. (45) other than the mix and the posting transfer, which table 8 groups with posting. In the status quo the penalty opportunity cost widens the wedge and backpay narrows it. Both grow with firm size up to several thousand workers and largely offset, and the markdown and the wage penalty are small, so the net change in the wedge stays near zero below a few hundred workers and rises only at the largest firms. Under perfect enforcement the wedge rises at every size. At small firms rationing raises it several times over and the penalty opportunity cost adds to it, while the markdown falls because firms that ration no longer recruit. Figures 18 and 19 in

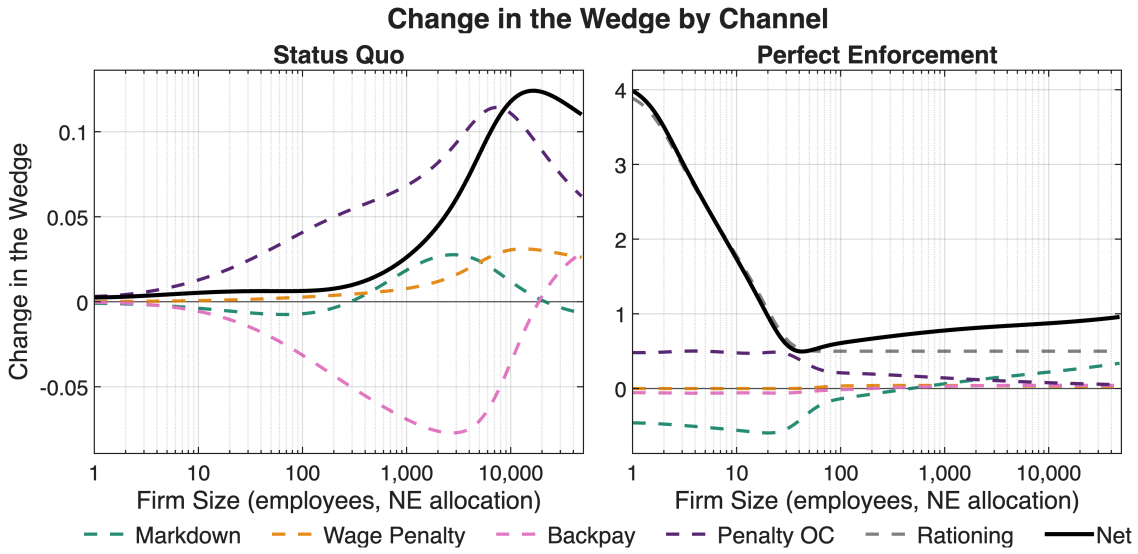


Figure 10: The change in the wedge at the no-enforcement postings, $G_j^E - G_j^{NE}$, by channel, against firm size in the no-enforcement allocation, in w_0 per worker. Left: status quo. Right: perfect enforcement. The dashed lines are the hiring forces of eq. (45) and, under perfect enforcement, the rationing line of eq. (147); Net is their sum, the bracket that eq. (48) weights by the employment gap. The sanction line is near zero and not drawn, and the mix and the posting transfer are in fig. 18. "Penalty OC" is the penalty opportunity cost, the cost of paying backpay and sanctions at the inspection date.

appendix D show posting, the net cost of enforcement and the employment gap by firm size.

Firms also change the wages and compliance they post, which moves the markdown directly. The status quo barely moves it: the markdown line contributes -0.03 log points. At the largest firms, higher wages and lower compliance move the markdown in opposite directions. Under perfect enforcement the markdown line contributes 0.29, because firms that ration no longer recruit. [markdown change by size, split into wage and compliance parts]

The small net effect of the status quo does not mean that the individual components are small. The hiring forces lower welfare by -0.44 log points, mostly through the penalty opportunity cost (-0.5). The three backpay lines partly offset it, 0.21 in total: the backpay the firm pays for one more worker contributes -0.78, while the value of backpay to workers contributes 0.58 through recruiting and 0.41 through sorting. Backpay pulls workers toward the large firms where the markdown is largest, which is the direction the planner moves them. The posting mix worsens welfare (-0.94), as expected, but large firms deliver less utility in the status quo, since workers there are paid partly in backpay. At those firms a util costs more to deliver than it is worth at the mean wage, and the posting transfer therefore

contributes 2.2 log points. Finally, most of the net enforcement cost is the opportunity cost of paying backpay and sanctions at the inspection date (-0.44).

These results show that counting only the costs of enforcement overstates the welfare loss of the status quo. The last panel of table 8 lists the costs of enforcement line by line: the wage penalty, the backpay the firm pays for one more worker, the sanction, the penalty opportunity cost, the posting mix, agency spending and, under perfect enforcement, rationing. Together they give a welfare change of -2.8 log points. These are the costs that the literature has focused on, but they leave out the gains to workers and the role of pre-existing misallocation. The status quo in fact changes welfare by 0.27 log points, or -1.9 without the posting transfer, so the costs alone overstate the loss by as much as 3.1 log points.

At first glance, these regulations look costly in a developing country like Peru, but imperfect enforcement mitigates most of their welfare cost. Indeed, under perfect enforcement, welfare changes by -15% of wages. The hiring forces carry most of the loss (-10 log points), through rationing (-8) and the penalty opportunity cost (-2). Formal firms turn away 7.9 million workers to the informal sector. We measure the share of the perfect-enforcement loss that the status quo avoids as $1 - \Delta\mathcal{W}^{SQ}/\Delta\mathcal{W}^{PE}$, where a value above 100% means that welfare rises under the status quo. The status quo mitigates 102% of the loss. Without the posting transfer, it mitigates 86%.

6.2.2 *The Incidence of the Welfare Change*

One rationale for regulations is that they protect workers, even at the expense of firm owners. We find that the status quo raises welfare for both workers and firm owners, while perfect enforcement hurts firms and helps workers.

Table 19 and fig. 20 in appendix D collect the totals and split the welfare change at each firm size between workers and firm owners. In the status quo, profits increase by 0.23 log points and workers gain 0.05. Workers receive less regulated compensation and amenities as large firms comply less, but higher wages and a higher backpay value more than offset this. Owners pay more in penalties but gain from higher output and lower costs. The gains and losses sort by size: below a size of 30 both firms and their workers lose, and above size 100 both gain.

Under perfect enforcement the increased utility from regulations and backpay offsets the large, negative effects of labor rationing and wage reductions for workers, resulting in a

	Status quo		Perfect enforcement	
	w_0, 000s	wage, log pts	w_0, 000s	wage, log pts
Hiring forces				
Markdown	−8.2	−0.03	72.2	0.29
Wage penalty	−28.4	−0.11	−57.8	−0.23
Backpay: paid for marginal worker	−193.9	−0.78	−700.1	−2.82
Backpay: recruiting	143.4	0.58	–	–
Backpay: sorting	102.6	0.41	688.3	2.77
Penalty opportunity cost	−124.3	−0.50	−484.3	−1.95
Sanction	−1.4	−0.01	−60.7	−0.24
Rationing	–	–	−1,994.4	−8.04
Total	−110.2	−0.44	−2,536.8	−10.23
Posting				
Mix	−233.2	−0.94	−274.7	−1.11
Transfer	544.4	2.19	−784.0	−3.16
Total	311.2	1.25	−1,058.8	−4.27
Net enforcement cost (NEC)				
Penalty opportunity cost	−109.9	−0.44	−387.4	−1.56
Backpay paid less its value	−21.2	−0.09	−11.4	−0.05
Agency spending	−3.7	−0.01	−93.4	−0.38
Total	−134.8	−0.54	−492.2	−1.98
Welfare change	66.2	0.27	−4,087.8	−16.48
Welfare change, no posting transfer	−478.2	−1.93	−3,303.7	−13.32
Costs of enforcement alone				
Wage penalty (hiring)	−28.4	−0.11	−57.8	−0.23
Backpay paid for marginal worker (hiring)	−193.9	−0.78	−700.1	−2.82
Penalty opportunity cost (hiring)	−124.3	−0.50	−484.3	−1.95
Penalty opportunity cost (NEC)	−109.9	−0.44	−387.4	−1.56
Sanction (hiring)	−1.4	−0.01	−60.7	−0.24
Mix (posting)	−233.2	−0.94	−274.7	−1.11
Agency spending (NEC)	−3.7	−0.01	−93.4	−0.38
Rationing (hiring)	–	–	−1,994.4	−8.04
Total	−694.7	−2.80	−4,052.9	−16.34

Table 8: Sources of the welfare change from no enforcement (eq. (48)). Values are in thousands of w_0, and "wage, log pts" is the wage equivalent of eq. (38). The three backpay rows split the backpay line of eq. (45): the backpay the firm pays for one more worker, $\mathcal{P}_\ell^{\text{bp}}$, what that backpay saves the firm in recruiting, $-\eta\mu_j\vartheta\ell_j B_\ell$, and workers' sorting on backpay, $-\vartheta B_j/\mathsf{m}$, each term weighted by $-\bar{\ell}_j^{\text{gap}}$ and summed over firms. The last panel lists the costs of enforcement alone, each row copied from the line named in parentheses ("hiring" is the hiring forces and NEC the net enforcement cost), and its total is the costs-alone welfare change. Rows may not add to totals because of rounding.

welfare gain of 0.28 log points. On the other hand, firm profits decline (-17) largely due to output loss. Workers at firms below ten workers lose on net, because those firms ration the most, while workers at larger firms gain. For a planner to implement perfect enforcement, it would need to weight workers' welfare more than firm owners' and the public's by a factor of 60.

6.2.3 *Welfare across Enforcement Intensities*

The status quo and perfect enforcement are two points on a continuum. We trace welfare between them by scaling the complaint cost scale ρ_0 of eq. (3): multiplying ρ_0 by m divides the mean complaint cost $1/\rho(\ell)$ by m at every firm size, with the regulations, the sanction schedule and the related groups held fixed. No enforcement is the limit $m \to 0$ and perfect enforcement the limit $m \to \infty$ (section 3.4). At each multiplier we solve the general equilibrium and measure its welfare change from no enforcement as in section 6.1.3.[21]

Figure 11 shows that welfare is an inverted U in enforcement intensity. It peaks at $m = 3$, where welfare is 0.36% of wages above no enforcement, against 0.27% in the status quo. Welfare stays above its no-enforcement level up to $m = 10$ and falls below it by $m = 20$. At $m = 100$ it is -2.5% of wages, 15% of the way to the -15% of perfect enforcement. Moderately easier complaining would therefore raise welfare, while the losses of section 6.2.1 under perfect enforcement build up only once complaints become much cheaper than in the status quo.

[21] The grid has 18 multipliers from 0.01 to 100. The equilibrium solver did not converge at $m = 300$, 1,000 and 3,000, and fig. 11 omits those three points.

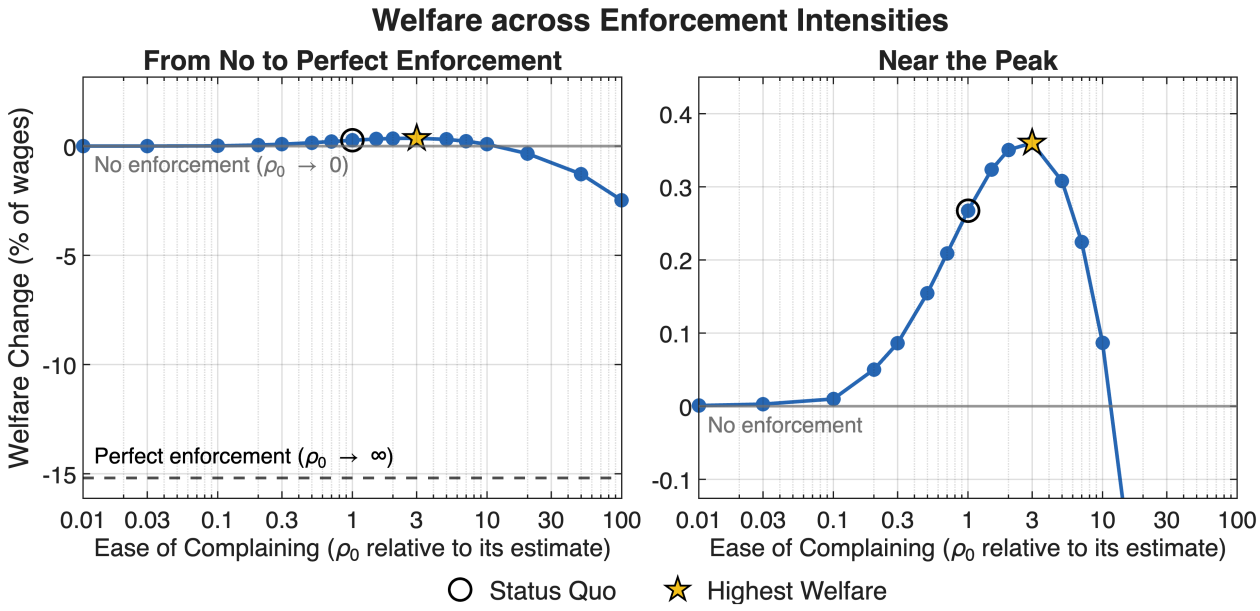


Figure 11: Welfare change from no enforcement against the complaint cost scale ρ_0, as a multiple m of its estimate, in percent of wages. The left panel spans the grid, with perfect enforcement as the dashed line; the right panel zooms in on the peak. The circle is the status quo ($m = 1$) and the star the highest welfare on the grid.

7 Conclusion

This paper asks what the enforcement of labor regulations does to aggregate outcomes when enforcement runs through worker complaints. In our model a firm chooses its compliance, a shorted worker complains when the value of the regulation exceeds the cost of complaining, and inspections deliver backpay to workers. The model gives a closed-form likelihood for Peru's inspection records, which lets us correct for selection into those records. Compliance rises with firm size, from 0.39 at the smallest firms to 0.93 above a thousand workers, and workers pay for it by accepting lower wages.

The welfare accounting turns on the fact that these regulations are enforced against firms that already pay workers less than their marginal product. Relative to no enforcement, the status quo raises welfare by 0.27% of wages, and counting only the costs of enforcement overstates the loss by as much as 3% of wages. Perfect enforcement dramatically reshapes the economy: firms bunch below statutory thresholds, formal employment falls from 25% to 13% of the workforce, and welfare falls by 15% of wages. Labor regulations and their enforcement therefore cannot be evaluated separately. A government with limited capacity

can design enforcement in tandem with labor regulations to seek to improve welfare. However, our analysis does not include other costs of labor regulations and their enforcement on firms, such as the effect on firm entry and effects on the firm life cycle. We leave these for future work.

References

Almeida, R. and P. Carneiro (2009). Enforcement of labor regulation and firm size. *Journal of Comparative Economics 37*(1), 28–46.

Almeida, R. and P. Carneiro (2012). Enforcement of labor regulation and informality. *American Economic Journal: Applied Economics 4*(3), 64–89.

Almeida, R. and L. Ronconi (2016). Labor inspections in the developing world: Stylized facts from the Enterprise Survey. *Industrial Relations: A Journal of Economy and Society 55*(3), 468–489.

Almunia, M. and D. Lopez-Rodriguez (2018). Under the radar: The effects of monitoring firms on tax compliance. *American Economic Journal: Economic Policy 10*(1), 1–38.

Amodio, F. and N. de Roux (2024). Measuring labor market power in developing countries: Evidence from Colombian plants. *Journal of Labor Economics 42*(4), 949–977.

Amodio, F., P. Medina, and M. Morlacco (2025). Labor market power, self-employment, and development. *American Economic Review 115*(9), 3014–3057.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017). Measuring the sensitivity of parameter estimates to estimation moments. *Quarterly Journal of Economics 132*(4), 1553–1592.

Armangué-Jubert, T., N. Guner, and A. Ruggieri (2025). Labor market power and development. *American Economic Review: Insights 7*(2), 177–195.

Azar, J. and I. Marinescu (2024). Monopsony power in the labor market. In *Handbook of Labor Economics*, Volume 5, pp. 761–827. Elsevier.

Bachas, P., R. N. Fattal Jaef, and A. Jensen (2019). Size-dependent tax enforcement and compliance: Global evidence and aggregate implications. *Journal of Development Economics 140*, 203–222.

Berger, D., K. Herkenhoff, and S. Mongey (2022). Labor market power. *American Economic Review 112*(4), 1147–1193.

Berger, D., K. Herkenhoff, and S. Mongey (2025). Minimum wages, efficiency, and welfare. *Econometrica 93*(1), 265–301.

Bergquist, L. F., D. Lashkari, and E. Verhoogen (2026). Wedges: A microeconomic perspective on misallocation. In P. Dupas, P. Goldberg, and R. Pande (Eds.), *Handbook of Development Economics*, Volume 6. Elsevier. Forthcoming.

Bosch, M., G. Cruces, S. González, and M. T. Silva-Porto (2026). Large firms and the intensive margin of labor informality: Evidence from an enforcement intervention in Peru. *Journal of Development Economics 180*, 103679.

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

Brooks, W. J., J. P. Kaboski, Y. A. Li, and W. Qian (2021). Exploitation of labor? classical monopsony power and labor's share. *Journal of Development Economics 150*, 102627.

Brotherhood, L., D. Da Mata, C. Santos, N. Guner, and P. Kircher (2024). Labor market regulation and informality. Idb working paper, Inter-American Development Bank.

Caldwell, S., A. Dube, and S. Naidu (2026). Monopsony makes it big. *Journal of Economic Literature*. Forthcoming.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records. *Quarterly Journal of Economics 126*(2), 749–804.

do Prado, T., M. Santos, and B. Van Doornik (2025). Enforcing compliance with labor regulations and firm outcomes: Evidence from Brazil. *Journal of Development Economics 176*, 103493.

Feinmann, J., R. Hsu Rocha, and M. Lauletta (2026). Payments under the table: Employer-employee collusion in Brazil. Working paper.

Felix, M. (2026). Oligopsony in dual labor markets. Working paper.

Garicano, L., C. Lelarge, and J. Van Reenen (2016). Firm size distortions and the productivity distribution: Evidence from France. *American Economic Review 106*(11), 3439–3479.

Gruber, J. (1994). The incidence of mandated maternity benefits. *American Economic Review 84*(3), 622–641.

Guner, N., G. Ventura, and Y. Xu (2008). Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics 11*(4), 721–744.

Gusti, A. and C. Carvalho (2024). Labor inspections, firm adjustments and employee costs. Working paper.

Harju, J., T. Matikka, and T. Rauhanen (2019). Compliance costs vs. tax incentives: Why do entrepreneurs respond to size-based regulations? *Journal of Public Economics 173*, 139–164.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics 124*(4), 1403–1448.

International Monetary Fund (2024). Peru: Selected issues. IMF Country Report 24/134, International Monetary Fund, Washington, DC.

Kleven, H. J., C. T. Kreiner, and E. Saez (2016). Why can modern governments tax so much? An agency model of firms as fiscal intermediaries. *Economica 83*(330), 219–246.

Kumler, T., E. Verhoogen, and J. Frías (2020). Enlisting employees in improving payroll tax compliance: Evidence from Mexico. *Review of Economics and Statistics 102*(5), 881–896.

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

Levy, S. and G. Cruces (2021). Time for a new course: An essay on social protection and growth in Latin America. Working paper.

Lipsey, R. G. and K. Lancaster (1956). The general theory of second best. *Review of Economic Studies 24*(1), 11–32.

Maloney, W. F. (2004). Informality revisited. *World Development 32*(7), 1159–1178.

Manning, A. (2011). Imperfect competition in the labor market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4, pp. 973–1041. Elsevier.

Mas, A. and A. Pallais (2017). Valuing alternative work arrangements. *American Economic Review 107*(12), 3722–3759.

O'Donoghue, T. and M. Rabin (1999). Doing it now or later. *American Economic Review 89*(1), 103–124.

Pignatti, C. (2026). Labor-law compliance in an informal labor market. *World Bank Economic Review*, lhag019.

Samaniego de la Parra, B. and L. Fernández Bujanda (2024). Increasing the cost of informal employment: Evidence from Mexico. *American Economic Journal: Applied Economics 16*(1), 377–411.

Schuessler, A. A. (2000). *A Logic of Expressive Choice*. Princeton, NJ: Princeton University Press.

Sorkin, I. (2018). Ranking firms using revealed preference. *Quarterly Journal of Economics 133*(3), 1331–1393.

Summers, L. H. (1989). Some simple economics of mandated benefits. *American Economic Review: Papers and Proceedings 79*(2), 177–183.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review 108*(8), 2015–2047.

Viollaz, M. (2018). Are labour inspections effective when labour regulations vary according to the size of the firm? Evidence from Peru. *International Labour Review 157*(2), 213–242.

World Bank (2004). *Doing Business in 2004: Understanding Regulation.* Washington, DC: World Bank and Oxford University Press.

Appendix Contents

A Proofs and Derivations **61**

A.1 Proof of Proposition 1 66

A.2 The Expected Sanction Base 73

A.3 The Backpay Value 76

A.4 Derivation of the Escape Probability and the Expected Complaint Cost . . . 76

A.5 Derivation of the Expected Backpay Cost 77

A.6 Microfoundation of the Opportunity Cost Curve 81

A.7 Worker Utility 84

A.8 Labor Supply Derivation 84

A.9 The Wage Elasticity of Labor Supply 85

A.10 The Firm's Optimality Conditions across Enforcement Regimes 86

B Contextual Details **99**

B.1 Size-Dependent Regulation 99

B.2 Bunching at Twenty Workers 102

C Estimation Details **103**

C.1 Block 1: Statutory and Institutional Parameters 103

C.2 Block 2: The Complaint and Inspection Process 107

C.3 Block 3: Costs, Production, and the Labor Market 114

D Welfare Accounting **119**

D.1 The Welfare Function 119

D.2 The Planner's Problem 122

D.3 The Wedge and the Welfare Change 124

D.4 What Moves the Wedge 127

D.5 Perfect Enforcement and Rationing 130

D.6 Incidence 133
D.7 The Inspection Agency 133
D.8 Results at the Estimate 134

A Proofs and Derivations

Table 9 lists every object of the model in the order it is introduced, with where it is defined and what kind of object it is. A *parameter* is set in Block 1 or estimated in Block 2 or 2 and is fixed across firms; a *variable* is chosen by a firm or a worker or drawn by nature; a *function* is a mapping built from those two; and a *static* object is a set, an index or a constant that the model defines once and never varies. Case distinguishes two objects that would otherwise collide: $d \in \{1, 2, 3\}$ is a firm's observed outcome and $D = \sum_{k \in \mathcal{K}} \Delta_k$ is the utility of the full regulatory package.

Table 9: The objects of the model, in the order they are introduced.

Symbol	**Object**	**Kind**	**Where**
Regulations, complaints and the enforcement funnel			
$\mathcal{K}^w$, $\mathcal{K}^a$, $\mathcal{K}$	Regulated compensation, regulated amenities, and all regulations	Static	section 4.2.1
K^w, K^a	How many regulations are in each set	Parameter	section 4.2.1
δ_k^w	Compensation rate of regulation k, as a fraction of the base wage	Parameter	(81)
$\mathcal{R}^0$, $\mathcal{R}^m$, $\mathcal{R}_k$	Always-checked set, related group m, and the regulations an inspection opens on a complaint about k	Static	section 4.2.2
$\mathcal{D}_k$	The other regulations whose complaint pulls k into the investigation	Static	section 4.2.2
λ	Compliance: the probability the firm honors an obligation	Variable	section 3.2.2
S_{ik}, R_{ik}	Whether worker i is shorted on regulation k, and its complement	Variable	(15)
n_k	Workers shorted on regulation k	Variable	(15)
ω_{ik}	Worker i's cost of complaining about regulation k	Variable	(3)
$\rho(\ell) = \rho_0 \ell^{\rho_1}$	Complaint hazard, with scale ρ_0 and size elasticity ρ_1	Function (ρ_0, ρ_1 parameters)	(3)
m_k, $\psi_k(\lambda)$	Silence probability given a shortfall, and unconditionally	Function	(19), (19)
$\mathcal{T}$	The regulations complained about	Variable	(20)
$a_k(\lambda, \ell)$	Probability that nobody complains about regulation k	Function	(20)
$\mathscr{I}(\lambda, \ell)$	Inspection probability	Function	(21)
$\mathcal{R}$	The surfaced set: investigated regulations carrying a shortfall	Variable	section 4.2.2
$\text{esc}_k(\lambda, \ell)$	Escape probability: a shortfall on k never surfaces	Function	(23)

continued on the next page

Symbol	Object	Kind	Where
$\mathcal{C}$, $\gamma(\ell) = \gamma_0 \ell^{-\gamma_1}$	Resolution capacity and the rate of its exponential law	Variable; function (γ_0, γ_1 parameters)	(24)
r, r^*	Resolution order, and the position at which the window closes	Variable	(25)
$\mathcal{A}$	The recorded set: surfaced regulations left unresolved	Variable	(26)
$\mathcal{N}$, $\sigma(\mathcal{N}, \ell)$	Surfaced burden of shorted workers, and the probability the window absorbs it	Variable; function	(27)
$d \in \{1, 2, 3\}$	Observed outcome: not inspected, inspected, sanctioned	Variable	(55)–(57)
P^1, P^2, P^3	Probabilities of the three outcomes	Function	(55)–(57)
$M_k(x)$; a_k, b_k, d_k	Generator of the shorted count, and the three of its values that carry the likelihood	Function	proposition 1
$f(\omega \mid d = 3)$, f	Pmf of the recorded set given a sanction; regulations resolved in-window	Function; variable	proposition 1
$\chi_k(\ell)$	Whether unit k is citable at size ℓ	Function	appendix B.1
Workers and the labor market			
β	Weight on the log wage in worker utility	Parameter	(80)
β_k^a	Worker's value of regulated amenity k	Parameter	(82)
$\tilde{w}_i$, W_{ij}	Realized wage, and realized utility of worker i at firm j	Variable	(81), (80)
Δ_k, D	Value of regulation k to a worker, and the sum over $\mathcal{K}$	Static	(82)
W_j	Delivered utility: what a posting is worth in expectation	Function	(5)
H_k	Expected complaint cost on regulation k	Function	(72)
$B(\lambda, \ell)$	Backpay value: expected restoration net of the expected complaint cost	Function	(71)
ϑ	Backpay weight: the weight workers place on the backpay value when they sort	Parameter	(4)
V_j	The value workers sort on, $W_j + \vartheta B$	Function	(4)
ε_{ij}, u_{ij}	Idiosyncratic taste for firm j, and total utility of the job	Variable	(1)
N_j, IV	Labor supply to firm j, and the market inclusive value	Function; variable	(2)
η_j	Labor supply elasticity, $1 - N_j/L$	Function	(84)
ξ_0, V_0	Taste for the informal sector, and the value it offers	Parameter; function	section 3.1.1, (16)
L, J, $\mathcal{J}$	Workforce, the number of formal firms, and the set of them	Parameter; static	section 3.3
Production, costs and penalties			
z_j, μ_z, σ_z	Firm productivity, and the Pareto bound and tail index it is drawn from	Variable; parameters	(6)
α, y_j	Returns to scale, and value-added production	Parameter; function	(6)

continued on the next page

Symbol	Object	Kind	Where
w_j, ℓ_j	Base wage and head count, the firm's other two choices	Variable	(7)
w^r	Minimum wage	Parameter	(7)
w_0, z_0, ℓ_0	Informal wage, informal productivity and informal employment	Variable	(16)
$\pi(z_j)$	Profit, which is operating surplus	Function	(7)
$\Lambda(\lambda)$	Expected factor on the base wage from delivered compensation	Function	(10)
c^a	Resource cost of delivering one regulated amenity unit	Parameter	(8)
q, F, q^*	Shadow opportunity cost of honoring an obligation, its distribution, and the firm's threshold	Variable; function	section 3.2.2, (79)
$\mathsf{b}(w)$	Cost of promised obligations at full compliance	Function	(11)
$\mathcal{X}(\lambda)$, ϕ, $\mathcal{X}_1$	Opportunity cost curve, its scale, and its level at full compliance, $\mathcal{X}_1 = \phi$	Function; parameter; static	(11), (12)
$c(w, \lambda)$	Per-worker cost of a posting	Function	(8)
$\Phi(\lambda, \ell)$	Expected wage factor a worker takes home once restoration is complete	Function	(74)
$\mathcal{P}^{\mathrm{bp}}$	Backpay transfer; the firm's outlay on penalties adds the penalty opportunity cost, $(1 + \mathcal{X}_1)(\mathcal{P}^{\mathrm{bp}} + \mathcal{P}^{\mathrm{s}})$	Function	(73), (13)
S_k, g_k, ν, ζ	Sanction on a recorded regulation, with severity, count curvature and size scaling	Function; parameters	(14)
$\mathcal{P}^{\mathrm{s}}$, Z_k	Expected sanction, and the expected sanction base on regulation k	Function	(64), (65)
$\mathcal{P}(w, \lambda, \ell)$	Expected penalty: backpay plus sanction	Function	(9)
Welfare accounting			
$\mathsf{m} = \beta/\bar{w}$	Marginal utility of money at the mean wage $\bar{w}$, so $1/\mathsf{m}$ converts utils to money	Static	section 6.1.1
v_j, T_j	The value workers sort on, and the planner's lump-sum transfer to each worker at firm j, in money	Variable	(132)
$\mathbf{p}_j$, $\mathbf{p}$	Posting: a firm's base wage and compliance (w_j, λ_j), and the postings of all firms, with a regime superscript	Variable	(36)
U, $U^{\mathrm{taste}}(\ell)$	Worker utility, and its part that comes from tastes and the informal option	Function	(135)
$\mathcal{W}$, $\mathcal{F}(\ell; \mathbf{p})$	Welfare, and the surplus: welfare before the net enforcement cost	Function	(35), (36)
NEC	Net enforcement cost	Function	(37)
x	A welfare change as a uniform wage raise, in log points	Function	(38)
$\mathcal{W}^*$, ℓ^*	Welfare and the allocation at the planner benchmark	Static	(136)
$MRPL_j$	Marginal revenue product of labor	Function	(18)

continued on the next page

Symbol	Object	Kind	Where
$\mathcal{V}_j(\ell)$	Inverse labor supply: the value that sorts ℓ_j workers to firm j against ℓ_0 informal	Function	(39)
$G_j(\ell; \mathbf{p})$	Marginal surplus of moving a worker from the informal sector to firm j	Function	(39)
G_j^{NE}, G_j^{E}	The wedge: the marginal surplus at each regime's allocation, at the no-enforcement postings	Static	(41), (42)
$\mathrm{DWL}(\ell)$	Deadweight loss: the distance of ℓ from the planner's optimum	Function	(140)
$\bar{\ell}_j^{\mathrm{gap}}$	Firm j's employment gap: workers the planner would add at j, averaged along the path	Static	(44)
$\Delta c_j^{\mathrm{net}}$	Change in the posting's cost per worker, net of the change in what it delivers; equals $\mathrm{Mix}_j + \mathrm{Transfer}_j$	Static	(145)
$e(W)$, $e'(W)$	Least cost of delivering utility W without enforcement, and the marginal cost of a util	Function	(46)
Mix_j, $\mathrm{Transfer}_j$	The posting change per worker, split into the extra cost of the enforced mix over the least cost of delivering W_j^E, and the cost of the change in delivered utility net of its value at m	Static	(47)
μ_j, κ_j	Multipliers on the labor supply constraint and the minimum wage. μ_j is the recruiting cost: the markdown without enforcement, the markdown plus the wage penalty with it, and zero at a firm that rations. $\eta\mu_j$ is the cost of a util	Variable	(85)
v_j^{eff}, ℓ_0^{free}	Effective value that sorts realized employment under rationing, and informal employment under free sorting on V	Function	(49)
c^{insp}	The inspection agency's cost per inspection	Parameter	(149)

Throughout this appendix we write a product over a set of regulations with the set as a subscript: for a subset $\mathcal{K}' \subseteq \mathcal{K}$,

$$a_{\mathcal{K}'} \;:=\; \prod_{k \in \mathcal{K}'} a_k, \qquad a_{\emptyset} \;:=\; 1,$$

and likewise $b_{\mathcal{K}'}$ and $d_{\mathcal{K}'}$, so that $a_m = a_{\mathcal{R}^m}$ and $b_0 = b_{\mathcal{R}^0}$ as in proposition 1. We write $\mathrm{Bin}_k(n) := \binom{\ell}{n}(1-\lambda)^n \lambda^{\ell-n}$ for the law of the shorted count n_k in eq. (15). A complaint on a regulation of the always-checked set opens only that set, $\mathcal{R}_k = \mathcal{R}^0$ for $k \in \mathcal{R}^0$, so the set $\mathcal{D}_k$ of other regulations whose complaint pulls k into the investigation (section 4.2.2) is $\mathcal{R}^m \setminus \{k\}$

for $k \in \mathcal{R}^m$ and $\mathcal{K} \setminus \{k\}$ for $k \in \mathcal{R}^0$. We write

$$a_{-k} \;:=\; \prod_{j \in \mathcal{D}_k} a_j \tag{50}$$

for the probability that no complaint arrives on any other regulation whose investigation reaches k: $a_{-k} = a_{\mathcal{R}^m \setminus \{k\}}$ for $k \in \mathcal{R}^m$ and $a_{-k} = a_{\mathcal{K} \setminus \{k\}}$ for $k \in \mathcal{R}^0$. Only a regulation with a shorted worker can surface, so $k \in \mathcal{R}$ requires $n_k \geq 1$.

Lemma 1 (Master generators and count laws)**.** *Define the count generator*

$$M_k(x) \;:=\; \mathbb{E}\big[x^{n_k}\big] \;=\; \big(\lambda + (1-\lambda)\, x\big)^{\ell}. \tag{51}$$

For regulation k and any x, the generators restricted to the complaint event $\{k \in \mathcal{T}\}$ and to the surfacing event $\{k \in \mathcal{R}\}$ are

$$\begin{aligned} M_k^{\mathcal{T}}(x) \;&:=\; \mathbb{E}\big[x^{n_k}\mathbb{1}(k \in \mathcal{T})\big] \;=\; M_k(x) - M_k(m_k x), \\ M_k^{\mathcal{R}}(x) \;&:=\; \mathbb{E}\big[x^{n_k}\mathbb{1}(k \in \mathcal{R})\big] \;=\; \big(M_k(x) - M_k(0)\big) - a_{-k}\big(M_k(m_k x) - M_k(0)\big), \end{aligned} \tag{52}$$

with the count laws, for $n \geq 1$,

$$\mathbb{P}\big(n_k = n,\; k \in \mathcal{T}\big) \;=\; \mathrm{Bin}_k(n)\big(1 - m_k^n\big), \qquad \mathbb{P}\big(n_k = n,\; k \in \mathcal{R}\big) \;=\; \mathrm{Bin}_k(n)\big(1 - m_k^n\, a_{-k}\big). \tag{53}$$

Proof. The proof uses three properties of the primitives.

1. $\mathbb{P}(k \notin \mathcal{T} \mid n_k) = m_k^{n_k}$: each of the n_k shorted workers stays silent independently with the silence probability m_k of eq. (19), and a worker who is not shorted has nothing to complain about.

2. $\mathbb{E}[x^{n_k}] = \big(\lambda + (1-\lambda)x\big)^{\ell}$: the probability generating function of the binomial count in eq. (15).

3. The shortfall and complaint draws are independent across regulations. The event that no complaint arrives on the other regulations reaching k is therefore independent of n_k and of the complaints on k, and has probability a_{-k} (eq. (50)).

For the complaint event,

$$\begin{aligned} M_k^{\mathcal{T}}(x) &= \mathbb{E}\big[x^{n_k}\big] - \mathbb{E}\big[x^{n_k}\mathbb{1}(k \notin \mathcal{T})\big] \\ &= \mathbb{E}\big[x^{n_k}\big] - \mathbb{E}\big[x^{n_k}\mathbb{P}(k \notin \mathcal{T} \mid n_k)\big] && \text{law of iterated expectations} \\ &= \mathbb{E}\big[x^{n_k}\big] - \mathbb{E}\big[(m_k x)^{n_k}\big] && \text{Property 1} \\ &= M_k(x) - M_k(m_k x) && \text{Property 2.} \end{aligned}$$

For the surfacing event, a regulation with no shorted worker never surfaces. A regulation with $n_k \geq 1$ fails to surface if and only if no complaint arrives on k and none arrives on any other regulation whose investigation reaches it. By Properties 1 and 3,

$$\mathbb{P}(k \notin \mathcal{R} \mid n_k) \;=\; m_k^{n_k}\, a_{-k} \qquad (n_k \geq 1).$$

We repeat the three lines above on the event $\{n_k \geq 1\}$, where $\mathbb{E}\big[x^{n_k}\mathbb{1}(n_k \geq 1)\big] = M_k(x) - M_k(0)$ because $M_k(0) = \lambda^\ell = \mathbb{P}(n_k = 0)$, and obtain $M_k^{\mathcal{R}}$. Evaluating the two joint probabilities at a single count $n \geq 1$ gives the count laws of eq. (53). □

Scope limits. Setting $M = 1$ and $\mathcal{R}^0 = \emptyset$ puts every regulation in one group, so any complaint audits every regulation and $a_{-k} = a_{\mathcal{K}\setminus\{k\}}$. At the other extreme, singleton groups with an empty always-checked set give $a_{-k} = 1$: an inspection prompted by k never reaches another regulation, and the surfaced set is the set of regulations complained about. The partition we impose sits between the two.

A.1 *Proof of Proposition 1*

Proposition 2 (Likelihood of the Recorded Set)**.** *Suppose that*

(i) each worker is shorted on each regulation independently with probability $1 - \lambda$*, so that the shorted-worker count* n_k *follows eq.* (15)*;*

(ii) a shorted worker complains if and only if $\Delta_k > \omega_{ik}$*, with complaint costs* ω_{ik} *drawn independently across workers and regulations from eq.* (3)*;*

(iii) a complaint on $k \in \mathcal{R}^m$ *triggers an inspection that investigates* $\mathcal{R}_k = \mathcal{R}^m \cup \mathcal{R}^0$*;*

(iv) the firm resolves the surfaced regulations in a random order, with a resolution capacity $\mathcal{C}$ drawn independently from eq. (24), and the regulations it has not reached when the cumulative burden exceeds $\mathcal{C}$ form the recorded set $\mathcal{A}$ of eq. (26);

Then the log-likelihood of the observed outcomes of firms j with compliance λ_j and size ℓ_j is

$$\mathscr{L} \;=\; \sum_{j} \Big[\ln P^{d_j}(\lambda_j, \ell_j) \;+\; \mathbb{1}(d_j = 3)\, \ln f\big(\omega_j \mid d_j = 3; \lambda_j, \ell_j\big) \Big], \tag{54}$$

where $P^d := \mathbb{P}(d)$ are the outcome probabilities and $f(\omega \mid d = 3)$, is the pmf of $\omega = \{n_k\}_{k\in\mathcal{A}}$, with the closed forms below.

Closed form. Let $M_k(x) := \mathbb{E}\big[x^{n_k}\big] = \big(\lambda + (1-\lambda)x\big)^{\ell}$ be the probability generating function of the shorted-worker count in eq. (15), and drop the argument ℓ in $\gamma(\ell)$ for simplicity. Three of its values carry the likelihood: $a_k := M_k(m_k)$, the probability that nobody complains about k; $b_k := M_k(e^{-\gamma})$, the probability that the window absorbs k's shorted workers; and $d_k := M_k(m_k e^{-\gamma})$, the probability of both. Write $a_m := \prod_{k\in\mathcal{R}^m} a_k$ for the product over group m, for $m = 0, \ldots, M$, and b_m, d_m likewise. Then the outcome probabilities are

$$P^1 := \mathbb{P}(d = 1) \;=\; \underbrace{\prod_{k\in\mathcal{K}} a_k}_{\text{no complaint}} \quad , \tag{55}$$

$$P^2 := \mathbb{P}(d = 2) \;=\; \underbrace{b_0}_{\substack{\text{always-checked}\\ \text{set in scope}}} \prod_{m=1}^{M} \Big(\underbrace{a_m}_{\substack{\text{group}\\ \text{never opened}}} + \underbrace{b_m - d_m}_{\substack{\text{group opened,}\\ \text{burden absorbed}}} \Big) \;-\; \underbrace{d_0 \prod_{m=1}^{M} a_m}_{\text{no inspection}}, \tag{56}$$

$$P^3 := \mathbb{P}(d = 3) \;=\; 1 - P^1 - P^2. \tag{57}$$

For a sanctioned firm ($d = 3$) let $\omega = \{n_k\}_{k\in\mathcal{A}}$ be the firm-wide counts $n_k \geq 1$ on the recorded set $\mathcal{A}$ of eq. (26), with $\mathsf{a} := |\mathcal{A}|$. The joint probability of the recorded set and the sanction is

$$\mathbb{P}(\omega,\, d = 3; \lambda, \ell) \;=\; \underbrace{\sum_{j\in\mathcal{A}} \big(1 - e^{-\gamma n_j}\big)}_{\text{window closes inside } \mathcal{A}} \;\; \sum_{\mathsf{f}\geq 0} \underbrace{\varpi(\mathsf{f}, \mathsf{a})}_{\text{order weight}} \;\; \underbrace{\Pi_{\mathsf{f}}(\lambda, \ell)}_{\mathsf{f} \text{ resolved in-window}} \quad ,$$

and the pmf of the recorded set given a sanction is $f(\omega \mid d = 3) := \mathbb{P}(\omega,\, d = 3; \lambda, \ell)/P^3$, with order weight $\varpi(\mathsf{f}, \mathsf{a}) = \mathsf{f}!\,(\mathsf{a} - 1)!/(\mathsf{f} + \mathsf{a})!$ and Π_{f} constructed below.

Proof. **The likelihood.** A firm contributes the probability of its observed outcome d and, if it is sanctioned, the probability of its recorded set given the sanction, $\mathbb{P}(d)\, f(\omega \mid d=3)^{\mathbb{1}(d=3)}$. The shortfalls, complaint costs, orders and capacities of (i)–(iv) are drawn independently across firms, so the log-likelihood sums these contributions over firms, which is eq. (28). It remains to derive the closed forms of P^d (Part 1) and of $f(\omega \mid d=3)$ (Part 2).

Part 1 (outcome probabilities).

No inspection (P^1). A firm is not inspected if and only if no regulation is complained about, $\mathcal{T}=\emptyset$. The events $\{k \notin \mathcal{T}\}$ are independent across regulations (Property 3 of lemma 1) and $\mathbb{P}(k \notin \mathcal{T}) = a_k$ by eq. (20), so $P^1 = \prod_{k\in\mathcal{K}} a_k$, which is eq. (55). The related groups play no role here: they govern what an inspection reaches, not whether one happens.

Reduction of P^2. An inspected firm avoids a sanction if and only if its capacity $\mathcal{C}$ covers the whole surfaced burden $\mathcal{N}$. The capacity is independent of the shortfall and complaint draws, so given those draws the event has probability $\sigma(\mathcal{N},\ell) = e^{-\gamma\mathcal{N}}$ (eq. (27)). By the law of iterated expectations,

$$P^2 \;=\; \mathbb{E}\big[\mathbb{1}(\mathcal{T}\neq\emptyset)\, e^{-\gamma\mathcal{N}}\big]. \tag{58}$$

Step 1: the burden decomposes by group. Write $\mathcal{N}_g := \sum_{k\in\mathcal{R}^g} n_k$ for the shorted workers in group g, including the always-checked set $g=0$, and $O_m := \{\mathcal{T}\cap\mathcal{R}^m \neq \emptyset\}$ for the event that group $m\geq 1$ is *opened*, that is, some regulation in it draws a complaint. An opened group brings all of its shorted workers into the burden, the always-checked set brings its shorted workers in whenever the firm is inspected, and a regulation with no shorted worker adds nothing. Hence

$$\mathcal{N} \;=\; \sum_{m=1}^{M} \mathbb{1}(O_m)\,\mathcal{N}_m \;+\; \mathbb{1}(\mathcal{T}\neq\emptyset)\,\mathcal{N}_0. \tag{59}$$

Step 2: the weight of one group. If group m is not opened, which happens with probability a_m, it contributes nothing to $\mathcal{N}$ and its window factor is one. If it is opened, its window factor is $e^{-\gamma\mathcal{N}_m}$, and

$$\mathbb{E}\big[e^{-\gamma\mathcal{N}_m}\mathbb{1}(O_m)\big] \;=\; \mathbb{E}\big[e^{-\gamma\mathcal{N}_m}\big] - \mathbb{E}\big[e^{-\gamma\mathcal{N}_m}\mathbb{1}(O_m^c)\big] \;=\; b_m - d_m.$$

The first expectation is $\prod_{k\in\mathcal{R}^m}\mathbb{E}[e^{-\gamma n_k}] = \prod_{k\in\mathcal{R}^m} M_k(e^{-\gamma}) = b_m$ by Property 2 of lemma 1. The second is $\prod_{k\in\mathcal{R}^m}\mathbb{E}\big[e^{-\gamma n_k}\mathbb{1}(k\notin\mathcal{T})\big] = \prod_{k\in\mathcal{R}^m} M_k(m_k e^{-\gamma}) = d_m$ by Property 1, since O_m^c is the intersection of the no-complaint events of the group's regulations. The weight of the

group is therefore

$$\mathbb{E}\big[e^{-\gamma \mathbb{1}(O_m)\mathcal{N}_m}\big] \;=\; \underbrace{a_m}_{\text{group never opened}} \;+\; \underbrace{b_m - d_m}_{\text{group opened, burden absorbed}},$$

which is the bracket of eq. (56).

Step 3: the always-checked set and the one inadmissible pattern. The always-checked set is in scope if and only if $\mathcal{T} \neq \emptyset$, an event that depends on every group, so it cannot enter as an independent factor. We instead price it as if it were always in scope and correct the one pattern this misprices:

$$\mathbb{1}(\mathcal{T} \neq \emptyset)\, e^{-\gamma \mathcal{N}} \;=\; e^{-\gamma \mathcal{N}_0} \prod_{m=1}^{M} e^{-\gamma \mathbb{1}(O_m)\mathcal{N}_m} \;-\; \mathbb{1}(\mathcal{T} = \emptyset)\, e^{-\gamma \mathcal{N}_0}.$$

The identity holds because no group is opened on $\{\mathcal{T} = \emptyset\}$, so the product equals one there and the two terms cancel; on $\{\mathcal{T} \neq \emptyset\}$ the second term vanishes and the first is $e^{-\gamma \mathcal{N}}$ by eq. (59). The draws are independent across groups, so the first term has expectation $b_0 \prod_{m=1}^{M} (a_m + b_m - d_m)$ by Step 2. The event $\{\mathcal{T} = \emptyset\}$ is the intersection of the no-complaint events of every regulation, so the second term has expectation $\mathbb{E}\big[\prod_{k \in \mathcal{R}^0} e^{-\gamma n_k} \mathbb{1}(k \notin \mathcal{T})\big] \prod_{m=1}^{M} a_m = d_0 \prod_{m=1}^{M} a_m$.

Step 4: assemble. Subtracting the two expectations gives eq. (56), and $P^3 = 1 - P^1 - P^2$ because the three outcomes partition. This proves Part 1.

Part 2 (recorded-set pmf). We compute the joint probability $\mathbb{P}(\omega,\, d = 3)$ of a given recorded set $\mathcal{A}$ with counts ω, and $f(\omega \mid d = 3)$ follows on dividing by P^3. The outcome $d = 3$ is the intersection of two events: the firm is inspected, $\mathcal{T} \neq \emptyset$, and the window closes before the surfaced burden is resolved, $\mathcal{C} < \mathcal{N}$. Both are implied by the recorded set being the nonempty set $\mathcal{A}$. A recorded regulation is surfaced, so the firm was inspected, and it is unresolved, so the window closed. Conversely, an inspected firm whose window closes records at least one regulation. Hence

$$\mathbb{P}(\omega,\, d = 3) \;=\; \mathbb{P}\big(\omega,\; \mathcal{T} \neq \emptyset,\; \mathcal{C} < \mathcal{N}\big) \;=\; \mathbb{P}\big(\text{the recorded set is } \mathcal{A} \text{ with counts } \omega\big),$$

and the two events are carried by two different parts of the computation. The sanction $\mathcal{C} < \mathcal{N}$ enters through the capacity in Step 1: the window bracket of eq. (63) is the probability that

the capacity runs out exactly at the first recorded regulation, which cannot happen unless $\mathcal{C} < \mathcal{N}$. The inspection $\mathcal{T} \neq \emptyset$ enters through the complaint patterns in Steps 2 to 4: every group that holds a recorded regulation must have drawn a complaint, and the one pattern with no complaint anywhere is subtracted. Neither event needs a separate factor.

The data show the recorded set and its counts but not the surfaced set, since a regulation resolved in-window leaves no trace. The latent objects to sum over are therefore the *resolved set* $\mathcal{F} := \mathcal{R} \setminus \mathcal{A}$ and its counts. We condition first on the shortfall and complaint draws, which fix $\mathcal{T}$, $\mathcal{R}$ and every count, and use that the capacity $\mathcal{C}$ and the order r are independent of them. Write $\Theta_{\mathcal{A}} := \sum_{j \in \mathcal{A}} \left(1 - e^{-\gamma n_j}\right)$ for the window bracket of eq. (63).

Step 1: the order and capacity factor. Fix $\mathcal{R} = \mathcal{A} \cup \mathcal{F}$ with $\mathsf{f} := |\mathcal{F}|$, and fix the counts on $\mathcal{R}$. Write $\mathcal{N}_{\mathcal{F}} := \sum_{k \in \mathcal{F}} n_k$ for the resolved burden. By eqs. (25) and (26), the recorded set equals $\mathcal{A}$ if and only if the order places all of $\mathcal{F}$ before all of $\mathcal{A}$, some $j \in \mathcal{A}$ comes first among $\mathcal{A}$, and the capacity satisfies $\mathcal{N}_{\mathcal{F}} \leq \mathcal{C} < \mathcal{N}_{\mathcal{F}} + n_j$. Neither the order within $\mathcal{F}$ nor the order of $\mathcal{A}$ after j matters. For each first recorded regulation j there are $\mathsf{f}!\,(\mathsf{a}-1)!$ favorable orders out of $(\mathsf{f}+\mathsf{a})!$, and the capacity condition has probability $e^{-\gamma \mathcal{N}_{\mathcal{F}}} - e^{-\gamma(\mathcal{N}_{\mathcal{F}} + n_j)} = e^{-\gamma \mathcal{N}_{\mathcal{F}}}\left(1 - e^{-\gamma n_j}\right)$ because $\mathcal{C}$ is exponential with rate γ. Summing over j,

$$\mathbb{P}\big(\mathcal{A} \mid \mathcal{R},\ \text{counts}\big) \;=\; \underbrace{\varpi(\mathsf{f},\mathsf{a})}_{\text{order weight}} \quad \underbrace{\Big(\prod_{k \in \mathcal{F}} e^{-\gamma n_k}\Big)}_{\text{resolved burden absorbed}} \quad \underbrace{\Theta_{\mathcal{A}}}_{\text{window closes inside } \mathcal{A}}, \tag{60}$$

so $\varpi(\mathsf{f},\mathsf{a})$ is the probability that the f resolved regulations come first and a given recorded regulation comes next. Summing eq. (60) over every nonempty $\mathcal{A} \subseteq \mathcal{R}$ returns $\mathbb{P}(\mathcal{C} < \mathcal{N}) = 1 - e^{-\gamma \mathcal{N}}$: the order spreads the sanction across recorded sets without changing its probability. The resolved burden enters only through $\prod_{k \in \mathcal{F}} e^{-\gamma n_k}$, which factors across regulations because the capacity is exponential.

Step 2: admissible complaint patterns. A regulation surfaces if and only if it has a shorted worker and its group is in scope. Every group that holds a recorded regulation must therefore have been opened. A group that holds none may be opened or not, and the always-checked set may draw complaints or not. The one inadmissible pattern is the one in which no group is opened and no complaint arrives on the always-checked set, since then there is no inspection. A regulation in scope with $n_k \geq 1$ that is not recorded was resolved, so it belongs to $\mathcal{F}$.

Step 3: the sum over patterns factors by group. The shortfall and complaint draws are independent across regulations, and $\prod_{k \in \mathcal{F}} e^{-\gamma n_k}$ is a product over regulations. The sum over

admissible patterns therefore factors into one polynomial per group, in a formal variable Ξ that marks a resolved regulation. Within a group in scope, each regulation carries a *full weight*, summed over all of its complaint patterns, and a *silent weight*, on the pattern in which nobody complains about it:

	full weight	silent weight
$k \in \mathcal{A}$ (recorded, count observed) :	$\mathrm{Bin}_k(n_k)$,	$\mathrm{Bin}_k(n_k)\, m_k^{n_k}$,
$k \notin \mathcal{A}$ (count summed over) :	$\lambda^\ell + \big(b_k - \lambda^\ell\big)\Xi$,	$\lambda^\ell + \big(d_k - \lambda^\ell\big)\Xi$.

The recorded row is the count law at the observed n_k, and its silent weight carries $\mathbb{P}(k \notin \mathcal{T} \mid n_k) = m_k^{n_k}$ (Property 1 of lemma 1), so full less silent is the count law $\mathbb{P}(n_k = n,\ k \in \mathcal{T})$ of eq. (53) at $n = n_k$. In the other row, $\lambda^\ell = \mathbb{P}(n_k = 0)$ is the weight of a regulation with no shorted worker, which cannot surface and so is never resolved, and it carries no Ξ. The coefficient $b_k - \lambda^\ell = \mathbb{E}\big[e^{-\gamma n_k}\mathbb{1}(n_k \geq 1)\big]$ is the weight of a regulation that surfaces and is resolved, with its window factor, and $d_k - \lambda^\ell = \mathbb{E}\big[e^{-\gamma n_k}\mathbb{1}(n_k \geq 1)\mathbb{1}(k \notin \mathcal{T})\big]$ is the same weight on the event that nobody complains about k. Collecting the rows over group g, with recorded part $\mathcal{A}^g := \mathcal{A} \cap \mathcal{R}^g$,

$$\begin{aligned} \Pi_g^1(\Xi) &:= \prod_{k \in \mathcal{A}^g} \mathrm{Bin}_k(n_k) \prod_{k \in \mathcal{R}^g \setminus \mathcal{A}} \big(\lambda^\ell + (b_k - \lambda^\ell)\,\Xi\big), \\ \Pi_g^2(\Xi) &:= \prod_{k \in \mathcal{A}^g} \mathrm{Bin}_k(n_k)\, m_k^{n_k} \prod_{k \in \mathcal{R}^g \setminus \mathcal{A}} \big(\lambda^\ell + (d_k - \lambda^\ell)\,\Xi\big). \end{aligned} \tag{61}$$

Π_g^1 sums the weights over every complaint pattern of group g, and Π_g^2 is the weight of the pattern in which the whole group is silent.

Step 4: group weights and the inadmissible pattern. An opened group has weight $\Pi_g^1(\Xi) - \Pi_g^2(\Xi)$: all of its complaint patterns less the one in which the whole group is silent. This is where the inspection enters for a group that holds a recorded regulation. A group with no recorded regulation may instead stay closed, with weight a_g and nothing resolved. The always-checked set is in scope on every inspection, so its weight is $\Pi_0^1(\Xi)$ over both of its complaint patterns. The group weights are therefore

$$\Upsilon_g(\Xi) := \Pi_g^1(\Xi) - \Pi_g^2(\Xi) + \mathbb{1}\big(\mathcal{A}^g = \emptyset\big)\, a_g \quad (g \geq 1), \qquad \Upsilon_0(\Xi) := \Pi_0^1(\Xi).$$

The product $\prod_{g=0}^{M} \Upsilon_g(\Xi)$ counts one pattern too many: the one in which every group stays

closed and the whole always-checked set is silent, so that $\mathcal{T} = \emptyset$ and there is no inspection. That pattern is possible only when every recorded regulation lies in the always-checked set, and its weight in the product is $\big(\prod_{m=1}^{M} a_m\big)\Pi_0^2(\Xi)$. Removing it gives the joint polynomial

$$\Pi(\Xi) \;:=\; \prod_{g=0}^{M} \Upsilon_g(\Xi) \;-\; \mathbb{1}\big(\mathcal{A} \subseteq \mathcal{R}^0\big)\Big(\prod_{m=1}^{M} a_m\Big)\,\Pi_0^2(\Xi), \tag{62}$$

of degree at most $|\mathcal{K}| - \mathsf{a}$, since each regulation outside $\mathcal{A}$ contributes one affine factor. Its coefficient on Ξ^{f} is Π_{f}: the probability of the counts ω and of a complaint pattern that surfaces exactly f other regulations, weighted by $\prod_{k\in\mathcal{F}} e^{-\gamma n_k}$.

Step 5: assemble. The order weight of eq. (60) depends on $\mathcal{F}$ only through $\mathsf{f} = |\mathcal{F}|$, and the window bracket $\Theta_{\mathcal{A}}$ depends only on the observed counts. Grouping the sum over patterns by f and reading off the coefficients of $\Pi(\Xi)$ gives, with $\mathsf{a} := |\mathcal{A}|$,

$$\mathbb{P}(\omega,\, d=3;\lambda,\ell) \;=\; \underbrace{\sum_{j\in\mathcal{A}}\big(1-e^{-\gamma n_j}\big)}_{\text{window closes inside }\mathcal{A}} \;\sum_{\mathsf{f}\geq 0}\; \underbrace{\varpi(\mathsf{f},\mathsf{a})}_{\text{order weight}} \;\; \underbrace{\Pi_{\mathsf{f}}(\lambda,\ell)}_{\mathsf{f}\text{ resolved in-window}}, \tag{63}$$

and dividing by P^3 gives $f(\omega \mid d=3)$. The order weight $\varpi(\mathsf{f},\mathsf{a})$ is the probability that the f resolved regulations come first in the order and a given recorded regulation comes next, and Π_{f} is the probability of the counts ω and of a complaint pattern that surfaces exactly f other regulations, each weighted by the probability that the window absorbs its burden. No coefficient is negative. Every full weight dominates its silent weight coefficient by coefficient, since $b_k \geq d_k$ and $m_k^{n_k} \leq 1$, and all coefficients are nonnegative, so $\Pi_g^1 - \Pi_g^2$ has nonnegative coefficients; the last subtraction in eq. (62) removes a pattern that the product already counts. Summing the result over every nonempty $\mathcal{A}$ and every ω returns P^3 of Part 1: the recorded sets and their counts partition the sanction event. □

Special cases. With a single regulation in one group (we drop the subscript k), the recorded set is $\mathcal{A} = \{1\}$, only $\mathsf{f} = 0$ survives with order weight $\varpi(0,1) = 1$, and eq. (63) reduces to $\mathrm{Bin}(n)\big(1-m^n\big)\big(1-e^{-\gamma n}\big)$: the count law, times the probability that somebody complains, times the probability that the window closes. Summed over $n \geq 1$ this is $1 - a - b + d$, which is P^3 from Part 1 with $b_0 = d_0 = 1$. With one group and an empty always-checked set, $\Pi(\Xi) = \Pi_1^1(\Xi) - \Pi_1^2(\Xi)$ is a difference of two products over all regulations, which is the likelihood without related groups.

A.2 *The Expected Sanction Base*

The expected sanction of section 3.2.5 is a function of the firm's compliance and size,

$$\mathcal{P}^{\mathrm{s}}(\lambda,\ell) \;=\; \ell^{\zeta}\sum_{k\in\mathcal{K}} g_k\, Z_k(\lambda,\ell), \tag{64}$$

where $Z_k(\lambda,\ell)$ is the expected sanction base: the probability that k is sanctioned, weighted by the factor $1-e^{-\nu n_k}$.

The sanction on a recorded regulation k is S_k of eq. (14), evaluated at the shorted count n_k of eq. (15). We write $\varsigma_k := 1-e^{-\nu n_k}$ for the count factor, so the expected sanction base in eq. (64) is $Z_k(\lambda,\ell) := \mathbb{E}\big[\mathbb{1}(k\in\mathcal{A})\,\varsigma_k\big]$, where the event $\{k\in\mathcal{A}\}$ already carries a failed resolution, $\mathcal{C}<\mathcal{N}$ (eq. (27)). This establishes its closed form. A regulation is recorded when it has shorted workers, it surfaces, and the window closes before the firm reaches it. Splitting the indicator along those three events,

$$Z_k(\lambda,\ell) \;=\; \underbrace{\mathbb{E}\big[\varsigma_k\,\mathbb{1}(n_k\ge 1)\big]}_{\text{base on } k\text{'s own count}} \;-\; \underbrace{\mathbb{E}\big[\varsigma_k\,\mathbb{1}(n_k\ge 1,\,k\notin\mathcal{R})\big]}_{k \text{ never surfaces}} \;-\; \underbrace{\mathbb{E}\big[\varsigma_k\,\mathbb{1}(k\in\mathcal{R}\setminus\mathcal{A})\big]}_{k \text{ surfaces, resolved in-window}}\;. \tag{65}$$

Every term is an expectation over the same shorted count and the same complaint and resolution draws, and each is an evaluation of the count generator M_k of lemma 1, because the count factor $\varsigma_k = 1-e^{-\nu n_k}$ of eq. (14) is itself a difference of two powers of n_k. We therefore extend the three values a_k, b_k, d_k of proposition 1 by the tilt $e^{-\nu}$ of the count factor,

$$a_k^{(\nu)} \;:=\; M_k\big(m_k e^{-\nu}\big), \qquad b_k^{(\nu)} \;:=\; M_k\big(e^{-(\gamma+\nu)}\big), \qquad d_k^{(\nu)} \;:=\; M_k\big(m_k e^{-(\gamma+\nu)}\big),$$

so that the four moments of the count factor we need are

$$\begin{aligned} \mathbb{E}\big[\varsigma_k\big] \;&=\; 1-M_k\big(e^{-\nu}\big), & \mathbb{E}\big[e^{-\gamma n_k}\varsigma_k\big] \;&=\; b_k-b_k^{(\nu)},\\ \mathbb{E}\big[\mathbb{1}(k\notin\mathcal{T})\,\varsigma_k\big] \;&=\; a_k-a_k^{(\nu)}, & \mathbb{E}\big[\mathbb{1}(k\notin\mathcal{T})\,e^{-\gamma n_k}\varsigma_k\big] \;&=\; d_k-d_k^{(\nu)}. \end{aligned} \tag{66}$$

The first row prices the count factor on k's own count, with and without the window factor $e^{-\gamma n_k}$, and the second row prices it on the event that nobody complains about k. Next we extend b_j and d_j to a *resolution position* $\tilde{r}\in[0,1]$, the position of a regulation in the

resolution order r read on the unit interval,

$$b_j(\tilde{r}) \;:=\; (1-\tilde{r}) + \tilde{r}\, b_j, \qquad d_j(\tilde{r}) \;:=\; (1-\tilde{r})\, a_j + \tilde{r}\, d_j, \tag{67}$$

so that $b_j(1) = b_j$, $d_j(1) = d_j$, $b_j(0) = 1$ and $d_j(0) = a_j$. Write

$$\bar{\sigma}_m(\tilde{r}) \;:=\; a_m + b_{\mathcal{R}^m}(\tilde{r}) - d_{\mathcal{R}^m}(\tilde{r})$$

for the weight of group m at that position: the expected window factor of the group's burden that lies ahead of position $\tilde{r}$, with $\bar{\sigma}_m(1) = a_m + b_m - d_m$ the bracket of eq. (56). The *burden operator* carries the burden resolved ahead of k,

$$\Gamma_k(\tilde{r}) \;:=\; \begin{cases} b_{\mathcal{R}^0}(\tilde{r}) \displaystyle\prod_{m' \neq m} \bar{\sigma}_{m'}(\tilde{r}) \Big(\big(b_k - b_k^{(\nu)}\big)\, b_{\mathcal{R}^m \setminus \{k\}}(\tilde{r}) - \big(d_k - d_k^{(\nu)}\big)\, d_{\mathcal{R}^m \setminus \{k\}}(\tilde{r}) \Big), & k \in \mathcal{R}^m, \\ \big(b_k - b_k^{(\nu)}\big)\, b_{\mathcal{R}^0 \setminus \{k\}}(\tilde{r}) \displaystyle\prod_{m=1}^{M} \bar{\sigma}_m(\tilde{r}) \;-\; \big(d_k - d_k^{(\nu)}\big)\, d_{\mathcal{R}^0 \setminus \{k\}}(\tilde{r}) \prod_{m=1}^{M} a_m, & k \in \mathcal{R}^0. \end{cases} \tag{68}$$

The three terms of eq. (65) are then $\mathbb{E}[\varsigma_k]$, $\mathbb{E}\big[\mathbb{1}(k \notin \mathcal{T})\varsigma_k\big]\, a_{-k}$ and $\int_0^1 \Gamma_k(\tilde{r})\, d\tilde{r}$, so

$$Z_k \;=\; \underbrace{1 - M_k\big(e^{-\nu}\big)}_{\text{base on } k\text{'s own count}} \;-\; \underbrace{\big(a_k - a_k^{(\nu)}\big)\, a_{-k}}_{k \text{ never surfaces}} \;-\; \underbrace{\int_0^1 \Gamma_k(\tilde{r})\, d\tilde{r}}_{k \text{ surfaces, resolved in-window}} . \tag{69}$$

Proof. Step 1: recording given the order. Burdens are nonnegative, so the cumulative burden rises along the order, and a surfaced regulation k is recorded if and only if the cumulative burden through its own position exceeds the capacity: $k \in \mathcal{A}$ if and only if $k \in \mathcal{R}$ and $\mathcal{C} < n_k + \sum_{j \in \mathcal{R} \setminus \{k\}} n_j \mathbb{1}\big(r(j) < r(k)\big)$. Averaging over the exponential capacity $\mathcal{C}$, a surfaced regulation is resolved with probability

$$\mathbb{P}\big(k \notin \mathcal{A} \mid r,\ \text{counts},\ \mathcal{R}\big) \;=\; e^{-\gamma n_k} \prod_{j \in \mathcal{R} \setminus \{k\}} e^{-\gamma n_j \mathbb{1}(r(j) < r(k))},$$

so the third term of eq. (65) is $\mathbb{E}\big[\mathbb{1}(k \in \mathcal{R})\, \varsigma_k\, e^{-\gamma n_k} \prod_{j \in \mathcal{R} \setminus \{k\}} e^{-\gamma n_j \mathbb{1}(r(j) < r(k))}\big]$, an expectation over the counts, the complaints and the order.

Step 2: the average over orders. Every surfaced regulation draws an independent uniform mark on $[0, 1]$, and r ranks the surfaced set by these marks. Given that k's mark is $\tilde{r}$, each

other surfaced regulation is ahead of k independently with probability $\tilde{r}$, so

$$\mathbb{E}_r\Big[\prod_{j\in\mathcal{R}\setminus\{k\}} e^{-\gamma n_j \mathbb{1}(r(j)<r(k))}\Big] \;=\; \int_0^1 \prod_{j\in\mathcal{R}\setminus\{k\}} \Big((1-\tilde{r}) + \tilde{r}\, e^{-\gamma n_j}\Big)\, d\tilde{r}. \tag{70}$$

The mark $\tilde{r}$ is k's resolution position, and $\tilde{r}=1$ puts every other surfaced regulation ahead of k.

Step 3: the first two terms. The count factor ς_k is zero at $n_k=0$, so the first term of eq. (65) is $\mathbb{E}[\varsigma_k]$. Given $n_k\geq 1$, regulation k fails to surface if and only if nobody complains about k and no complaint arrives on the other regulations whose investigation reaches it (lemma 1). These events are independent, so the second term is $\mathbb{E}\big[\mathbb{1}(k\notin\mathcal{T})\varsigma_k\big]\, a_{-k}$.

Step 4: the third term factors by group. By Steps 1 and 2 the third term is the integral over $\tilde{r}$ of

$$\mathbb{E}\Big[\mathbb{1}(k\in\mathcal{R})\,\varsigma_k\, e^{-\gamma n_k} \prod_{j\in\mathcal{R}\setminus\{k\}} \big((1-\tilde{r}) + \tilde{r}e^{-\gamma n_j}\big)\Big].$$

A regulation outside $\mathcal{R}$ carries no burden, and a regulation in scope with $n_j=0$ has factor one, so each factor depends only on whether the regulation's group is in scope. By Property 2 of lemma 1, $\mathbb{E}\big[(1-\tilde{r})+\tilde{r}e^{-\gamma n_j}\big] = b_j(\tilde{r})$, and by Property 1, $\mathbb{E}\big[\mathbb{1}(j\notin\mathcal{T})\big((1-\tilde{r})+\tilde{r}e^{-\gamma n_j}\big)\big] = d_j(\tilde{r})$. A group m' that does not contain k contributes $a_{m'}$ if it is closed and $b_{\mathcal{R}^{m'}}(\tilde{r}) - d_{\mathcal{R}^{m'}}(\tilde{r})$ if it is opened, a total of $\bar{\sigma}_{m'}(\tilde{r})$, exactly as in Step 2 of Part 1 of the proof of proposition 1. The group of $k\in\mathcal{R}^m$ must be opened, so it contributes $\mathbb{E}\big[e^{-\gamma n_k}\varsigma_k\big]\, b_{\mathcal{R}^m\setminus\{k\}}(\tilde{r})$ less the same weight on the event that the whole group is silent, $\mathbb{E}\big[\mathbb{1}(k\notin\mathcal{T})e^{-\gamma n_k}\varsigma_k\big]\, d_{\mathcal{R}^m\setminus\{k\}}(\tilde{r})$. The always-checked set is then in scope and contributes $b_{\mathcal{R}^0}(\tilde{r})$. For $k\in\mathcal{R}^0$, we price every regulation of $\mathcal{R}^0$ as in scope and subtract the pattern in which nobody complains at all, in which every group is closed, with weight $\prod_{m=1}^M a_m$, and the always-checked set is silent, exactly as in Step 3 of Part 1 of that proof. Collecting the cases and substituting eq. (66) gives $\Gamma_k(\tilde{r})$ of eq. (68).

Step 5: assemble. The four moments in eq. (66) follow from Properties 1 and 2 of lemma 1. For example, $\mathbb{E}\big[e^{-\gamma n_k}\varsigma_k\big] = \mathbb{E}\big[e^{-\gamma n_k}\big] - \mathbb{E}\big[e^{-(\gamma+\nu)n_k}\big] = b_k - b_k^{(\nu)}$, and $\mathbb{E}\big[\mathbb{1}(k\notin\mathcal{T})e^{-\gamma n_k}\varsigma_k\big] = \mathbb{E}\big[m_k^{n_k}e^{-\gamma n_k}\big] - \mathbb{E}\big[m_k^{n_k}e^{-(\gamma+\nu)n_k}\big] = d_k - d_k^{(\nu)}$. Combining Steps 3 and 4, and exchanging expectation and integral since the integrand is bounded, gives eq. (69). □

Each per-regulation factor $b_j(\tilde{r})$ and $d_j(\tilde{r})$ in eq. (68) is affine in $\tilde{r}$, so Γ_k is a polynomial in $\tilde{r}$ of degree at most $|\mathcal{K}|-1$, and the $\lceil|\mathcal{K}|/2\rceil$-point Gauss–Legendre rule on $[0,1]$ that the

estimation code applies to this integral evaluates it exactly. Two limits check eq. (69). When resolution always succeeds ($\gamma \to 0$), $b_j(\tilde{r}) \to 1$, $d_j(\tilde{r}) \to a_j$, $\bar{\sigma}_m(\tilde{r}) \to 1$, $b_k^{(\nu)} \to M_k(e^{-\nu})$ and $d_k^{(\nu)} \to a_k^{(\nu)}$, so $\Gamma_k \to 1 - M_k(e^{-\nu}) - \big(a_k - a_k^{(\nu)}\big)a_{-k}$ and $Z_k \to 0$: every surfaced regulation is resolved and nothing is recorded. When resolution always fails ($\gamma \to \infty$), b_k, $b_k^{(\nu)}$, d_k and $d_k^{(\nu)}$ all tend to λ^ℓ, so $\Gamma_k \to 0$ and $Z_k \to 1 - M_k(e^{-\nu}) - \big(a_k - a_k^{(\nu)}\big)a_{-k} = \mathbb{E}\big[\mathbb{1}(k \in \mathcal{R})\,\varsigma_k\big]$: every surfaced regulation is recorded.

A.3 *The Backpay Value*

The backpay value of section 3.1.3 sums the restored utility upon complaint and inspection net of the expected complaint cost over shorted worker-regulation pairs:

$$B(\lambda, \ell) \;=\; \sum_{k \in \mathcal{K}} (1-\lambda) \Big[\underbrace{\Delta_k\big(1 - \mathrm{esc}_k(\lambda, \ell)\big)}_{\text{expected restoration}} \;-\; \underbrace{H_k}_{\substack{\text{expected} \\ \text{complaint cost}}} \Big] \tag{71}$$

A shortfall on k is never investigated with the escape probability $\mathrm{esc}_k(\lambda, \ell)$ of eq. (23), so the expected restoration is $\Delta_k(1 - \mathrm{esc}_k)$. The worker pays the complaint cost whenever they complain, with expectation $H_k := \mathbb{E}\big[\omega_{ik}\,\mathbb{1}(\omega_{ik} < \Delta_k)\big]$. Appendix A.4 derives both, and eqs. (96) and (99) give the compliance and size derivatives of B.

A.4 *Derivation of the Escape Probability and the Expected Complaint Cost*

This establishes eq. (23) and eq. (72). We condition throughout on a worker-regulation pair (i, k) being shorted, $S_{ik} = 1$.

(i) Escape. The shorted pair is made whole if and only if regulation k surfaces. Restoration does not wait on the order of resolution: a surfaced regulation is restored whether it is resolved in-window or after a sanction. Given $S_{ik} = 1$ we have $n_k \geq 1$, so k surfaces if and only if a complaint arrives on k or on some other regulation whose investigation reaches k:

$$\{k \notin \mathcal{R}\} \;=\; \{k \notin \mathcal{T}\} \,\cap \bigcap_{j \in \mathcal{D}_k} \{j \notin \mathcal{T}\}.$$

The complaint draws are independent across regulations, so the probability factors. For k itself we must condition on the pair being shorted. The shorted worker is silent with the silence probability m_k of eq. (19), and each of the $\ell - 1$ coworkers is silent on k with

probability ψ_k (eq. (19)), either not shorted on k or shorted and silent, which gives $m_k\psi_k^{\ell-1}$. Every $j \in \mathcal{D}_k$ draws no complaint with probability a_j (eq. (20)), since the conditioning event says nothing about the draws on j. Multiplying gives $\mathrm{esc}_k = m_k\,\psi_k^{\ell-1} a_{-k}$, which is eq. (23). At $\ell = 1$ the worker is alone, so $a_j = \psi_j$ and $\mathrm{esc}_k = m_k \prod_{j\in\mathcal{D}_k} \psi_j$: her own silence on k and on the regulations that reach it decides restoration.

(ii) Expected complaint cost. A shorted pair pays its complaint cost ω_{ik} only on the complaint event $\{\omega_{ik} < \Delta_k\}$, where Δ_k is the backpay on k. With $\omega_{ik} \sim \mathrm{Exp}\big(\rho(\ell)\big)$ and $m_k = e^{-\rho\Delta_k}$ (eqs. (3) and (19)), writing ρ for $\rho(\ell)$ as elsewhere and integrating by parts,

$$\begin{aligned} H_k \;&=\; \int_0^{\Delta_k} u\,\rho e^{-\rho u}\,du \;=\; \Big[-u\,e^{-\rho u}\Big]_0^{\Delta_k} + \int_0^{\Delta_k} e^{-\rho u}\,du \\ &=\; \tfrac{1}{\rho}\big(1-e^{-\rho\Delta_k}\big) - \Delta_k e^{-\rho\Delta_k} \;=\; \tfrac{1}{\rho}\big(1-m_k\big) - \Delta_k m_k. \end{aligned} \tag{72}$$

A.5 *Derivation of the Expected Backpay Cost*

The backpay transfer of section 3.2.4 splits into a compensation block and an amenity block,

$$\mathcal{P}^{\mathrm{bp}}(w,\lambda,\ell) \;=\; \ell\Big[\; \underbrace{w\big(\Phi(\lambda,\ell)-\Lambda(\lambda)\big)}_{\text{regulated compensation}} \;+\; \underbrace{c^a(1-\lambda)\sum_{k\in\mathcal{K}^a}\big(1-\mathrm{esc}_k\big)}_{\text{regulated amenities}} \;\Big], \tag{73}$$

where

$$\Phi(\lambda,\ell) \;:=\; \mathbb{E}\Big[\prod_{k\in\mathcal{K}^w}\big(1+\delta_k^w\,\mathbb{1}(\text{worker receives } k\text{, on time or as backpay})\big)\Big] \tag{74}$$

is the expected wage factor a worker takes home once backpay is paid: the worker receives premium k unless they are shorted on k and no inspection investigates it. Each block is what the firm owes less what escapes restoration, with the escape probability esc_k of eq. (23).

This establishes eq. (73) and gives the closed form of the restored wage factor $\Phi(\lambda,\ell)$ in eq. (74). We set $\delta_k^w := 0$ for $k \in \mathcal{K}^a$ throughout, so that a product over $\mathcal{K}$ of compensation factors is a product over $\mathcal{K}^w$.

Aggregation. Workers are symmetric, so the backpay transfer is ℓ times the expected backpay owed to a single worker, whom we call worker 1.

Amenity block. Each restored unit of amenity $k \in \mathcal{K}^a$ costs the firm c^a. Worker 1 is

restored on k if and only if she is shorted on k, which has probability $1-\lambda$, and k surfaces, which has probability $1-\mathrm{esc}_k$ given the shortfall by eq. (23). Summing over the amenities gives the second term of eq. (73). This block is linear in the escape probabilities, so their marginal values are enough.

Compensation block. Before restoration, worker 1 takes home the realized wage $\tilde{w}_1$ of eq. (81), whose expectation is $w\,\Lambda(\lambda)$ with Λ the realized-wage factor of eq. (10). After restoration she receives premium k unless she is shorted on k and k never surfaces, so her factor on k is the random variable inside eq. (74),

$$\Phi_k \;:=\; 1+\delta_k^w\big(1-S_{1k}\,\mathbb{1}(k\notin\mathcal{R})\big), \tag{75}$$

and $\Phi=\mathbb{E}\big[\prod_{k\in\mathcal{K}^w}\Phi_k\big]$. The backpay she is owed therefore has expectation $w\big(\Phi-\Lambda(\lambda)\big)$, which is the first term of eq. (73). The premia compound, so Φ does not factor across regulations: the same complaints surface several regulations at once.

A switch for each group. Write $O_g:=\{\mathcal{T}\cap\mathcal{R}^g\neq\emptyset\}$ for the event that some regulation in group g draws a complaint, for $g=0,\dots,M$, and O_g^c for its complement; for $g\geq 1$ this is the opened event of Part 1 of the proof of proposition 1. A shorted regulation $k\in\mathcal{R}^m$ surfaces if and only if O_m occurs, and a shorted regulation $k\in\mathcal{R}^0$ surfaces if and only if $\mathcal{T}\neq\emptyset$, where $\mathbb{1}(\mathcal{T}=\emptyset)=\prod_{g=0}^{M}\mathbb{1}(O_g^c)$. If a shortfall on k would surface, Φ_k is the *owed factor* $\Phi_k^1:=1+\delta_k^w$. If it would not, Φ_k is the *delivered factor* $\Phi_k^0:=1+\delta_k^w R_{1k}$, because a worker who received k keeps it and a worker who was shorted gets nothing back. For every group $m\geq 1$ and for the always-checked set,

$$\prod_{k\in\mathcal{R}^m}\Phi_k \;=\; \mathbb{1}(O_m)\,\Phi^1_{\mathcal{R}^m}+\mathbb{1}(O_m^c)\,\Phi^0_{\mathcal{R}^m},\qquad \prod_{k\in\mathcal{R}^0}\Phi_k \;=\; \Phi^1_{\mathcal{R}^0}+\mathbb{1}(\mathcal{T}=\emptyset)\,\big(\Phi^0_{\mathcal{R}^0}-\Phi^1_{\mathcal{R}^0}\big).$$

Multiplying across groups, and using $\mathbb{1}(O_m^c)\big[\mathbb{1}(O_m)\Phi^1_{\mathcal{R}^m}+\mathbb{1}(O_m^c)\Phi^0_{\mathcal{R}^m}\big]=\mathbb{1}(O_m^c)\Phi^0_{\mathcal{R}^m}$,

$$\prod_{k\in\mathcal{K}}\Phi_k \;=\; \Phi^1_{\mathcal{R}^0}\prod_{m=1}^{M}\Big[\mathbb{1}(O_m)\,\Phi^1_{\mathcal{R}^m}+\mathbb{1}(O_m^c)\,\Phi^0_{\mathcal{R}^m}\Big] \;+\; \mathbb{1}(O_0^c)\big(\Phi^0_{\mathcal{R}^0}-\Phi^1_{\mathcal{R}^0}\big)\prod_{m=1}^{M}\mathbb{1}(O_m^c)\,\Phi^0_{\mathcal{R}^m}.$$

Every factor now depends on the draws of a single group, and the draws are independent across groups, so the expectation passes through group by group.

The per-regulation moments. The owed factor is deterministic, so $\mathbb{E}\big[\mathbb{1}(O_g^c)\,\Phi^1_{\mathcal{R}^g}\big]=a_g\,\Phi^1_{\mathcal{R}^g}$.

For the delivered factor on the no-complaint event,

$$\bar{\Phi}^0_k \;:=\; \mathbb{E}\big[\mathbb{1}(k \notin \mathcal{T})\,\Phi^0_k\big] \;=\; a_k + \delta^w_k\,\lambda\,\psi_k^{\ell-1}. \tag{76}$$

The first term is $\mathbb{P}(k \notin \mathcal{T}) = a_k$. In the second, worker 1 receives k with probability λ, and a worker who is not shorted cannot complain, so only the $\ell - 1$ coworkers must be silent on k, each with probability ψ_k. Independence across regulations then gives $\mathbb{E}\big[\mathbb{1}(O^c_g)\,\Phi^0_{\mathcal{R}^g}\big] = \bar{\Phi}^0_{\mathcal{R}^g}$. On an amenity, $\Phi^1_k = 1$ and $\bar{\Phi}^0_k = a_k$: amenities carry no premium, but a complaint about them still opens their group.

Assemble. Taking expectations group by group,

$$\Phi(\lambda,\ell) \;=\; \Phi^1_{\mathcal{R}^0} \prod_{m=1}^{M} \Big[\underbrace{(1-a_m)\,\Phi^1_{\mathcal{R}^m}}_{\text{group opened}} + \underbrace{\bar{\Phi}^0_{\mathcal{R}^m}}_{\text{group silent}} \Big] \;+\; \underbrace{\big(\bar{\Phi}^0_{\mathcal{R}^0} - a_0\,\Phi^1_{\mathcal{R}^0}\big) \prod_{m=1}^{M} \bar{\Phi}^0_{\mathcal{R}^m}}_{\text{no complaint at all}}. \tag{77}$$

Each group is opened with probability $1 - a_m$, in which case the worker is restored in full on its regulations, or stays silent, in which case she keeps only what the firm delivered. The always-checked set is restored whenever the firm is inspected, and the last term replaces its owed factor by the delivered one when nobody complains. As ℓ grows with $\lambda < 1$ and $\rho_1 > -1$, every a_k and $\psi_k^{\ell-1}$ vanishes and $\Phi \to \Lambda(1)$: the firm pays all it owes. When nobody ever complains ($m_k = 1$), $\bar{\Phi}^0_k = 1 + \delta^w_k\lambda$ and $\Phi = \Lambda(\lambda)$: the firm pays no backpay.

Compliance Lowers the Backpay Transfer

Lemma 2 (Compliance and the backpay transfer)**.** *If the compensation rates satisfy $\bar{\delta} := \sum_{k\in\mathcal{K}^w} \delta^w_k \le 1$, the backpay transfer per worker $\mathcal{P}^{\mathrm{bp}}/\ell$ of eq. (73) is weakly decreasing in compliance λ at every wage w and size ℓ.*

Proof. Amenity block. The factor $1 - \lambda$ falls with λ, and $1 - \mathrm{esc}_k$ falls with λ because $\partial_\lambda \mathrm{esc}_k \ge 0$ by eq. (97), so $c^a(1-\lambda)\sum_{k\in\mathcal{K}^a}(1-\mathrm{esc}_k)$ is decreasing.

Compensation block. Let $r_k := 1 - S_{1k}\,\mathbb{1}(k \notin \mathcal{R})$ indicate that worker 1 receives k, on time or as backpay, and write $\Pi(\mathcal{K}') := \prod_{j\in\mathcal{K}'}(1+\delta^w_j)$ and $\delta(\mathcal{K}') := \sum_{j\in\mathcal{K}'}\delta^w_j$ for a set $\mathcal{K}' \subseteq \mathcal{K}^w$. By eqs. (10) and (74), $\Phi - \Lambda(\lambda) = \mathbb{E}[G]$, where

$$G \;:=\; \prod_{k\in\mathcal{K}^w} \big(1 + \delta^w_k\,r_k\big) \;-\; \prod_{k\in\mathcal{K}^w} \big(1 + \delta^w_k\,R_{1k}\big)$$

is the restored premium of worker 1. Hold the complaint costs ω_{ik} fixed. The shortfalls S_{ik} are independent with $\mathbb{P}(S_{ik}=0)=\lambda$, so $\mathbb{E}[G]$ is affine in each pair's probability and

$$\frac{d\,\mathbb{E}[G]}{d\lambda} \;=\; \sum_{i=1}^{\ell}\sum_{k\in\mathcal{K}} \mathbb{E}\big[G_0^{ik}-G_1^{ik}\big],$$

where G_s^{ik} is G with S_{ik} set to s and every other draw unchanged. Honoring a pair removes a possible complaint and adds none, so it weakly shrinks the complained set $\mathcal{T}$ and hence the surfaced set $\mathcal{R}$.

(i) Other pairs. If $i\neq 1$ or $k\in\mathcal{K}^a$, honoring (i,k) leaves every R_{1j} unchanged and weakly lowers every r_j, so $G_0^{ik}-G_1^{ik}\le 0$.

(ii) Worker 1's compensation pairs. Fix the other draws and suppose $S_{1k}=1$. Split $\mathcal{K}^w$ into the regulations worker 1 is honored on, $\mathcal{K}_h$, shorted on and restored on, $\mathcal{K}_r$, and shorted on and not restored on, $\mathcal{K}_u$, so that $k\in\mathcal{K}_r\cup\mathcal{K}_u$. Honoring $(1,k)$ sets $R_{1k}=r_k=1$ and weakly lowers every other r_j, so

$$G_0^{1k}-G_1^{1k} \;\le\; \delta_k^w\Big[(1-r_k)\,\Pi(\mathcal{K}_h)\,\Pi(\mathcal{K}_r)-\Pi(\mathcal{K}_h)\Big] \;=\; \begin{cases} \delta_k^w\,\Pi(\mathcal{K}_h)\big(\Pi(\mathcal{K}_r)-1\big), & k\in\mathcal{K}_u,\\ -\,\delta_k^w\,\Pi(\mathcal{K}_h), & k\in\mathcal{K}_r. \end{cases}$$

The first case is positive: honoring k raises the wage factor on which the restored premia compound. The difference $G_0^{1k}-G_1^{1k}$ does not depend on S_{1k}, which is independent of the other draws, so $\mathbb{E}[G_0^{1k}-G_1^{1k}]=\mathbb{E}\big[S_{1k}(G_0^{1k}-G_1^{1k})\big]/(1-\lambda)$. Summing over $k\in\mathcal{K}^w$,

$$\sum_{k\in\mathcal{K}^w}\mathbb{E}\big[G_0^{1k}-G_1^{1k}\big] \;\le\; \frac{1}{1-\lambda}\,\mathbb{E}\Big[\Pi(\mathcal{K}_h)\Big(\delta(\mathcal{K}_u)\big(\Pi(\mathcal{K}_r)-1\big)-\delta(\mathcal{K}_r)\Big)\Big].$$

Telescoping the product gives $\Pi(\mathcal{K}_r)-1\le\delta(\mathcal{K}_r)\,\Pi(\mathcal{K}_r)$, and $\Pi(\mathcal{K}_r)\le e^{\delta(\mathcal{K}_r)}$, so with $x:=\delta(\mathcal{K}_r)$ and $\delta(\mathcal{K}_u)\le\bar\delta-x$ the bracket is at most $x\big[(\bar\delta-x)e^{x}-1\big]$. The map $x\mapsto(\bar\delta-x)e^{x}$ has derivative $(\bar\delta-x-1)e^{x}\le 0$ for $x\ge 0$ when $\bar\delta\le 1$, so $(\bar\delta-x)e^{x}\le\bar\delta\le 1$ and the bracket is nonpositive.

Adding (i) and (ii), $d\,\mathbb{E}[G]/d\lambda\le 0$, so the compensation block $w\big(\Phi-\Lambda(\lambda)\big)$ is decreasing in λ at every $w\ge 0$. □

The condition is needed only because the premia compound. Without it the sign can reverse: with one worker, no related groups and $\lambda=0$, the compensation block has slope

$\sum_k \delta_k^w \left[(1-s_k)\prod_{j\neq k}(1+\delta_j^w s_j)-1\right]$, where $s_k = 1-m_k$, and this is positive when the rates are large and the s_k small. The statutory rates of table 2 sum to $\bar{\delta} = 0.47$.

A.6 *Microfoundation of the Opportunity Cost Curve*

This derives the opportunity cost curve $\mathcal{X}$ of eq. (11) from one primitive, the distribution F of the firm's shadow opportunity cost q, and establishes the two results the cost and penalty blocks use: the properties of $\mathcal{X}$ in lemma 3, and the scaling $(1+\mathcal{X}_1)$ on restored obligations in eq. (13). The opportunity cost is a firm-side resource cost. Workers and inspectors never see it, so every worker-side object of the main text is unchanged.

Environment and the curve

Within a period a worker's regulatory obligations come due at distinct moments of the firm's production cycle. Obligations arrive as a continuum of units. At each due date the firm faces a shadow opportunity cost $q \sim F$ on $[1,\infty)$ with $F^{-1}(0) = 1$, drawn independently across obligation units.[22] Paying one dollar of obligations in state q costs the firm q dollars: the transfer itself, plus the opportunity cost $q-1$. The floor $q=1$ is the frictionless state, and larger q are tighter states. The firm commits to a state-contingent payment rule: it honors an obligation due in state q if and only if q lies in a payment set $\mathcal{Q} \subseteq [1,\infty)$.

Step 1: the optimal rule is a threshold. The benefit of honoring an obligation, whether it attracts workers or avoids a penalty, does not depend on the state in which the obligation is honored. The cost per dollar, q, rises with the state. Fix any payment frequency $\mathbb{P}(q \in \mathcal{Q})$. Among the payment sets with that frequency, the one with the lowest expected cost collects the cheapest states, so it is a lower interval: pay if and only if $q \leq q^*$. The firm's choice is therefore the cutoff q^*.

Step 2: compliance is a probability. Whether an obligation was honored depends only on the state at its due date. Seen from the worker's or the inspector's side, an obligation unit is therefore honored with probability

$$\lambda \;=\; \mathbb{P}(q \leq q^*) \;=\; F(q^*). \tag{78}$$

[22] Independence across obligation units is what preserves the independence of shortfalls across worker-regulation pairs that the complaint and count blocks rely on. A single firm-level draw per period would make all shortfalls within a firm-period move together and break the binomial count law of eq. (15).

Compliance is the marginal law of a deterministic threshold rule against a random state, and every obligation unit, compensation or amenity, is read against the same F under the same cutoff. This is what makes λ a probability rather than an intensity.

Step 3: the curve as an integral. Under cutoff q^* the expected opportunity cost per dollar of obligation is $\int_1^{q^*}(q-1)\,dF(q)$: the cost $q-1$ averaged over the states in which the firm pays. We change variables to the quantile $u = F(q)$, so that $q = F^{-1}(u)$ and $dF(q) = du$, and the range $q \in [1, q^*]$ maps to $u \in [0, \lambda]$ by eq. (78):

$$\int_1^{q^*} (q-1)\,dF(q) \;=\; \int_0^{\lambda} \big(F^{-1}(u) - 1\big)\,du \;=\; \mathcal{X}(\lambda),$$

which is the curve of eq. (11). The cost of promised obligations $\mathsf{b}(w)$ multiplies it because each obligation unit is a dollar of that bill, so the compliance cost is $\mathsf{b}(w)\mathcal{X}(\lambda)$.

Lemma 3 (Properties of the opportunity cost curve)**.** *The curve $\mathcal{X}$ of eq. (11) satisfies $\mathcal{X}(0) = 0$; $\mathcal{X}'(\lambda) = F^{-1}(\lambda) - 1 \geq 0$ is nondecreasing, so $\mathcal{X}$ is convex; $\mathcal{X}'(0) = 0$; and $\mathcal{X}(1) = \mathbb{E}[q] - 1$, the mean opportunity cost. Choosing $\mathcal{X}$ is equivalent to choosing F, one-to-one, through $F^{-1}(u) = 1 + \mathcal{X}'(u)$.*

Proof. Level and slope. An integral over an empty range is zero, so $\mathcal{X}(0) = 0$. By the fundamental theorem of calculus, $\mathcal{X}'(\lambda) = F^{-1}(\lambda) - 1$, which is nonnegative because the support of q starts at 1.

Convexity. A quantile function is nondecreasing, so $\mathcal{X}'$ is nondecreasing and $\mathcal{X}$ is convex. At the origin, $F^{-1}(0) = 1$ gives $\mathcal{X}'(0) = 0$: the first obligations are honored in the frictionless state.

Level at full compliance. For $U \sim \mathrm{Unif}[0,1]$ the quantile transform $F^{-1}(U)$ has the law of q, so

$$\mathcal{X}(1) \;=\; \int_0^1 \big(F^{-1}(u) - 1\big)\,du \;=\; \mathbb{E}\big[F^{-1}(U)\big] - 1 \;=\; \mathbb{E}[q] - 1.$$

One-to-one. Given $\mathcal{X}$, the map $u \mapsto 1 + \mathcal{X}'(u)$ is nondecreasing on $[0,1]$ and equals 1 at $u = 0$, so it is the quantile function of a distribution on $[1, \infty)$, and that distribution is F. Given F, eq. (11) returns $\mathcal{X}$. □

In the main text the opportunity cost $q - 1$ is exponential with mean ϕ,

$$F(q) \;=\; 1 - e^{-(q-1)/\phi} \quad (q \geq 1), \qquad F^{-1}(u) \;=\; 1 - \phi \ln(1-u), \tag{79}$$

which gives the closed forms of eq. (12). The marginal cost is $\mathcal{X}'(\lambda) = F^{-1}(\lambda) - 1 = -\phi \ln(1-\lambda)$, and integrating from zero,

$$\mathcal{X}(\lambda) \;=\; -\phi \int_0^{\lambda} \ln(1-u)\, du \;=\; \phi\big[(1-\lambda)\ln(1-\lambda) + \lambda\big],$$

with $\mathcal{X}_1 = \mathcal{X}(1) = \phi$, the mean of an exponential opportunity cost with mean ϕ. The marginal cost $-\phi \ln(1-\lambda)$ rises without bound as $\lambda \to 1$ while the level stays at ϕ: the last obligations are honored in states that are arbitrarily tight but arbitrarily rare.

The convexity is the selection logic itself. The firm fills its compliance quota from the cheapest states upward, so the marginal unit of λ is paid in the λ-quantile state, and each additional unit is bought in a weakly worse state than the last.

Derivation of the Scaling on Penalties

This establishes eq. (13): the firm's outlay on its penalties is $(1 + \mathcal{X}_1)\big(\mathcal{P}^{\mathrm{bp}} + \mathcal{P}^{\mathrm{s}}\big)$, with $\mathcal{P}^{\mathrm{bp}}$ the backpay transfer of eq. (73) and $\mathcal{P}^{\mathrm{s}}$ the sanction of eq. (64).

Step 1: the cost on a dollar paid at inspection. Condition on a shorted pair being restored, or on a sanction being imposed. The firm pays the owed transfer or the fine at the inspection date, in whatever state prevails then: a fresh draw $q \sim F$, independent of the events that surfaced the regulation. Payment is ordered rather than timed, so the firm has no selection margin over this state. Taking expectations over q, the expected opportunity cost per dollar paid is $\mathbb{E}[q] - 1 = \mathcal{X}(1) = \mathcal{X}_1$ by lemma 3.[23]

Step 2: linearity in the restored bill. The rate $\mathcal{X}_1$ is a constant, so by linearity of expectations the expected opportunity cost is $\mathcal{X}_1$ times the expected amount paid. That amount is the backpay transfer plus the sanction, $\mathcal{P}^{\mathrm{bp}} + \mathcal{P}^{\mathrm{s}}$, whatever their internal structure.

Step 3: assemble. Adding the payments to their opportunity cost, the outlay is $\mathcal{P}^{\mathrm{bp}} + \mathcal{P}^{\mathrm{s}} + \mathcal{X}_1\big(\mathcal{P}^{\mathrm{bp}} + \mathcal{P}^{\mathrm{s}}\big) = (1 + \mathcal{X}_1)\big(\mathcal{P}^{\mathrm{bp}} + \mathcal{P}^{\mathrm{s}}\big)$, which is eq. (13). The scaling is exact because the rate is a constant that enters linearly: no moment of a nonlinear function of the random obligations is needed.

[23] Were the firm able instead to route restored obligations through its regular payment queue, the cost per dollar on backpay would be the average over its *selected* states, $\mathcal{X}(\lambda)/\lambda$, which is strictly below $\mathcal{X}_1$ for $\lambda < 1$, and the payment channel would largely wash out of the compliance margin. We assume the restoration order forecloses deferral; the deferral variant replaces $\mathcal{X}_1$ by $\mathcal{X}(\lambda)/\lambda$ in eq. (13) and remains closed-form.

A.7 *Worker Utility*

A worker receives regulation k unless she is shorted on it, $R_{ik} = 1 - S_{ik}$. Her utility is log in realized compensation and linear in amenities,

$$W_{ij} = \beta \ln(\tilde{w}_i) + \sum_{k \in \mathcal{K}^a} \beta_k^a \, R_{ik}, \tag{80}$$

where her realized wage compounds the premia she receives,

$$\tilde{w}_i \;=\; w \prod_{k \in \mathcal{K}^w} \left(1 + \delta_k^w\right)^{R_{ik}}. \tag{81}$$

Taking logs, $\beta \ln \tilde{w}_i = \beta \ln w + \sum_{k \in \mathcal{K}^w} \beta \ln(1+\delta_k^w) \, R_{ik}$, so an honored compensation regulation adds a fixed number of utils, just as an honored amenity does:

$$\Delta_k = \begin{cases} \beta \ln(1 + \delta_k^w) & \text{if } k \in \mathcal{K}^w \\ \beta_k^a & \text{if } k \in \mathcal{K}^a \end{cases} \tag{82}$$

Each regulation is honored with probability $\mathbb{E}[R_{ik}] = \lambda$, so expected utility is $\beta \ln w + \lambda \sum_{k \in \mathcal{K}} \Delta_k$, which is eq. (5).

A.8 *Labor Supply Derivation*

This derives the labor supply curve of eq. (2) and the semi-elasticity η_j from the worker's problem in eq. (1). Worker i chooses among the J formal firms and the informal sector. Her utility at formal firm j is $V_j + \varepsilon_{ij}$, and in the informal sector it is $V_0 + \varepsilon_{i0}$, with the informal value V_0 of eq. (16), which includes the informal taste ξ_0. The taste shocks ε_{ij} are i.i.d. type-I extreme value with scale one across all $J + 1$ options.

Labor supply. With type-I extreme value shocks, the probability that an option is the utility maximizer is its exponentiated value over the sum of exponentiated values, the logit choice probability. Worker i therefore chooses formal firm j with probability

$$\mathbb{P}(i \text{ chooses } j) \;=\; \frac{\exp(V_j)}{\exp(V_0) + \sum_{j' \in \mathcal{J}} \exp(V_{j'})} \;=\; \exp\left(V_j - IV\right),$$

where the second equality writes the denominator as $\exp(IV)$ with the inclusive value IV of eq. (2). The choice probabilities do not depend on i, so the labor supply to firm j from a workforce of L workers is $N_j = L\, e^{V_j - IV}$, and N_j/L is firm j's share of the workforce. Taking logs,

$$\ln N_j \;=\; V_j - IV + \ln L,$$

which is eq. (2). The informal sector's employment follows the same rule with its own value, $\ln N_0 = V_0 - IV + \ln L$. Subtracting it from eq. (2) cancels the two market-wide terms IV and $\ln L$ and gives the estimating equation

$$\ln N_j - \ln N_0 \;=\; V_j - V_0. \tag{83}$$

Formal firms carry no firm-specific taste, so ξ_0 enters only through V_0, the common intercept.

Labor supply semi-elasticity. The value V_j enters $\ln N_j$ directly and through the inclusive value. Differentiating the inclusive value in eq. (2),

$$\frac{\partial IV}{\partial V_j} \;=\; \frac{\exp(V_j)}{\exp(V_0) + \sum_{j' \in \mathcal{J}} \exp(V_{j'})} \;=\; \frac{N_j}{L},$$

so the semi-elasticity of labor supply with respect to the value the firm offers is

$$\eta_j \;:=\; \frac{\partial \ln N_j}{\partial V_j} \;=\; 1 - \frac{\partial IV}{\partial V_j} \;=\; 1 - \frac{N_j}{L}. \tag{84}$$

A firm that raises its value pulls workers from every other option in proportion to their shares, and its own share N_j/L is the part of the market it cannot pull from.

A.9 *The Wage Elasticity of Labor Supply*

This derives the total elasticity of eq. (34) and the markdown it implies. Consider a small rise in a firm's productivity, which moves its wage, compliance and size together. The value of a job in eq. (4) depends on all three, so it changes by

$$dV_j \;=\; \beta\, d\ln w_j \;+\; \big(D + \vartheta B_\lambda\big)\, d\lambda_j \;+\; \vartheta\, \ell_j B_\ell\, d\ln \ell_j,$$

where $D = \sum_{k \in \mathcal{K}} \Delta_k$. Labor supply responds to the value with the semi-elasticity of eq. (84), so $d \ln N_j = \eta_j \, dV_j$. In equilibrium the firm's size is its labor supply, $\ell_j = N_j$, so the last term feeds back into the left-hand side. Collecting $d \ln N_j$ and dividing by $d \ln w_j$,

$$\frac{d \ln N_j}{d \ln w_j} \;=\; \frac{1}{1 - \eta_j \vartheta \, \ell_j B_\ell} \left[\eta_j \beta + \eta_j \big(D + \vartheta B_\lambda \big) \frac{d\lambda_j}{d \ln w_j} \right],$$

which is eq. (34). Every term is in closed form except $d\lambda_j / d \ln w_j$, which we measure as the difference in compliance over the difference in log wages between adjacent firms on the productivity grid.

Markdowns. The markdown at a posting is $w_j c_w / (\beta \eta_j)$, the per-worker cost of a wage increase over the partial elasticity (eq. (45)). If compliance is assumed fixed, the total elasticity is read as the wage elasticity, and the implied markdown is

$$\frac{w_j \, c_w}{d \ln N_j / d \ln w_j} \;\leq\; \frac{w_j \, c_w}{\beta \eta_j},$$

which is smaller whenever compliance rises with the wage. Figure 12 plots both by firm size. The gap closes at the largest firms: above a thousand workers the markdown is 57% of the wage, against 48% if compliance is assumed fixed.

A.10 *The Firm's Optimality Conditions across Enforcement Regimes*

This gives the firm's optimality conditions in each enforcement regime of section 3.4. We first derive the limits in table 10, then the first-order conditions of the Lagrangian in eq. (85), and then, regime by regime, the penalty and backpay derivatives that the main text leaves in symbolic form.

The limits. The regimes differ only in the complaint cost scale ρ_0, which enters every enforcement object through the silence probability $m_k = e^{-\rho \Delta_k}$ (eq. (19)), with $\rho = \rho_0 \ell^{\rho_1}$. Both limits require $\Delta_k > 0$ for every regulation, so that m_k moves with ρ. Table 10 collects the limits.

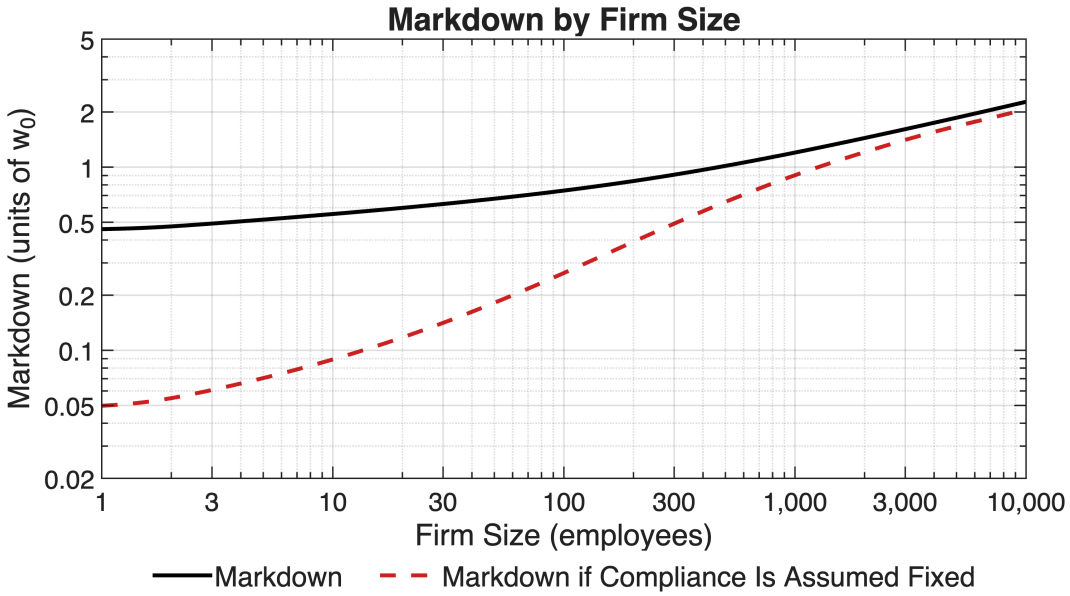


Figure 12: The markdown by firm size in the status quo, in units of the informal wage w_0, against the markdown implied when compliance is assumed fixed.

Table 10: The three enforcement regimes as limits of the complaint cost.

	No enforcement $\rho_0 \to 0$	Perfect enforcement $\rho_0 \to \infty$
Silence given a shortfall, m_k (19)	1	0
Escape probability, esc_k (23)	1	0
Inspection probability, $\mathscr{I}$ (21)	0	$1-\lambda^{\ell\lvert\mathcal{K}\rvert}$
Expected complaint cost, H_k (72)	0	0
Backpay value, B (71)	0	$(1-\lambda)\sum_{k\in\mathcal{K}}\Delta_k$
Restored wage factor, Φ (74)	$\Lambda(\lambda)$	$\Lambda(1)$
Backpay transfer, $\mathcal{P}^{\text{bp}}$ (73)	0	$\ell\big[w\big(\Lambda(1)-\Lambda(\lambda)\big)+c^a K^a(1-\lambda)\big]$
Expected sanction, $\mathcal{P}^{\text{s}}$ (64)	0	(64), every shorted regulation investigated

No enforcement, $\rho_0 \to 0$. The complaint rate ρ goes to zero, so $m_k \to 1$: a shorted worker never complains. The unconditional silence $\psi_k = \lambda+(1-\lambda)m_k$ (eq. (19)) and the no-complaint factor $a_k = \psi_k^\ell$ (eq. (20)) then go to one, so the escape probability $\text{esc}_k = m_k\psi_k^{\ell-1}a_{-k}$ (eq. (23)) goes to one and the inspection probability $\mathscr{I} = 1-\prod_k a_k$ (eq. (21)) goes to zero. The expected complaint cost of eq. (72) vanishes: $(1-m_k)/\rho = (1-e^{-\rho\Delta_k})/\rho \to \Delta_k$ and $\Delta_k m_k \to \Delta_k$, so $H_k \to 0$. With $\text{esc}_k = 1$ and $H_k = 0$, the backpay value of eq. (71) is $B = 0$.

In the restored wage factor of eq. (77), $a_k = 1$ closes every group, and the delivered factor on the no-complaint event in eq. (76) becomes $\bar{\Phi}^0_k = 1 + \delta^w_k \lambda$, so $\Phi = \prod_{k \in \mathcal{K}^w}(1 + \delta^w_k \lambda) = \Lambda(\lambda)$. Nothing surfaces, so the backpay transfer of eq. (73) and the expected sanction of eq. (64) are zero, and $\mathcal{P} = 0$.

Perfect enforcement, $\rho_0 \to \infty$. The complaint rate goes to infinity, so $m_k \to 0$: every shorted worker complains. Then $\text{esc}_k = 0$, since its leading factor m_k is zero, and every shortfall surfaces on its own regulation. The unconditional silence is $\psi_k = \lambda$, so $a_k = \lambda^\ell$, and a firm is uninspected only when no worker-regulation pair is shorted: $\mathscr{I} = 1 - \lambda^{\ell|\mathcal{K}|}$. The expected complaint cost vanishes again, now because $1/\rho \to 0$ and $\Delta_k m_k \to 0$, so $B = (1 - \lambda) \sum_{k \in \mathcal{K}} \Delta_k = (1 - \lambda) D$. In eq. (76), $\bar{\Phi}^0_k = \lambda^\ell + \delta^w_k \lambda \cdot \lambda^{\ell - 1} = a_k(1 + \delta^w_k)$, so each group bracket of eq. (77) collapses to the owed factor of its regulations and the last term vanishes: $\Phi = \prod_{k \in \mathcal{K}^w}(1 + \delta^w_k) = \Lambda(1)$, and the firm restores every shortfall. The backpay transfer is then $\ell[w(\Lambda(1) - \Lambda(\lambda)) + c^a K^a (1 - \lambda)]$, and the expected sanction is eq. (64) with the surfaced set $\mathcal{R} = \{k : n_k \geq 1\}$.

The Lagrangian and first-order conditions

The firm problem in eq. (7) has the same form in every regime, and only the penalty $\mathcal{P}$ and the backpay value B change across them. We write labor supply as $N_j = L\, e^{V_j - IV}$ (eq. (2)), attach the multiplier $\mu \geq 0$, the recruiting cost, to the labor supply constraint $\ell \leq N_j$ and $\kappa \geq 0$ to the minimum wage $w \geq w^r$, and form the Lagrangian

$$\mathcal{L}_j \;=\; z_j \ell^\alpha - c(w, \lambda)\, \ell - \mathcal{P}(w, \lambda, \ell) + \mu \big[N_j(w, \lambda, \ell) - \ell \big] + \kappa\, (w - w^r). \tag{85}$$

Labor supply responds to the firm's choices only through the value $V_j = \beta \ln w + \lambda D + \vartheta B(\lambda, \ell)$ of eq. (4). By eq. (84), $\partial N_j / \partial V_j = \eta N_j$, and the chain rule gives

$$\frac{\partial N_j}{\partial w} \;=\; \eta N_j \frac{\beta}{w}, \qquad \frac{\partial N_j}{\partial \lambda} \;=\; \eta N_j \left(D + \vartheta B_\lambda \right), \qquad \frac{\partial N_j}{\partial \ell} \;=\; \eta N_j\, \vartheta B_\ell,$$

where B_λ and B_ℓ are the derivatives of the backpay value. When the constraint binds, $\mu > 0$ and $N_j = \ell$ by complementary slackness. When it is slack, $\mu = 0$ and the derivatives of N_j drop out. In both cases we may replace N_j by ℓ in the terms that carry μ. Differentiating

eq. (85) with respect to w, λ and ℓ and setting each derivative to zero,

$$c_w\,\ell + \mathcal{P}_w \;=\; \mu\eta\ell\,\frac{\beta}{w} + \kappa, \tag{86}$$

$$c_\lambda\,\ell + \mathcal{P}_\lambda \;=\; \mu\eta\ell\,\big(D + \vartheta B_\lambda\big), \tag{87}$$

$$\alpha z_j \ell^{\alpha-1} - c - \mathcal{P}_\ell \;=\; \mu\,\big(1 - \eta\ell\,\vartheta B_\ell\big). \tag{88}$$

Each left-hand side is a marginal cost and each right-hand side is the value of the constraint it relaxes. Raising the wage or compliance recruits $\mu\eta\ell$ times the utility it adds per worker. Hiring one more worker tightens the labor supply constraint by one, net of the backpay value that the larger size delivers to every worker.

The mix of compliance and the wage. When the wage is above the minimum, $\kappa = 0$. Dividing eq. (87) by eq. (86) cancels $\mu\eta\ell$ and gives the marginal rate of technical substitution between compliance and the wage,

$$\underbrace{\frac{c_\lambda + \mathcal{P}_\lambda/\ell}{D + \vartheta B_\lambda}}_{\text{cost of a util through compliance}} \;=\; \underbrace{\frac{c_w + \mathcal{P}_w/\ell}{\beta/w}}_{\text{cost of a util through the wage}} \;=\; \eta\mu, \tag{89}$$

which is eq. (17). Each ratio is a per-worker marginal cost over the utility it buys a worker, so the firm equates the cost of a util across its two instruments, and the common value is $\eta\mu$. At the minimum wage $\kappa > 0$, and eq. (86) gives $(c_w + \mathcal{P}_w/\ell)/(\beta/w) = \eta\mu + \kappa w/(\beta\ell) > \eta\mu$: the wage is the more expensive instrument, and the firm would cut it if it could.

The per-worker cost in eq. (8) is $c(w,\lambda) = \Lambda(\lambda)w + c^a K^a \lambda + \mathsf{b}(w)\mathcal{X}(\lambda)$, with the cost of promised obligations $\mathsf{b}(w) = w(\Lambda(1) - 1) + c^a K^a$ (eq. (11)). It is linear in w with $\partial_w \mathsf{b} = \Lambda(1) - 1$, so its derivatives are the same in every regime,

$$c_w \;=\; \underbrace{\Lambda(\lambda)}_{\text{delivered wage}} + \underbrace{\big(\Lambda(1) - 1\big)\mathcal{X}(\lambda)}_{\text{compliance cost}}, \qquad c_\lambda \;=\; \underbrace{w\,\Lambda'(\lambda) + c^a K^a}_{\text{delivery}} + \underbrace{\mathsf{b}(w)\,\mathcal{X}'(\lambda)}_{\text{compliance cost}}. \tag{90}$$

The delivered wage factor $\Lambda(\lambda) = \prod_{k\in\mathcal{K}^w}(1 + \delta_k^w \lambda)$ (eq. (10)) differentiates through its logarithm, $\Lambda'(\lambda) = \Lambda(\lambda)\sum_{k\in\mathcal{K}^w} \delta_k^w/(1 + \delta_k^w\lambda)$, and $\mathcal{X}'(\lambda) = -\phi\ln(1-\lambda)$ from eq. (12).

The labor choice. Rearranging eq. (88), a firm on its labor supply curve hires until

$$\underbrace{\alpha z_j \ell^{\alpha-1}}_{MRPL} = \underbrace{c(w,\lambda)}_{\text{per-worker cost}} + \underbrace{\mathcal{P}_\ell}_{\text{size penalty}} + \underbrace{\mu}_{\text{recruiting cost}} - \underbrace{\eta\mu\,\vartheta\,\ell\,B_\ell}_{\substack{\text{value delivered}\\ \text{by inspection}}}, \tag{91}$$

which is eq. (18). To hire one more worker the firm must raise the value of the job to all of its workers, which costs the recruiting cost μ. One util raises labor supply by $\eta\ell$ workers, so attracting one more worker takes $1/(\eta\ell)$ utils for each of the ℓ workers, at a cost of $\eta\mu$ each, which is μ. A larger firm is also inspected more often, which raises the backpay value by B_ℓ for each of its ℓ workers and delivers that part of the value for free.

A firm that *rations* hires fewer workers than sort to it, $\ell < N_j$, so $\mu = 0$ by complementary slackness. Then eq. (86) reads $\kappa = c_w\ell + \mathcal{P}_w$. Both terms are positive: $c_w > 0$ by eq. (90), and $\mathcal{P}_w = (1+\mathcal{X}_1)\ell\big(\Phi - \Lambda(\lambda)\big) \geq 0$ because the restored wage factor is at least the delivered one. So $\kappa > 0$ and the minimum wage binds: a firm rations only at the minimum wage. Compliance no longer buys workers, so eq. (87) sets it to minimize cost plus penalty, $c_\lambda + \mathcal{P}_\lambda/\ell = 0$. Both recruiting terms leave eq. (91), and the firm hires until $MRPL = c + \mathcal{P}_\ell$.

No Enforcement

Under no enforcement $\mathcal{P} = 0$ and $B = 0$ (table 10), so $\mathcal{P}_w = \mathcal{P}_\lambda = \mathcal{P}_\ell = 0$ and $B_\lambda = B_\ell = 0$. Substituting eq. (90) into eq. (89), the mix condition becomes

$$\underbrace{\frac{\overbrace{w\,\Lambda'(\lambda) + c^a K^a}^{\text{delivery}} + \overbrace{\mathsf{b}(w)\,\mathcal{X}'(\lambda)}^{\text{compliance cost}}}{D}}_{\text{delivered utility}} = \underbrace{\frac{\overbrace{\Lambda(\lambda)}^{\text{delivered wage}} \quad + \overbrace{\big(\Lambda(1)-1\big)\mathcal{X}(\lambda)}^{\text{compliance cost}}}{\beta/w}}_{\text{utility of the wage}}, \tag{92}$$

and the labor choice in eq. (91) becomes

$$\alpha z_j \ell^{\alpha-1} = c(w,\lambda) + \underbrace{\mu}_{\text{markdown}}. \tag{93}$$

A firm still complies with regulations, but only to attract workers, and the recruiting cost μ, the gap between the marginal product and the per-worker cost, is the monopsony markdown. A firm that rationed would set $\lambda = 0$ and $w = w^r$. With $\mu = 0$, compliance is set by $c_\lambda = 0$, but $c_\lambda \geq w\Lambda'(\lambda) + c^a K^a > 0$ at every λ, so the firm stops at the corner, and the wage is at

the minimum by the argument above. Both instruments would then only cost it money.

Status Quo

Enforcement adds the penalty derivatives and the backpay value. We take the wage and compliance derivatives first and the size derivatives second. Throughout, the penalty is $\mathcal{P} = (1+\mathcal{X}_1)\big(\mathcal{P}^{\mathrm{bp}} + \mathcal{P}^{\mathrm{s}}\big)$ (eq. (9)), with the backpay transfer

$$\mathcal{P}^{\mathrm{bp}} \;=\; \ell\Big[w\big(\Phi(\lambda,\ell) - \Lambda(\lambda)\big) + c^a(1-\lambda)\sum_{k\in\mathcal{K}^a}\big(1-\mathrm{esc}_k\big)\Big]$$

of eq. (73) and the expected sanction $\mathcal{P}^{\mathrm{s}} = \ell^\zeta \sum_{k\in\mathcal{K}} g_k Z_k$ of eq. (64). The restored wage factor Φ, the escape probabilities esc_k and the sanction bases Z_k depend on (λ,ℓ) but not on the wage.

The wage and compliance derivatives. Only the compensation block of the backpay transfer carries w, and it is linear in w, so

$$\frac{\mathcal{P}_w}{\ell} \;=\; \underbrace{(1+\mathcal{X}_1)\big(\Phi - \Lambda(\lambda)\big)}_{\text{wage backpay per unit wage}}. \tag{94}$$

For compliance, we differentiate each block of $\mathcal{P}^{\mathrm{bp}}/\ell$ in turn. The compensation block gives $w\big(\Phi_\lambda - \Lambda'(\lambda)\big)$. In the amenity block, the product rule on $(1-\lambda)(1-\mathrm{esc}_k)$ gives $-(1-\mathrm{esc}_k) - (1-\lambda)\partial_\lambda \mathrm{esc}_k$: higher compliance leaves fewer shortfalls to restore, and each remaining shortfall is less likely to surface. Adding the sanction,

$$\frac{\mathcal{P}_\lambda}{\ell} \;=\; (1+\mathcal{X}_1)\Big[\underbrace{w\big(\Phi_\lambda - \Lambda'(\lambda)\big)}_{\text{wage backpay}} - c^a\sum_{k\in\mathcal{K}^a}\Big(\underbrace{1-\mathrm{esc}_k}_{\text{fewer shortfalls}} + \underbrace{(1-\lambda)\,\partial_\lambda \mathrm{esc}_k}_{\text{fewer surface}}\Big)\Big] \;+\; \underbrace{(1+\mathcal{X}_1)\frac{\mathcal{P}^{\mathrm{s}}_\lambda}{\ell}}_{\text{sanction}}. \tag{95}$$

The backpay value $B = \sum_{k\in\mathcal{K}}(1-\lambda)\big[\Delta_k(1-\mathrm{esc}_k) - H_k\big]$ (eq. (71)) differentiates the same way. The expected complaint cost H_k of eq. (72) depends on λ neither directly nor through the complaint rate, so the product rule gives

$$B_\lambda \;=\; -\sum_{k\in\mathcal{K}}\Big[\underbrace{\Delta_k\big(1-\mathrm{esc}_k\big) - H_k}_{\text{fewer shortfalls to restore}}\Big] \;-\; (1-\lambda)\sum_{k\in\mathcal{K}}\underbrace{\Delta_k\,\partial_\lambda \mathrm{esc}_k}_{\text{fewer restored}}. \tag{96}$$

Both need the compliance derivative of the escape probability. We differentiate $\ln \mathrm{esc}_k = \ln m_k + (\ell-1)\ln\psi_k + \ell\sum_{j\in\mathcal{D}_k}\ln\psi_j$, from eq. (23) and eq. (50). The silence probability m_k does not depend on λ, and $\partial_\lambda\psi_j = 1-m_j$ from eq. (19), so

$$\partial_\lambda \mathrm{esc}_k \;=\; \mathrm{esc}_k\Big[\underbrace{(\ell-1)\frac{1-m_k}{\psi_k}}_{\text{coworkers on } k} + \underbrace{\ell\sum_{j\in\mathcal{D}_k}\frac{1-m_j}{\psi_j}}_{\text{regulations that pull } k \text{ in}}\Big] \;\geq\; 0. \tag{97}$$

Compliance raises the escape probability through the silence of the coworkers on k and of the workers on the regulations whose complaint pulls k in. The derivatives Φ_λ and $\mathcal{P}^{\mathrm{s}}_\lambda$ are worked out below, in eq. (111) and eq. (116).

Enforcement moves the mix through three terms. The penalty saved lowers the cost of compliance, and wage backpay raises the cost of the wage, so both push the firm toward compliance. The backpay value lost lowers the return to compliance and pushes the other way, so the net effect is ambiguous in sign (section 5.1.1).

The size derivatives. Size moves every enforcement object. The penalty derivative is

$$\begin{aligned}
\mathcal{P}_\ell \;&=\; (1+\mathcal{X}_1)\big(\mathcal{P}^{\mathrm{bp}}_\ell + \mathcal{P}^{\mathrm{s}}_\ell\big), \qquad (98)\\
\mathcal{P}^{\mathrm{bp}}_\ell \;&=\; \underbrace{\frac{\mathcal{P}^{\mathrm{bp}}}{\ell}}_{\text{one more worker's backpay}} + \underbrace{\ell\Big(w\,\Phi_\ell - c^a(1-\lambda)\sum_{k\in\mathcal{K}^a}\partial_\ell \mathrm{esc}_k\Big)}_{\text{more surfaces at every worker}},\\
\mathcal{P}^{\mathrm{s}}_\ell \;&=\; \underbrace{\frac{\zeta\,\mathcal{P}^{\mathrm{s}}}{\ell}}_{\text{sanction schedule}} + \underbrace{\ell^\zeta\sum_{k\in\mathcal{K}} g_k\,\partial_\ell Z_k}_{\text{more recorded}}.
\end{aligned}$$

The backpay transfer is ℓ times a per-worker bracket, so the product rule gives the bracket itself, which is $\mathcal{P}^{\mathrm{bp}}/\ell$, plus ℓ times the bracket's own size derivative. The expected sanction is ℓ^ζ times the severity-weighted sum of the sanction bases, so the product rule gives the schedule term $\zeta\mathcal{P}^{\mathrm{s}}/\ell$ plus the change in the bases. The backpay value has no leading ℓ, and both the escape probability and the complaint cost move with size, so

$$B_\ell \;=\; -(1-\lambda)\sum_{k\in\mathcal{K}}\Big[\underbrace{\Delta_k\,\partial_\ell \mathrm{esc}_k}_{\text{more restored}} + \underbrace{\partial_\ell H_k}_{\text{complaint cost}}\Big]. \tag{99}$$

Size enters the escape probability three times: through the exponents $\ell - 1$ and ℓ, and through the complaint rate $\rho(\ell) = \rho_0 \ell^{\rho_1}$ inside every silence probability $m_j = e^{-\rho\Delta_j}$. Write $\rho_\ell := \partial_\ell \rho = \rho_1 \rho/\ell$, so that $\partial_\ell m_j = -\rho_\ell \Delta_j m_j$ and $\partial_\ell \psi_j = (1-\lambda)\,\partial_\ell m_j = -(1-\lambda)\,\rho_\ell\,\Delta_j m_j$. Differentiating $\ln \mathrm{esc}_k$ as above, now with respect to ℓ,

$$\partial_\ell \mathrm{esc}_k \;=\; \mathrm{esc}_k \Big[\underbrace{\ln \psi_k + \sum_{j \in \mathcal{D}_k} \ln \psi_j}_{\text{more workers who might complain}} \;+\; \underbrace{(\ell-1)\frac{\partial_\ell \psi_k}{\psi_k} + \ell \sum_{j \in \mathcal{D}_k} \frac{\partial_\ell \psi_j}{\psi_j} - \rho_\ell \Delta_k}_{\text{cheaper complaints}} \Big]. \tag{100}$$

The first bracket differentiates the exponents and is negative, since every $\psi_j \leq 1$. The second collects the rate channel and is negative when $\rho_1 \geq 0$, since then $\rho_\ell \geq 0$ and every $\partial_\ell \psi_j \leq 0$. So a shortfall at a larger firm is more likely to be restored. The complaint cost $H_k = (1 - m_k)/\rho - \Delta_k m_k$ depends on size only through ρ, with $dm_k/d\rho = -\Delta_k m_k$, so

$$\frac{dH_k}{d\rho} \;=\; -\frac{1-m_k}{\rho^2} + \frac{\Delta_k m_k}{\rho} + \Delta_k^2 m_k \;=\; \frac{1}{\rho^2}\Big[(\rho\Delta_k)^2 m_k - \big(1 - m_k - \rho\Delta_k m_k\big)\Big],$$

and multiplying by ρ_ℓ,

$$\partial_\ell H_k \;=\; \frac{\rho_\ell}{\rho^2}\Big[\underbrace{(\rho\Delta_k)^2 m_k}_{\text{more complain}} - \underbrace{\big(1 - m_k - \rho\Delta_k m_k\big)}_{\text{each pays less}} \Big]. \tag{101}$$

The subtracted term, $1 - e^{-\rho\Delta_k}(1 + \rho\Delta_k)$, is positive. The bracket is positive when $\rho\Delta_k$ is small and negative when it is large, so the sign of eq. (101) is ambiguous. The labor choice is eq. (91) with these derivatives,

$$\alpha z_j \ell^{\alpha-1} \;=\; c(w,\lambda) + \underbrace{(1+\mathcal{X}_1)\mathcal{P}_\ell^{\mathrm{bp}}}_{\text{backpay}} + \underbrace{(1+\mathcal{X}_1)\mathcal{P}_\ell^{\mathrm{s}}}_{\text{sanction}} + \mu - \underbrace{\eta\mu\,\vartheta\,\ell\,B_\ell}_{\substack{\text{value delivered} \\ \text{by inspection}}}. \tag{102}$$

Perfect Enforcement

Under perfect enforcement $\mathrm{esc}_k = 0$ and $H_k = 0$ at every (λ, ℓ), so $\Phi = \Lambda(1)$ and $B = (1-\lambda)D$ (table 10), and their derivatives are $\Phi_\lambda = \Phi_\ell = 0$, $\partial_\lambda \mathrm{esc}_k = \partial_\ell \mathrm{esc}_k = 0$, $B_\lambda = -D$ and $B_\ell = 0$.

Substituting into eq. (94), eq. (95) and eq. (98),

$$\begin{aligned}\frac{\mathcal{P}_w}{\ell} &= (1+\mathcal{X}_1)\big(\Lambda(1)-\Lambda(\lambda)\big),\\ \frac{\mathcal{P}_\lambda}{\ell} &= -(1+\mathcal{X}_1)\big(w\,\Lambda'(\lambda)+c^aK^a\big)+(1+\mathcal{X}_1)\frac{\mathcal{P}^{\mathrm{s}}_\lambda}{\ell},\\ \mathcal{P}_\ell &= (1+\mathcal{X}_1)\Big(\frac{\mathcal{P}^{\mathrm{bp}}}{\ell}+\mathcal{P}^{\mathrm{s}}_\ell\Big).\end{aligned} \tag{103}$$

In the compliance derivative the amenity sum is K^a and the wage backpay is $-w\Lambda'(\lambda)$, so the bracket is minus the delivery cost of compliance in eq. (90). Adding c_λ and $\mathcal{P}_\lambda/\ell$ therefore cancels the delivery cost against the backpay it saves, because a firm that shorts a worker pays her anyway, and leaves

$$c_\lambda+\frac{\mathcal{P}_\lambda}{\ell} = \mathsf{b}(w)\,\mathcal{X}'(\lambda)-\mathcal{X}_1\Big(w\,\Lambda'(\lambda)+c^aK^a-\frac{\mathcal{P}^{\mathrm{s}}_\lambda}{\ell}\Big)+\frac{\mathcal{P}^{\mathrm{s}}_\lambda}{\ell}.$$

On the wage side, adding c_w and $\mathcal{P}_w/\ell$ replaces the delivered wage $\Lambda(\lambda)$ by the full one, $\Lambda(1)$, and adds the penalty opportunity cost on the gap. With $D+\vartheta B_\lambda=(1-\vartheta)D$, the mix condition in eq. (89) is

$$\begin{aligned}&\frac{\overbrace{\mathsf{b}(w)\,\mathcal{X}'(\lambda)}^{\text{compliance cost}}-\overbrace{\mathcal{X}_1\big(w\,\Lambda'(\lambda)+c^aK^a-\mathcal{P}^{\mathrm{s}}_\lambda/\ell\big)}^{\text{penalty opportunity cost saved}}+\overbrace{\mathcal{P}^{\mathrm{s}}_\lambda/\ell}^{\text{sanction saved}}}{\underbrace{(1-\vartheta)\,D}_{\substack{\text{delivered utility net}\\ \text{of backpay value}}}}\\ &\qquad=\frac{\overbrace{\Lambda(1)}^{\text{full wage bill}}+\overbrace{\big(\Lambda(1)-1\big)\mathcal{X}(\lambda)+\mathcal{X}_1\big(\Lambda(1)-\Lambda(\lambda)\big)}^{\text{opportunity costs}}}{\beta/w}.\end{aligned} \tag{104}$$

Compliance now trades its opportunity cost against the penalty opportunity cost and the sanction it saves. The wage costs the full wage bill $\Lambda(1)$ whatever the firm complies with. If $\vartheta=1$, the worker is made whole in expectation whatever the firm does, so $V_j=\beta\ln w+D$ does not depend on compliance. The left denominator is then zero, and eq. (87) reduces to $c_\lambda+\mathcal{P}_\lambda/\ell=0$: the firm chooses compliance to minimize cost plus penalty, the same condition as a firm that rations.

On the labor choice, the backpay value no longer depends on size, so inspection delivers no value to recruits and the last term of eq. (91) drops out. The per-worker backpay $\mathcal{P}^{\mathrm{bp}}/\ell=$

$w\big(\Lambda(1)-\Lambda(\lambda)\big)+c^aK^a(1-\lambda)$ tops the delivered bill up to the full one, so $c+\mathcal{P}^{\mathrm{bp}}/\ell = \Lambda(1)w+c^aK^a+\mathsf{b}(w)\mathcal{X}(\lambda)$, and substituting $\mathcal{P}_\ell$ from eq. (103),

$$\alpha z_j \ell^{\alpha-1} \;=\; \underbrace{\Lambda(1)\,w + c^aK^a}_{\text{full bill}} + \underbrace{\mathsf{b}(w)\,\mathcal{X}(\lambda) + \mathcal{X}_1\frac{\mathcal{P}^{\mathrm{bp}}}{\ell}}_{\text{opportunity costs}} + \underbrace{(1+\mathcal{X}_1)\mathcal{P}^{\mathrm{s}}_\ell}_{\text{sanction}} + \mu. \tag{105}$$

Every worker costs the full bill, as if the firm complied fully, plus the opportunity costs and the sanction. A firm that rations drops the μ term and sets its compliance by $c_\lambda + \mathcal{P}_\lambda/\ell = 0$, which is the case for the smallest firms at the estimate (section 5.1).

Remaining derivatives

The main text leaves four derivatives in symbolic form: Φ_λ and Φ_ℓ of the restored wage factor in eq. (77), and $\partial_\lambda Z_k$ and $\partial_\ell Z_k$ of the expected sanction base in eq. (69). Both closed forms are built from values of the count generator $M_k(x) = \big(\lambda+(1-\lambda)x\big)^\ell$ (eq. (51)), so one rule differentiates every piece. We write ∂ for either ∂_λ or ∂_ℓ when a formula holds for both.

The generator rule. Every generator value in eq. (77) and eq. (69) is $M_k(x)$ at an argument of the form $x = m_k^i e^{-\mathsf{t}}$. The index $i \in \{0,1\}$ marks whether the silence probability m_k is present, and the tilt $\mathsf{t} \in \{0, \gamma, \nu, \nu+\gamma\}$ is built from the resolution rate $\gamma(\ell) = \gamma_0\ell^{-\gamma_1}$ and the count curvature ν: $a_k = M_k(m_k)$, $b_k = M_k(e^{-\gamma})$, $d_k = M_k(m_k e^{-\gamma})$, and the count-factor values $M_k(e^{-\nu})$, $a_k^{(\nu)} = M_k(m_k e^{-\nu})$, $b_k^{(\nu)} = M_k(e^{-(\gamma+\nu)})$ and $d_k^{(\nu)} = M_k(m_k e^{-(\gamma+\nu)})$ of eq. (66). Compliance enters only through the base $\lambda+(1-\lambda)x$. Size enters through the exponent ℓ, through m_k via the complaint rate $\rho(\ell)$, and through $\gamma(\ell)$, with

$$\rho_\ell \;=\; \frac{\rho_1\rho}{\ell}, \qquad \gamma_\ell \;=\; -\frac{\gamma_1\gamma}{\ell}, \qquad \partial_\ell m_k \;=\; -\rho_\ell\Delta_k\, m_k, \qquad \partial_\ell x \;=\; -x\big(i\,\rho_\ell\Delta_k + \mathsf{t}_\ell\big), \tag{106}$$

where $\mathsf{t}_\ell = \gamma_\ell$ if the tilt contains γ and $\mathsf{t}_\ell = 0$ otherwise. The last equality differentiates $\ln x = i\ln m_k - \mathsf{t}$. Differentiating the base with respect to λ gives $1-x$, and differentiating $\ln M_k(x) = \ell\ln\big(\lambda+(1-\lambda)x\big)$ with respect to ℓ gives a term from the exponent and a term

from the argument, so

$$\begin{aligned}\partial_\lambda M_k(x) &= \ell\,(1-x)\,\big(\lambda+(1-\lambda)x\big)^{\ell-1},\\ \partial_\ell M_k(x) &= M_k(x)\Big[\underbrace{\ln\big(\lambda+(1-\lambda)x\big)}_{\text{one more worker}} + \underbrace{\frac{\ell\,(1-\lambda)\,\partial_\ell x}{\lambda+(1-\lambda)x}}_{\substack{\text{complaint and}\\ \text{resolution rates}}}\Big].\end{aligned} \tag{107}$$

For example, at $x = m_k$ the base is the unconditional silence ψ_k of eq. (19), and $\partial_\lambda a_k = \ell(1-m_k)\psi_k^{\ell-1}$, which is the term that drives eq. (97). Products over a set differentiate through their logarithm: for $x_A = \prod_{k\in A} x_k$,

$$\partial x_A \;=\; x_A \sum_{k\in A} \frac{\partial x_k}{x_k}, \tag{108}$$

which covers a_m, a_{-k} (eq. (50)), $b_{\mathcal{R}^0}$, $\bar\Phi^0_{\mathcal{R}^m}$ and the tilted products below. Every factor is strictly positive when $\lambda > 0$, and also at $\lambda = 0$ whenever $m_k > 0$, so the ratios are well defined.

The restored wage factor. The owed factors $\Phi^1_k = 1+\delta^w_k$ are constants. The delivered factor on the no-complaint event, $\bar\Phi^0_k = a_k + \delta^w_k \lambda\psi_k^{\ell-1}$ (eq. (76)), moves through a_k and through the silence of the $\ell-1$ coworkers. Its compliance derivative applies the product rule to $\lambda\psi_k^{\ell-1}$ with $\partial_\lambda\psi_k = 1-m_k$, and its size derivative differentiates $\psi_k^{\ell-1}$ through its logarithm:

$$\partial_\lambda\bar\Phi^0_k \;=\; \partial_\lambda a_k + \delta^w_k\Big[\underbrace{\psi_k^{\ell-1}}_{\text{worker 1 receives } k} + \underbrace{\lambda(\ell-1)(1-m_k)\,\psi_k^{\ell-2}}_{\text{coworkers silent}}\Big], \tag{109}$$

$$\partial_\ell\bar\Phi^0_k \;=\; \partial_\ell a_k + \delta^w_k\,\lambda\,\psi_k^{\ell-1}\Big[\ln\psi_k + (\ell-1)\frac{\partial_\ell\psi_k}{\psi_k}\Big], \tag{110}$$

with $\partial_\ell\psi_k = -(1-\lambda)\rho_\ell\Delta_k m_k$ as in eq. (100). Write $\bar\Phi_{\mathcal{R}^m} := (1-a_m)\,\Phi^1_{\mathcal{R}^m} + \bar\Phi^0_{\mathcal{R}^m}$ for the bracket of group m in eq. (77), the expected restored factor on the group's regulations, so that $\partial\bar\Phi_{\mathcal{R}^m} = -\partial a_m\,\Phi^1_{\mathcal{R}^m} + \partial\bar\Phi^0_{\mathcal{R}^m}$, with the set products differentiated by eq. (108). The closed form in eq. (77) is then $\Phi = \Phi^1_{\mathcal{R}^0}\prod_m \bar\Phi_{\mathcal{R}^m} + \big(\bar\Phi^0_{\mathcal{R}^0} - a_0\Phi^1_{\mathcal{R}^0}\big)\prod_m \bar\Phi^0_{\mathcal{R}^m}$, a product over groups plus a correction for the always-checked set when nobody complains. Applying the

product rule to each,

$$\begin{aligned}\partial\Phi \;=\; & \underbrace{\Phi^{1}_{\mathcal{R}^0}\sum_{m=1}^{M}\partial\bar{\Phi}_{\mathcal{R}^m}\prod_{m'\neq m}\bar{\Phi}_{\mathcal{R}^{m'}}}_{\text{groups opened or silent}} \;+\; \underbrace{\big(\partial\bar{\Phi}^{0}_{\mathcal{R}^0}-\partial a_0\,\Phi^{1}_{\mathcal{R}^0}\big)\prod_{m=1}^{M}\bar{\Phi}^{0}_{\mathcal{R}^m}}_{\text{always-checked set, no complaint}} \\ & +\; \underbrace{\big(\bar{\Phi}^{0}_{\mathcal{R}^0}-a_0\,\Phi^{1}_{\mathcal{R}^0}\big)\sum_{m=1}^{M}\partial\bar{\Phi}^{0}_{\mathcal{R}^m}\prod_{m'\neq m}\bar{\Phi}^{0}_{\mathcal{R}^{m'}}}_{\text{other groups, no complaint}}\,.\end{aligned} \tag{111}$$

Setting $\partial=\partial_\lambda$ gives Φ_λ in eq. (95), and $\partial=\partial_\ell$ gives Φ_ℓ in eq. (98).

The expected sanction base. The closed form in eq. (69) has three terms: the own-count moment $1-M_k(e^{-\nu})$, the never-surfaces term $\big(a_k-a_k^{(\nu)}\big)a_{-k}$, and the integral over the resolution position $\tilde{r}$ of the burden operator $\Gamma_k(\tilde{r})$ of eq. (68), all built from the four moments of the count factor in eq. (66). The integrand is a polynomial in $\tilde{r}$ whose coefficients are smooth in (λ,ℓ), so the derivative passes through the integral:

$$\partial Z_k \;=\; \underbrace{-\,\partial M_k\big(e^{-\nu}\big)}_{\text{own count}} \;-\; \underbrace{\Big(\big(\partial a_k-\partial a_k^{(\nu)}\big)\,a_{-k}+\big(a_k-a_k^{(\nu)}\big)\,\partial a_{-k}\Big)}_{\text{never surfaces}} \;-\; \underbrace{\int_0^1\partial\Gamma_k(\tilde{r})\,d\tilde{r}}_{\text{resolved in-window}}\,. \tag{112}$$

The four moments of eq. (66) are differences of generator values, so eq. (107) gives their derivatives,

$$\begin{aligned}\partial\,\mathbb{E}\big[\varsigma_k\big] &= -\partial M_k\big(e^{-\nu}\big), & \partial\,\mathbb{E}\big[e^{-\gamma n_k}\varsigma_k\big] &= \partial b_k-\partial b_k^{(\nu)},\\ \partial\,\mathbb{E}\big[\mathbb{1}(k\notin\mathcal{T})\,\varsigma_k\big] &= \partial a_k-\partial a_k^{(\nu)}, & \partial\,\mathbb{E}\big[\mathbb{1}(k\notin\mathcal{T})\,e^{-\gamma n_k}\varsigma_k\big] &= \partial d_k-\partial d_k^{(\nu)},\end{aligned} \tag{113}$$

The tilted factors in eq. (67) are affine in their generator values at a fixed position $\tilde{r}$, so $\partial b_j(\tilde{r})=\tilde{r}\,\partial b_j$ and $\partial d_j(\tilde{r})=(1-\tilde{r})\,\partial a_j+\tilde{r}\,\partial d_j$; the group weight differentiates term by term, $\partial\bar{\sigma}_m(\tilde{r})=\partial a_m+\partial b_{\mathcal{R}^m}(\tilde{r})-\partial d_{\mathcal{R}^m}(\tilde{r})$; and the products over sets follow eq. (108). For $k\in\mathcal{R}^m$, the burden operator is the product of the always-checked factor $b_{\mathcal{R}^0}(\tilde{r})$, the weights of the other groups, and k's own group bracket $\big(b_k-b_k^{(\nu)}\big)\,b_{\mathcal{R}^m\setminus\{k\}}(\tilde{r})-\big(d_k-d_k^{(\nu)}\big)\,d_{\mathcal{R}^m\setminus\{k\}}(\tilde{r})$. The product rule on the first two factors is a log-derivative times Γ_k, and on the bracket it is a

sum of four terms:

$$\begin{aligned}\partial\Gamma_k(\tilde r) \;=\Big[\;\; &\underbrace{\frac{\partial b_{\mathcal{R}^0}(\tilde r)}{b_{\mathcal{R}^0}(\tilde r)}}_{\text{always-checked set}} \;+\; \underbrace{\sum_{m'\neq m}\frac{\partial\bar\sigma_{m'}(\tilde r)}{\bar\sigma_{m'}(\tilde r)}}_{\text{other groups}}\Big]\Gamma_k(\tilde r)\\ &+\; b_{\mathcal{R}^0}(\tilde r)\prod_{m'\neq m}\bar\sigma_{m'}(\tilde r)\Big[\underbrace{\big(\partial b_k-\partial b_k^{(\nu)}\big)\,b_{\mathcal{R}^m\setminus\{k\}}(\tilde r)+\big(b_k-b_k^{(\nu)}\big)\,\partial b_{\mathcal{R}^m\setminus\{k\}}(\tilde r)}_{k\text{'s own group}}\\ &\qquad-\;\underbrace{\big(\partial d_k-\partial d_k^{(\nu)}\big)\,d_{\mathcal{R}^m\setminus\{k\}}(\tilde r)+\big(d_k-d_k^{(\nu)}\big)\,\partial d_{\mathcal{R}^m\setminus\{k\}}(\tilde r)}_{k\text{'s own group, nobody complains}}\Big].\end{aligned} \tag{114}$$

For $k\in\mathcal{R}^0$ the two terms of eq. (68) differentiate the same way, each as a product rule on its leading pair and a log-derivative of its product over groups,

$$\begin{aligned}\partial\Gamma_k(\tilde r) \;=\;&\Big[\big(\partial b_k-\partial b_k^{(\nu)}\big)\,b_{\mathcal{R}^0\setminus\{k\}}(\tilde r)+\big(b_k-b_k^{(\nu)}\big)\,\partial b_{\mathcal{R}^0\setminus\{k\}}(\tilde r)\Big]\prod_{m=1}^{M}\bar\sigma_m(\tilde r)\\ &+\;\big(b_k-b_k^{(\nu)}\big)\,b_{\mathcal{R}^0\setminus\{k\}}(\tilde r)\prod_{m=1}^{M}\bar\sigma_m(\tilde r)\sum_{m=1}^{M}\frac{\partial\bar\sigma_m(\tilde r)}{\bar\sigma_m(\tilde r)}\\ &-\;\Big[\big(\partial d_k-\partial d_k^{(\nu)}\big)\,d_{\mathcal{R}^0\setminus\{k\}}(\tilde r)+\big(d_k-d_k^{(\nu)}\big)\,\partial d_{\mathcal{R}^0\setminus\{k\}}(\tilde r)\Big]\prod_{m=1}^{M}a_m\\ &-\;\big(d_k-d_k^{(\nu)}\big)\,d_{\mathcal{R}^0\setminus\{k\}}(\tilde r)\prod_{m=1}^{M}a_m\sum_{m=1}^{M}\frac{\partial a_m}{a_m}.\end{aligned} \tag{115}$$

Each $\partial\Gamma_k$ is again a polynomial in $\tilde r$ of degree at most $|\mathcal{K}|-1$, since differentiating with respect to λ or ℓ changes the coefficients and not the degree, so the rule that evaluates the integral in eq. (69) exactly evaluates its derivative exactly. The sanction derivatives in the main text are then

$$\mathcal{P}^{\mathrm{s}}_{\lambda} \;=\; \ell^{\zeta}\sum_{k\in\mathcal{K}}g_k\,\partial_\lambda Z_k, \qquad \mathcal{P}^{\mathrm{s}}_{\ell} \;=\; \frac{\zeta\,\mathcal{P}^{\mathrm{s}}}{\ell}+\ell^{\zeta}\sum_{k\in\mathcal{K}}g_k\,\partial_\ell Z_k, \tag{116}$$

where the second differentiates the size scaling ℓ^{ζ} of the sanction schedule in eq. (14) as well as the bases. Under perfect enforcement $m_k=0$, so every argument that carries m_k is zero: a_k, d_k, $a_k^{(\nu)}$ and $d_k^{(\nu)}$ collapse to $M_k(0)=\lambda^{\ell}$, hence the two moments on the no-complaint event in eq. (66) are zero and $\partial_\ell m_k=0$, and the formulas above apply with these values.

B Contextual Details

B.1 *Size-Dependent Regulation*

Peru relieves small firms of part of the labor code at a single size threshold of 20 workers: firms with 20 or more workers must establish an occupational safety and health committee and an internal safety and health regulation, and firms with more than 20 owe their workers a share of profits.[24] We read the schedule off the law rather than estimating it, and the table below carries it with its statutory citations.

We split each of these regulations into units, and each unit is either owed in full by a size class or not at all. The units of a regulation sum back to it, so a class owes exactly the fraction the table assigns it. This splits the eleven regulations into thirteen units, and the parameters of a split regulation are apportioned across its units so that the compensation factor, the amenity bill and the worker's valuation of the package are unchanged.[25]

The size class then determines *penalties* and nothing else. Write $\chi_k(\ell) \in \{0, 1\}$ for whether unit k is citable at size ℓ. A shortfall on a unit with $\chi_k = 0$ is not a violation: no worker can complain of it, no inspection can surface it, and it carries neither backpay nor sanction. So that unit leaves $\mathcal{T}$, $\mathcal{R}$ and $\mathcal{A}$. However, every unit sits in every firm's package, so the delivered wage factor of eq. (10), the promised bill of eq. (11), the per-worker cost of eq. (8) and the worker's valuation Δ_k of eq. (82) are all continuous in ℓ. Only the expected penalty is a step function.

The firm's objective is discontinuous in ℓ, so maximized profit is the upper envelope of two sub-problems, one on each side of the threshold. But the discontinuity now sits in $\mathcal{P}$ alone, and $\mathcal{P}$ is collected only when the firm is inspected. The notch on crossing the threshold is therefore determined by the step in the penalty weighted by the inspection probability $\mathscr{I}(\ell)$ of eq. (21). In contrast to work in contexts with strong enforcement (Garicano et al., 2016), low enforcement rates mean that a firm has little incentive to bunch at the threshold.

The rest of this subsection carries the schedule itself: the statutory table, the split into units, and the apportionment rule for each parameter.

[24] There are other thresholds based on revenue, but our only observable measure of size is number of workers.

[25] Both split regulations are amenities, which enter linearly, so their utilities and delivery costs split in proportion to each unit's share. The apportionment rule for each parameter, and the aggregates it holds fixed, are given below.

	Unit	u_k	$\ell < 20$	$\ell \geq 20$	Statute
1	Severance deposit (CTS)	1	1	1	D.S. 001-97-TR
2	Gratuities	1	1	1	Ley 27735
3	Health insurance	1	1	1	Ley 26790
4	Pension	1	1	1	D.L. 19990; Ley 29741
5	Wage timeliness	1	1	1	D.S. 003-97-TR
6	Working time, rest and leave	1	1	1	D.L. 713
7	Fundamental rights	1	1	1	Ley 27942
8	Workplace safety and health	1	1	1	Ley 29783, arts. 28–36
9a	Safety management: plan, accident notices, audits	0.051	1	1	Ley 29783
9b	Safety management: committee and internal regulation	0.949	0	1	Ley 29783, arts. 29, 34
10	Labor formalities	1	1	1	D.S. 018-2007-TR
11a	Social security and indirect benefits, other	0.779	1	1	
11b	Social security and indirect benefits: profit sharing	0.221	0	1	D.L. 892, art. 9

Table 11: The size schedule. Each unit is owed in full (1) or not at all (0) by a size class, and u_k is its share of its parent regulation, numbered as in table 2. The shares of each parent sum to one.

The schedule

Three rules turn on with head count in 2016, all at twenty workers. The safety and health committee (Ley 29783, art. 29) and the internal safety and health regulation (art. 34) bind at twenty or more workers, and profit sharing (D.L. 892, art. 9) binds at more than twenty. We place the threshold for all three at twenty, where the first two bind, so the classes are $\ell < 20$ and $\ell \geq 20$. The intervals partition strictly, since profit is not monotone in the citable set and a shared endpoint could be won by the wrong class's bundle. Relief for micro and small firms has depended on annual sales rather than head count since 2013, so the former thresholds at 10 and 100 workers no longer apply.

None of the three rules is a regulation of its own in the violations data. The committee and the internal regulation are cited under safety and health management and records, and profit sharing under social security and indirect benefits. Each of these two regulations therefore splits into a unit owed at every size and a unit owed only from twenty workers, and table 11 lists the thirteen units.

The split into units

Each partially relieved regulation is split by the layer-cake decomposition of its own schedule: one unit per distinct positive level, with each unit carrying the slab between consecutive levels and citable in exactly those classes that owe at least that level. With two classes, a

split regulation has one unit owed at every size and one owed only from twenty workers. The share u_k of the second unit is its statute's share of the parent's shorted-worker burden at firms of twenty or more workers, where the statute is owed. We measure it on the sanctions of complaint inspections from 2014 to 2019. At firms of twenty or more, citations of the committee and the internal regulation cover 171,625 of the 180,889 shorted workers cited under safety and health management, a share of 0.95, and citations of profit sharing cover 3,226 of the 14,592 shorted workers cited under social security, a share of 0.22. Some citations below twenty still carry these statutes' names: the committee citation also covers the supervisor who serves in its place in a smaller firm. The citable unit of each regulation absorbs them, so every regulation keeps a citable unit on both sides of twenty and no recorded violation has probability zero.

Apportionment

Each unit inherits its parent's parameters under a rule that leaves the worker's package and the firm's bill unchanged. No compensation regulation splits, so the rates δ_k^w and the promised bill of eq. (11) are unchanged. An amenity unit carries the share u_k of its parent's value, $\beta_k^a = u_k\,\beta_{\text{parent}}^a$, so the worker's valuation D of the package is unchanged. The unit also carries u_k as a weight on its delivery cost and on its resolution burden: a unit costs $u_k c^a$ to deliver, and a worker shorted on it adds u_k to the burden $\mathcal{N}$ of eq. (27). The amenity bill $c^a \sum_k u_k$ and the burden are therefore unchanged by the split. Each unit carries its parent's full sanction severity g_k. This is a modelling choice: a citation of the committee alone is one sanction for a safety management violation, not a fraction of one. It raises the sum of severities over the units by 13%, and splitting the severity in proportion to u_k would keep the sum unchanged.

Residual channels

The schedule is not exactly neutral in the enforcement block, in two places, and both scale with enforcement. We report them at the estimate for a firm that pays $0.97\,w_0$ with compliance 0.85.[26] First, a worker shorted on a unit that is not citable has nothing to recover and files no complaint. Crossing twenty therefore adds the two relieved units to the firm's complaint products: the inspection probability rises by 0.0031 at the threshold and the backpay

[26] Model 19.7_smm, `check_size_regs` at the estimate; the log is saved with the counterfactual results.

value by 26%. This jump is part of the notch of section 5.3. Second, splitting a regulation gives a worker two independent occasions to complain where there was one, so at full citability the split economy is not numerically identical to the unsplit one. At the same wage and compliance, the split changes the inspection probability by less than 0.01% at every size, and it changes the expected penalty by −1.8% at five workers, −0.8% at twelve, −0.1% at forty and +0.07% at three hundred.

B.2 *Bunching at Twenty Workers*

We measure bunching in the 2016 tax records (SUNAT), with one row per firm and 330,830 firms. Size is the firm's reported number of employees, floored at one and interpolated across years for a firm that skipped a filing. The floor pools the firms that report no employees with the genuine one-employee firms, so we drop size one from every fit. Twenty is the only head-count threshold in force in 2016: the safety and health committee and the internal safety regulation bind at twenty or more workers and profit sharing at more than twenty, so nineteen is the last size with no new mandate and we place the threshold at $L^* = 20$, the first regulated size.

We follow the excess-mass estimator of Chetty et al. (2011). Let c_L be the number of firms of size L. On the fit range $L \in [L^* - R,\, L^* + R]$ we regress

$$\ln c_L \;=\; \alpha + \sum_{p=1}^{Q} \beta_p \Big(\frac{L - L^*}{10}\Big)^p + \big(\gamma_5\, R5_L + \gamma_{10}\, R10_L\big)\big(1 - W_L\big) + \sum_{j=L^*-N}^{L^*+M} d_j\, \mathbb{1}[L = j] + e_L, \quad (117)$$

where $W_L := \mathbb{1}[L^* - N \le L \le L^* + M]$ marks the excluded window, the N sizes below the threshold where firms pile up and the M sizes above it that they vacate, and $R5_L$ and $R10_L$ indicate sizes that are multiples of five but not ten and multiples of ten. The window dummies d_j keep the bins inside the window from shaping the fitted curve, and the round-number terms absorb heaping at round sizes and are identified from bins outside the window. We use logs because the size density is close to Pareto, so $\ln c_L$ is nearly linear in L over a short range. The counterfactual count $\hat{c}_L$ switches the window dummies off and keeps the round-number terms on at every size, so heaping at twenty is not scored as bunching:

$$\hat{c}_L \;=\; \exp\Big(\hat{\alpha} + \sum_{p=1}^{Q} \hat{\beta}_p \Big(\frac{L - L^*}{10}\Big)^p + \hat{\gamma}_5\, R5_L + \hat{\gamma}_{10}\, R10_L\Big).$$

The excess mass is the extra firms below the threshold in units of one counterfactual bin,

$$b \;=\; \frac{\sum_{L=L^*-N}^{L^*-1} \left(c_L - \hat{c}_L\right)}{\frac{1}{N}\sum_{L=L^*-N}^{L^*-1} \hat{c}_L}. \tag{118}$$

Its standard error comes from a residual bootstrap with 1,000 replications: we resample the residuals of the bins outside the window, add them to the full fitted values with the window dummies on, so that the bunching in the data survives in the bootstrap samples, and re-estimate b.

We choose the polynomial degree Q, the half-width R of the fit range and the window (N, M) without looking at twenty. Over the 81 combinations of Q, R, N, M in $\{2,3,4\} \times \{8,10,12\} \times \{2,3,4\} \times \{2,3,4\}$, we take the one that minimizes the largest absolute placebo estimate at sizes 11, 15, 25, 30, 40, 101 and 200. It is $Q = 4$, $R = 12$, $N = 3$ and $M = 4$, so the window runs from 17 to 24 workers. At that choice $b = 0.64$ (0.25). The estimate is positive in all 81 combinations, and none of the placebo estimates is significant. Applied to the French firm-size counts of Garicano et al. (2016) at their fifty-worker threshold, the same estimator gives $b = 1.5$.

C Estimation Details

Table 12 lists every parameter, the block that sets it and how.

C.1 *Block 1: Statutory and Institutional Parameters*

We index the eleven regulations of section 4.1 in the order of table 2 throughout.

Regulations left out. Four compensation categories in the data are not modeled as separate regulations: the minimum wage, indemnities and severance payments, allowances, and bonuses. Together they make up about 6.5% of recorded violations. For each of them we either cannot observe the premium over the base wage or the premium does not scale the wage, so no rate δ_k^w can represent them. We also exclude night work and overtime due to low coverage. Night work accounts for 0.14% of recorded shorted workers and overtime for 1.8%.

We keep social security and indirect benefits, whose profit-sharing component applies

Symbol	Object	Where	How
K^w, K^a	Compensation / amenity counts	Block 1	Classification of the violations data
δ_k^w	Compensation rate of regulation k	Block 1	Statutory rates
g_k	Sanction severity, in soles	Block 1	Mean severity class, interpolated
$\nu,\ \zeta$	Sanction count curvature, size elasticity	Block 1	NLS on 2016 sanction records
$L,\ J,\ N_F$	Workforce, formal firms, formal jobs	Block 1	ENAHO 2016; employer–employee panel
$w_0,\ w^r$	Informal wage, minimum wage	Block 1	ENAHO 2016; statutory value
$\{\mathcal{R}^m\}_{m=0}^M$	Related groups partition	Block 2	Estimated positions on a line, then held fixed
$\tilde{\rho}_0 = \rho_0\beta$	Complaint-hazard scale at $\ell = 1$	Block 2	MLE
ρ_1	Size elasticity of the complaint hazard	Block 2	MLE
$\gamma_0,\ \gamma_1$	Resolution rate, its size gradient	Block 2	MLE
$\tilde{\beta}_k^a = \beta_k^a/\beta$	Amenity preference (scaled)	Block 2	MLE
λ_b	Compliance rate in size bin b	Block 2	MLE (limited)
α	Returns to scale	Block 3	SMM: profit share of value-added
$\mu_z,\ \sigma_z$	Pareto lower bound and tail index	Block 3	SMM: firm size distribution
ξ_0	Informal-sector taste level	Block 3	SMM: informal share of employment
β	Wage weight in worker utility	Block 3	SMM: wage–size profile
ϕ	Compliance cost scale	Block 3	SMM: compliance level
c^a	Per-unit amenity delivery cost	Block 3	SMM: compliance level
ϑ	Backpay weight	Block 3	SMM: compliance curvature at top

Table 12: The full parameter inventory.

only above 20 workers. We group this with the other amenities, due to the non-profit sharing components and because profit-sharing is not easily incorporated into the way we model regulated compensation.

The sanction sample. We estimate the sanction function using the 2016 sanction records merged with firm records from the tax authority (SUNAT). We keep sanctions from complaint-driven inspections, with one record per firm, year, regulation and regulation subcategory. The dependent variable is the sanction the firm actually paid, in soles. When one sanction covers several regulations, we divide the paid amount among them in proportion to their shares of the proposed sanction. Head count is grouped into $\ell \in \{10, 100, 1000\}$, and n is the number of shorted workers on the record. Dropping records with a nonpositive sanction or count leaves 8,010 records.

Each record carries a severity class $h \in \{1, 2, 3\}$ (Minor, Serious, Very Serious). Taking logs of eq. (14) and replacing the severity with a separate level s_h for each class gives

$$\ln \mathsf{S} \;=\; \zeta \ln \ell \;+\; \ln s_h \;+\; \ln\left(1 - e^{-\nu n}\right), \tag{119}$$

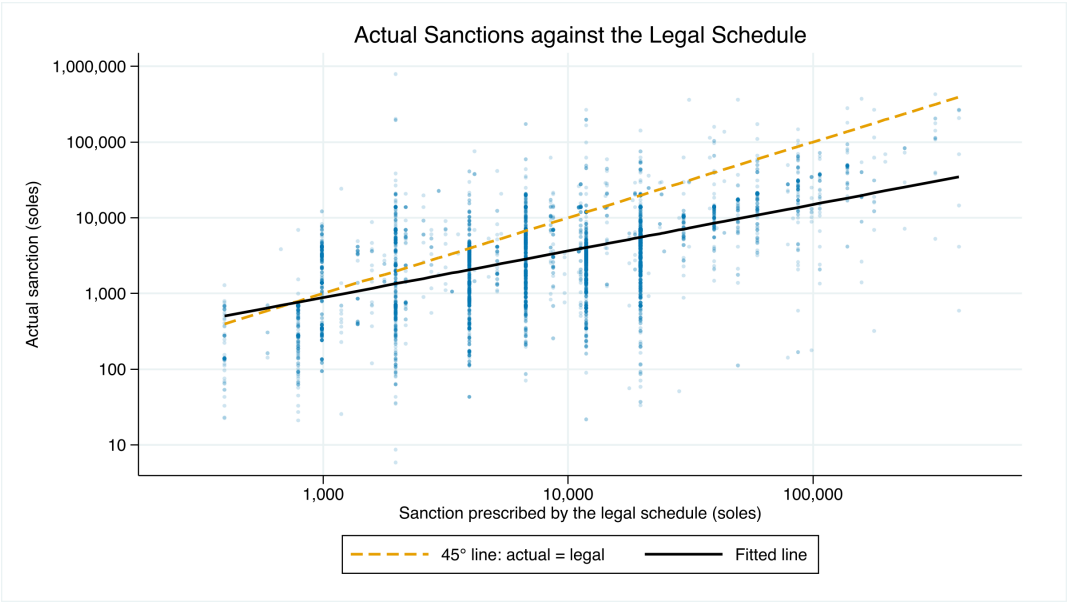


Figure 13: Realized sanctions against the sanction the legal table prescribes, 2016, on log scales. The fitted line is the regression of the log realized sanction on the log value in the legal table.

which we estimate by nonlinear least squares on the log paid sanction. The estimates, with standard errors in parentheses, are

$$\hat{\zeta} = 0.15 \ (0.02), \qquad \hat{\nu} = 0.48 \ (0.043), \qquad (\hat{s}_1, \hat{s}_2, \hat{s}_3) = (400, \ 2{,}442, \ 4{,}085) \text{ soles}.$$

All five are significant and $\hat{\nu}$ is stable across starting values.

Two regressions check that the schedule is predictable. Regressing the log realized sanction on the log value in the legal table gives a slope of 0.61 (0.026) with $R^2 = 0.32$ (fig. 13). For the fitted schedule in eq. (119), the same regression gives a slope of 0.55 (0.008) with $R^2 = 0.66$. Slopes below one mean that realized sanctions compress the spread of the legal table, and the fit of the second regression means that the schedule follows the table closely.

The severity g_k in eq. (14) is in soles, while the mean severity class $\bar{h}_k \in [1, 3]$ of regulation k is an average over its recorded violations in the full sample of inspections triggered by a complaint. We interpolate linearly between the two estimated levels that bracket $\bar{h}_k$. With $m = \lfloor \bar{h}_k \rfloor$, and $m = 2$ when $\bar{h}_k = 3$,

$$g_k \ = \ (m + 1 - \bar{h}_k)\, \hat{s}_m \ + \ (\bar{h}_k - m)\, \hat{s}_{m+1}, \tag{120}$$

Aggregate	Symbol	Value	Source
Total employment	L	14,719,456	ENAHO 2016, survey-weighted
Formal firms	J	332,579	Employer–employee panel, 2016
Formal employment	N_F	3,464,773	Employer–employee panel, 2016
Informal wage, monthly	w_0	S/ 843	ENAHO 2016, mean informal earnings
Minimum wage, monthly	w^r	S/ 817	2016 statutory minimum, annual mean

Table 13: Aggregates for 2016. The informal wage w_0 is the numeraire. Formal employment is $N_F/L = 0.24$ of the total, and $w^r/w_0 = 0.968$.

which returns $\hat{s}_h$ exactly at each integer class and is continuous and piecewise linear between classes. For pension, $\bar{h}_k = 2.8$ gives $g_k = 0.221 \times 2{,}442 + 0.779 \times 4{,}085 \approx 3{,}721$ soles. A sanction is a lump sum paid once per inspection, while every other monetary object in the model is a monthly flow and the inspection probability in eq. (21) is per firm-year. We therefore divide g_k by twelve months of the informal wage, $12\, w_0$, when it enters the firm's problem.

Aggregates. Table 13 reports the aggregates and their sources.

We compute L from the employment module (module 500) of the 2016 national household survey (ENAHO), pooling its twelve monthly samples. A person counts as employed if the survey classifies them as occupied in the reference week (`ocu500`= 1), and we keep the household's habitual residents. Weighting each person by the module's expansion factor (`fac500a`) gives 16,197,110 employed workers in 2016. The model has no public employer, so we subtract the 1,477,654 wage earners whose employer is the armed forces or police, the public administration or a public enterprise (question `p510`), which leaves $L = 14{,}719{,}456$. We take L and N_F from the same year, so the formal share N_F/L is 0.24. The survey's own classification of formal employment, applied to the same non-government workers, gives a formal share of 0.231.

We obtain J, N_F and the firm size distribution from the employer–employee panel,which agrees with the tax records. The employer–employee panel (the *planilla electrónica*) holds $J = 332{,}579$ firms with at least one worker in some month of 2016, and N_F sums each firm's mean monthly head count over the months it appears. The tax records (SUNAT) hold 330,830 private employers with a positive head count in 2016, within 0.5% of the panel's count. The two registries share no firm identifier, so we compare their totals only.

The statutory minimum wage was S/ 750 until 1 May 2016 and S/ 850 afterwards, which

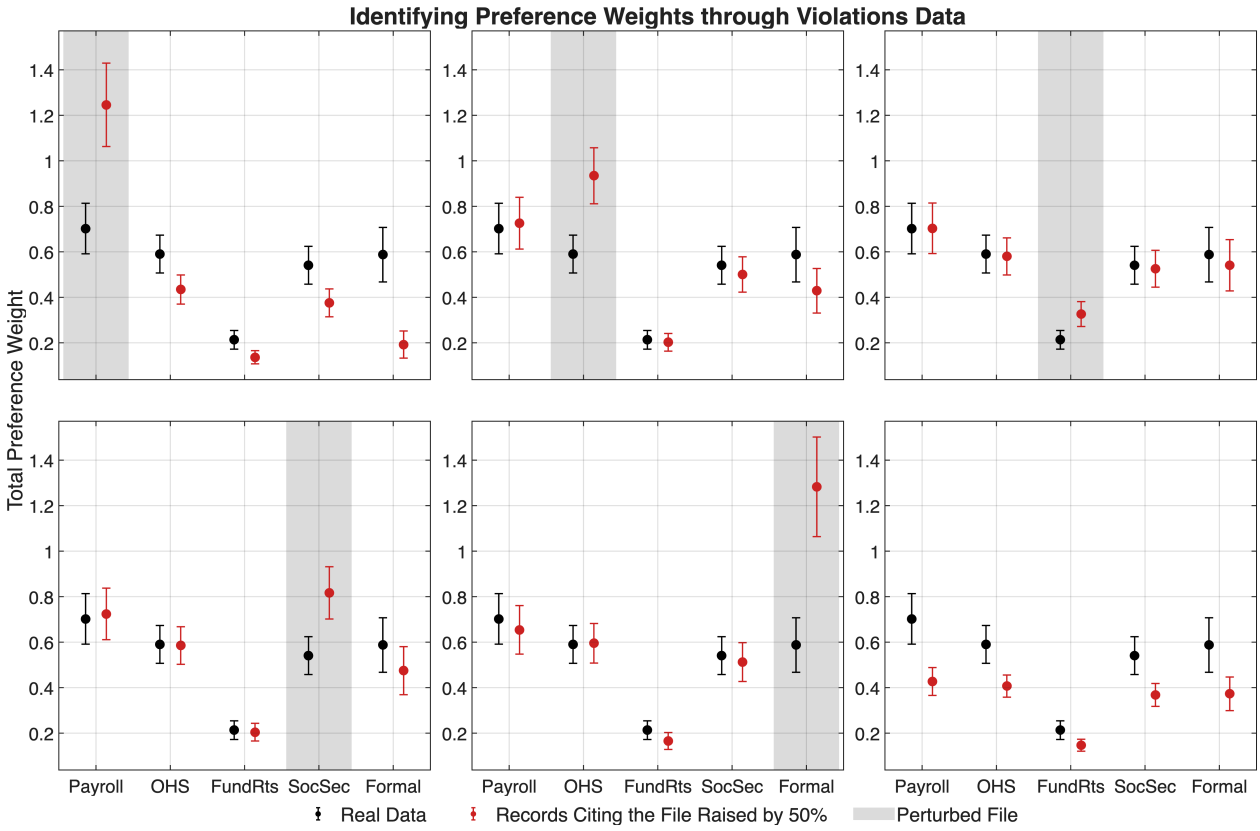


Figure 14: Amenity values re-estimated after raising the share of violation records citing one related group by 50%. Black is the estimate on the real data and red the estimate on the perturbed data, with 95% confidence intervals. The perturbed column is highlighted in gray on each panel's axis. The final panel perturbs regulated compensation, whose value is fixed by statute.

gives an annual mean of $w^r = (4 \times 750 + 8 \times 850)/12 = 817$. We use the annual mean because the wage profile that w^r is compared with in Block 3 is also an annual mean. At this value w^r almost exactly equals the wage in the two smallest size bins, since those firms pay the minimum wage all year.

C.2 *Block 2: The Complaint and Inspection Process*

Related groups. To estimate the related groups, we place each regulation k at a position u_k on a line and let a complaint pull in the regulations near it. Given the set $\mathcal{T}$ of regulations complained about, regulation j goes unexamined with probability

$$\mathbb{P}\big(j \text{ not examined} \mid \mathcal{T}\big) \;=\; (1 - b_j) \prod_{k \in \mathcal{T}} \Big(1 - e^{-(u_k - u_j)^2}\Big), \tag{121}$$

Regulation	**Group**	u_k	b_j
Severance deposit (CTS)	$\mathcal{R}^1$	0.00	0.00
Gratuities	$\mathcal{R}^1$	0.00	0.00
Wage timeliness	$\mathcal{R}^1$	0.04	0.03
Working time, rest and leave	$\mathcal{R}^1$	−0.02	0.03
Health insurance	$\mathcal{R}^2$	1.56	0.00
Pension	$\mathcal{R}^2$	1.60	0.00
Workplace safety and health	$\mathcal{R}^3$	12.51	0.04
Safety and health management	$\mathcal{R}^3$	12.65	0.00
Fundamental rights	$\mathcal{R}^4$	14.21	0.02
Social security	$\mathcal{R}^5$	25.16	0.17
Labor formalities	$\mathcal{R}^0$	11.15	0.75

Table 14: Estimated positions and always-check probabilities. Positions u_k on the line of eq. (121), with the severance deposit at zero and the scale of distances set to one, and the probabilities b_j that a regulation is checked on every visit. The group column is the partition we set from these estimates.

where $e^{-(u_k-u_j)^2}$ is the probability that a complaint about k pulls j into the investigation, equal to one at $k = j$ and falling with the distance between the two positions, and b_j is the probability that j is checked on every visit whatever the complaint. Two regulations at distance 0.14 are examined together with probability 0.98, at distance 0.83 with probability 0.5, and at distance 2 with probability 0.02. We normalize the severance deposit's position to zero and the scale of distances to one, and estimate the positions u_k and the probabilities b_j by maximum likelihood before the rest of Block 2.

Table 14 reports the estimates. Regulations whose positions clump together form a related group: the four payroll and leave regulations sit within 0.04 of zero, health insurance and pension near 1.6, and the two safety regulations near 13, while fundamental rights and social security each sit apart from every other regulation. Only one regulation is checked on most visits whatever the complaint: labor formalities is checked on 0.75 of visits whatever the complaint, against at most 0.17 for any other regulation, so it forms the always-checked set $\mathcal{R}^0$. We set these groups exogenously and estimate the rest of Block 2 holding them fixed.

The sanction gate. Figure 15 plots the probability that a first complaint inspection ends in a sanction against firm size. The resolution layer of section 4.2.3 fits this curve: the burden of shorted workers grows with size, while the resolution rate $\gamma(\ell) = \gamma_0 \ell^{-\gamma_1}$ (eq. (24)) grows with it too, and the estimates of γ_0 and γ_1 are what set the balance between them.

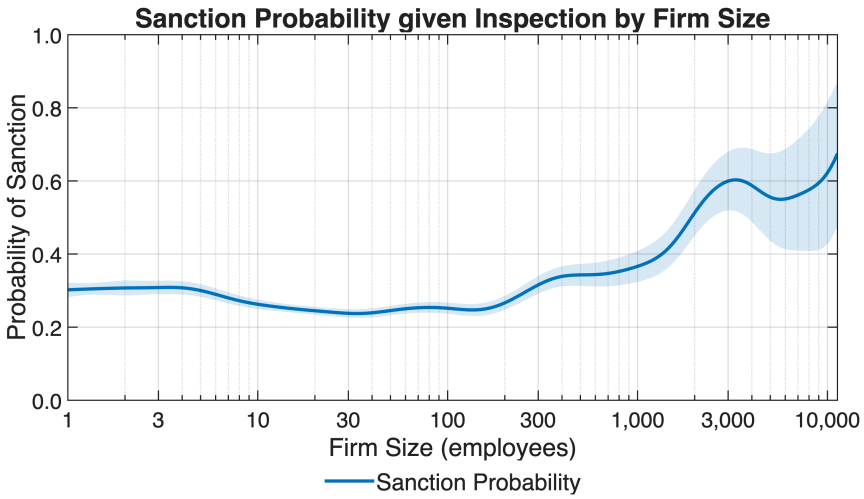


Figure 15: The probability that a first complaint inspection ends in a sanction, by firm size. Points are means within 0.25 log-size bins and the line is a cubic logit in log head count.

Size bins. The fifteen size bins of section 4.3 have lower edges

$$1,\ 3,\ 4,\ 5,\ 11,\ 21,\ 30,\ 40,\ 50,\ 101,\ 150,\ 200,\ 300,\ 500,\ 1000.$$

The smallest bin combines firms with one and two workers. At one worker the count of shorted workers can only be zero or one, so it carries no information about the shape of the binomial count law used to estimate compliance. The edges at 11 and 101 workers date from the former head-count thresholds of relief for small firms, and the threshold at 20 workers of appendix B.1 falls at the top of the 11–20 bin. The compliance and wage moments of Block 3 use the same bins, so the bins are chosen once, before any estimation.

The estimation sample. Table 15 reports the sample. Each firm contributes its years up to and including its first inspection, so the categorical log-likelihood uses every firm-year and the violations log-likelihood uses only the sanctioned firm-years whose recorded set we observe.

From the violation records to the count cells. Table 16 traces every violation record of 2014–2019 into the cells the violations log-likelihood fits. We drop three kinds of record before the data are the model's object. Pure obstruction is a procedural charge against the inspector rather than compensation owed to a worker, so it carries no regulation in the model. Planned inspections are not complaint-triggered, and the complaint process is what

Firm-years	**Count**
First-inspection panel, 2014–2019	1,965,447
Not inspected ($d = 1$)	1,943,360
Inspected, no sanction ($d = 2$)	16,067
Inspected and sanctioned ($d = 3$)	6,020
Sanctioned, recorded set observed	3,885

Table 15: The Block 2 estimation sample. Each firm contributes its years up to and including its first inspection. The categorical log-likelihood uses all firm-years, and the violations log-likelihood uses the sanctioned firm-years whose recorded set is observed.

Violation records, 2014–2019	**Records**	**Firms**
All records on inspections at formal firms	77,503	
less pure obstruction	53,550	
less planned-origin inspections	49,393	
less counts above the firm's head count	46,992	
The violations data	46,992	
on the 11 modeled regulations	42,330	6,252
at a firm's first inspection		3,885
Count cells fitted by $\mathscr{L}_{\text{viol}}$	9,229	3,885

Table 16: The violations data funnel. A record is one regulation cited at one inspection. The firm column is reported only where it is the binding count.

section 4.2.1 describes. A record whose shorted-worker count exceeds the firm's head count cannot be a draw from $n_k \sim \text{Binomial}(\ell, 1-\lambda)$, so we drop it as a measurement error. What remains is the violations data.

Of that, the eleven modeled regulations carry 90% of the records and 97% of the shorted-worker records, so the regulations we leave out are the ones inspectors cite against few workers at a time. The largest unmodeled group is compensation: minimum wage, indemnities and severance, allowances and bonuses together account for 6.5% of the violations data. The last step is the one that separates the violations sample from the panel. The modeled records sit in 9,318 firm-years at 6,252 distinct firms, but only a firm's *first* inspection is in the panel, so records from later inspections leave. Folding 3,553 repeat citations of the same regulation within a firm-year into counts then turns 3,885 recorded sets into the 9,229 count cells of $\mathscr{L}_{\text{viol}}$.

Sanctioned firm-years without a recorded set. Of the 6,020 first inspections that end in a sanction, we observe a usable recorded set for 3,885. The rest leave at a named gate of table 16, and in order of size the reasons are that no record maps to a modeled regulation, that the shorted-worker count exceeds the firm's head count, that the sanction came from a planned rather than a complaint inspection, and that the only records are night work or overtime. None of these gates reads the recorded set itself: the first three are properties of the inspection or of the regulation cited, and the fourth is the regulation set we chose in Block 1. We therefore treat the observed recorded sets as a random sample of the sanctioned firm-years, so O_{jt} is uncorrelated with the recorded set and $f(\omega \mid d = 3)$ is the likelihood of every observed recorded set. A sanctioned firm-year without a recorded set ($O_{jt} = 0$) still contributes, through $\ln P^3$ alone, as the last line of eq. (122).

Contributions by outcome. Firm-year (j, t) contributes

$$\mathscr{L}_{jt} \;=\; \begin{cases} \ln P^1_{jt} & d_{jt} = 1, \\ \ln P^2_{jt} & d_{jt} = 2, \\ \ln P^3_{jt} + \ln f(\omega_{jt} \mid d_{jt} = 3) & d_{jt} = 3,\ O_{jt} = 1, \\ \ln P^3_{jt} & d_{jt} = 3,\ O_{jt} = 0, \end{cases} \tag{122}$$

and the firm's contribution $\mathscr{L}_j$ is the sum of these terms over its years in the sample. Summing over firms splits the log-likelihood into its two parts,

$$\mathscr{L} \;=\; \underbrace{\sum_j \sum_{t=t^0_j}^{\tau_j} \ln P^{d_{jt}}_{jt}}_{\mathscr{L}_{\text{cat}}} \;+\; \underbrace{\sum_j \sum_{t=t^0_j}^{\tau_j} O_{jt}\, \ln f(\omega_{jt} \mid d_{jt} = 3)}_{\mathscr{L}_{\text{viol}}}, \tag{123}$$

with the inner sum running over all of the firm's panel years when it is never inspected. Every term has the closed form of proposition 1, so we evaluate the likelihood exactly and without simulation. The estimation treats every regulation as citable at every firm size, which is proposition 1 without the schedule of appendix B.1.

Two-stage estimator. Write $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_{15})$ for the compliance rates of stage A. The eleven parameters of stage B are θ_2 of eq. (30): $\tilde{\rho}_0$, ρ_1, γ_0, γ_1 and the seven amenity values

$\tilde{\beta}^a_k$. We write $\Theta_2 := (\boldsymbol{\lambda}, \theta_2)$ for the full vector that Block 2 estimates. The amenity values sit in stage B because they act on the probability of a complaint, which the categorical outcomes measure, and the violations log-likelihood alone does not bound them. Stages A and B in eq. (31)–eq. (32) alternate until no parameter changes by more than 10^{-4} in relative terms, and the estimate is the fixed point of this alternation. We start the alternation from four points scattered around a starting guess built from the data, and no parameter of the reported estimate lies at a numerical bound. All four starts converge to the same estimate: their log-likelihoods differ by less than 0.02, and every parameter agrees to at least four significant digits.

Coordinates free of β. The likelihood depends on worker preferences only through the exponent $\rho(\ell)\Delta_k$ of the silence weight in eq. (19). Substituting eq. (3) and eq. (82),

$$\rho(\ell)\,\Delta_k \;=\; \begin{cases} \tilde{\rho}_0\,\ell^{\rho_1}\ln\left(1+\delta^w_k\right) & k \in \mathcal{K}^w, \\ \tilde{\rho}_0\,\ell^{\rho_1}\,\tilde{\beta}^a_k & k \in \mathcal{K}^a, \end{cases} \qquad \tilde{\rho}_0 \;:=\; \rho_0\,\beta, \qquad \tilde{\beta}^a_k \;:=\; \beta^a_k/\beta, \tag{124}$$

so β does not enter the likelihood on its own and Block 2 estimates $\tilde{\rho}_0$ and $\tilde{\beta}^a_k$.

Inference. A firm contributes every year up to its first inspection, so the 1,965,447 firm-years of table 15 are not independent observations. We assume independence across firms and cluster on the firm. The estimate solves the first-order conditions of two different objectives, so its variance is the sandwich for the stacked conditions rather than an inverse information matrix. Stage A maximizes $\mathscr{L}_{\text{viol}}$ and stage B maximizes $\mathscr{L}$, and the estimate is the fixed point at which neither stage moves. We work in the coordinates of the search: $\operatorname{logit}\lambda_b$ for compliance, and logs for θ_2 except ρ_1, which stays in levels. Write $\mathbf{s}_j$ for firm j's contribution to the two stage gradients, stacked in these coordinates and summed over its years,

$$\mathbf{s}_j(\Theta_2) \;:=\; \begin{pmatrix} \partial\mathscr{L}_{\text{viol},j}/\partial \operatorname{logit}\boldsymbol{\lambda} \\ \partial\mathscr{L}_j/\partial\theta_2 \end{pmatrix}, \qquad \mathbf{s} \;:=\; \sum_j \mathbf{s}_j. \tag{125}$$

The upper block is the stage A condition in eq. (31) and the lower block is the stage B condition in eq. (32), so the fixed point of the alternation is the $\hat{\Theta}_2$ at which both blocks of $\mathbf{s}$ vanish.

Expanding $\mathbf{s}(\hat{\Theta}_2) = \mathbf{0}$ around the truth writes the estimate as a linear function of the

stacked score,

$$\mathbf{0} \;=\; \mathbf{s}(\hat{\Theta}_2) \;\approx\; \mathbf{s}(\Theta_2) \;-\; \mathbf{H}\,\big(\hat{\Theta}_2 - \Theta_2\big), \qquad \mathbf{H} \;:=\; -\,\partial \mathbf{s}/\partial \Theta_2',$$

so $\hat{\Theta}_2 - \Theta_2 \approx \mathbf{H}^{-1}\mathbf{s}(\Theta_2)$, and the variance of the estimate is the variance of the stacked score carried through $\mathbf{H}^{-1}$ on both sides. Let $\mathbf{I}$ and $\mathbf{I}_{\text{viol}}$ be the negative Hessians of $\mathscr{L}$ and $\mathscr{L}_{\text{viol}}$ in these coordinates at the estimate, with blocks indexed by λ and B. The Jacobian $\mathbf{H}$, the clustered variance $\mathbf{\Omega}$ of the stacked score, and the variance $\mathbf{V}$ of the estimate are

$$\mathbf{H} \;=\; \begin{pmatrix} \mathbf{I}_{\text{viol}}^{\lambda\lambda} & \mathbf{I}_{\text{viol}}^{\lambda B} \\ \mathbf{I}^{B\lambda} & \mathbf{I}^{BB} \end{pmatrix}, \qquad \mathbf{\Omega} \;=\; \sum_j \mathbf{s}_j\,\mathbf{s}_j', \qquad \mathbf{V} \;=\; \mathbf{H}^{-1}\,\mathbf{\Omega}\,\mathbf{H}^{-\top}. \tag{126}$$

The rows of $\mathbf{H}$ follow the objective that each stage differentiates. Stage A differentiates $\mathscr{L}_{\text{viol}}$, so both blocks of its row come from $\mathbf{I}_{\text{viol}}$, and stage B differentiates $\mathscr{L}$, so both blocks of its row come from $\mathbf{I}$. The lower-left block is therefore not the transpose of the upper-right one, and an asymmetric $\mathbf{H}$ is what rules out an inverse information matrix even before clustering.

We form $\mathbf{\Omega}$ from the scores themselves rather than from the information identity, which would set $\mathbf{\Omega}$ to a matrix of negative Hessians and let eq. (126) collapse. The identity holds only if the model is correctly specified, and the shorted counts are overdispersed relative to the binomial law of eq. (15) that generates them: a firm cited on one regulation is cited on its related regulations in the same year, and the same firm is cited again in later years. Both are dependence the likelihood does not model, and both inflate the true variance of the compliance score without touching its Hessian. Under the information identity the compliance standard errors are smaller than the clustered ones by more than an order of magnitude, so we report the clustered sandwich throughout. The outer product in eq. (126) sums over firms rather than firm-years, which is what carries the within-firm dependence.

Block 3 receives the compliance rates as data moments in natural units, and table 4 reports every parameter in natural units, while eq. (126) is in the coordinates of the search, so we convert $\mathbf{V}$ by the delta method. The delta method multiplies each entry of $\mathbf{V}$ by the derivative of the transformation applied to each of its two coordinates. Compliance is the inverse logit of its coordinate, whose derivative is $\lambda_b(1-\lambda_b)$. A parameter of θ_2 searched in logs is the exponential of its coordinate, whose derivative is the parameter itself, and ρ_1 is searched in levels and is left as it is. Writing ϱ_B for that derivative, the estimate $\hat{\theta}_B$ for a log

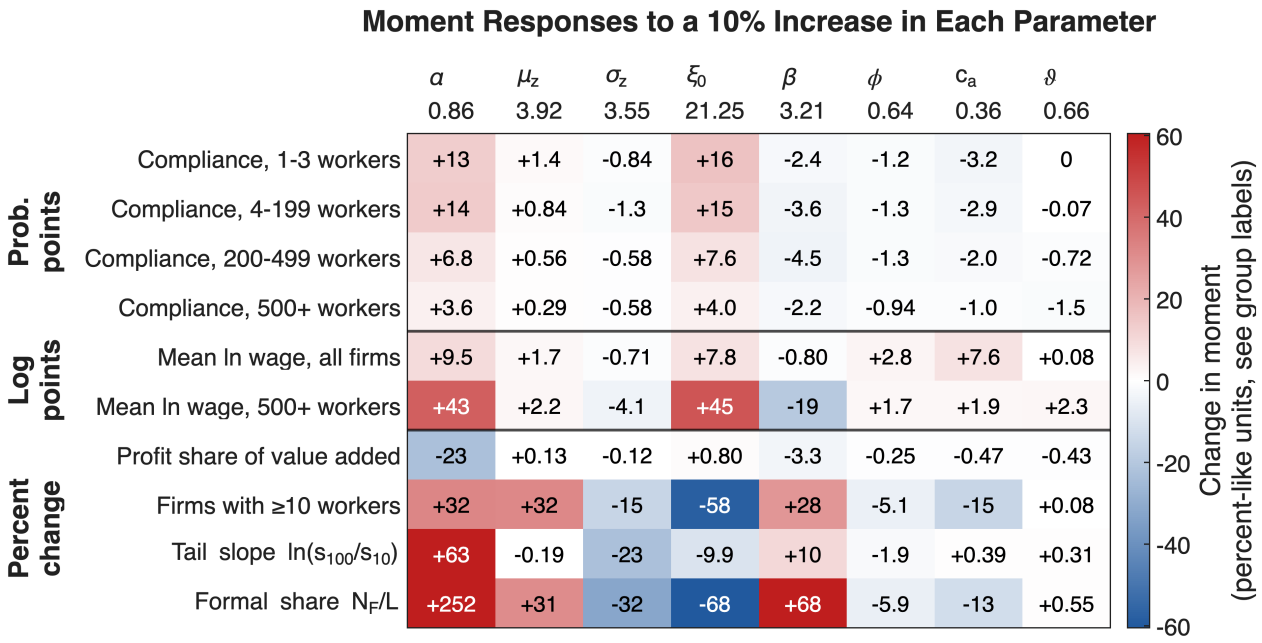


Figure 16: The change in each simulated moment in response to a 10% increase in each parameter, in the moment's own units. The code writes ψ for the opportunity cost scale ϕ.

coordinate and 1 for ρ_1,

$$[\mathbf{V}]_{bb'} \mapsto \hat{\lambda}_b(1-\hat{\lambda}_b)\,[\mathbf{V}]_{bb'}\,\hat{\lambda}_{b'}(1-\hat{\lambda}_{b'}),$$
$$[\mathbf{V}]_{bB} \mapsto \hat{\lambda}_b(1-\hat{\lambda}_b)\,[\mathbf{V}]_{bB}\,\varrho_B, \qquad [\mathbf{V}]_{BB'} \mapsto \varrho_B\,[\mathbf{V}]_{BB'}\,\varrho_{B'}.$$

This converted $\mathbf{V}$ is the one reported in table 4.

C.3 *Block 3: Costs, Production, and the Labor Market*

Local sensitivity. Figure 16 perturbs each parameter by 10% and shows the change in each simulated moment.

Searched coordinates. We search over the transformed parameters of table 17 and bound each by its box in both phases of the search.

Solving at a candidate. Each node solves eq. (7) taking the inclusive value IV as given, and we iterate until the IV implied by the nodes' choices through eq. (2) equals the value they took as given. The fixed point also determines the share of workers who choose the

Parameter	Transform	Box	Parameter	Transform	Box
α	logit	$[0.051,\ 0.989]$	ϕ	log	$[10^{-3},\ 10^{2}]$
μ_z	level	$[1,\ 20]$	c^a	log	$[10^{-4},\ 50]$
σ_z	log	$[0.30,\ 15]$	ξ_0	level	$[0,\ 38]$
β	log	$[0.1,\ 16]$	ϑ	logit	$[0.001,\ 0.999]$

Table 17: Searched coordinates and search box of the eight Block 3 parameters.

informal sector. The nodes represent the J formal firms of Block 1, which sets the size of the economy. When the fixed point has several solutions, we select the one with the lowest stable value of IV by starting every solve from the minimum wage, so the selected equilibrium does not depend on the order in which candidates are evaluated. A candidate whose moments are meaningless, because the formal sector nearly vanishes or fills the economy or the solve fails, receives a penalty in place of the criterion (below).

Compliance moment. Stage A in eq. (31) estimates λ_b from the violations log-likelihood of the firms in bin b. Applied to simulated firms $j \in b$ with compliance λ_j and head count ℓ_j, its probability limit λ_b^* solves

$$\sum_{j \in b} \mathbb{E}_{\lambda_j}\big[\mathbb{1}(d=3)\,\partial_\lambda \ln f(\omega \mid d=3;\lambda_b^*,\ell_j)\big] \;=\; 0, \tag{127}$$

where $\mathbb{E}_{\lambda_j}$ averages over the outcomes of firm j at its own compliance rate. Compliance enters the data-generating process only through the shorted counts of eq. (15). By Fisher's identity, each score is therefore proportional to the difference between the expected number of shorted workers given the recorded set and given the sanction alone. The root of eq. (127) has no closed form. Expanding each term to first order around λ_j, and evaluating the information at the bin's mean compliance, gives the information-weighted mean

$$\lambda_b^I \;=\; \frac{\sum_{j \in b} \iota(\ell_j, \bar{\lambda}_b)\,\lambda_j}{\sum_{j \in b} \iota(\ell_j, \bar{\lambda}_b)}, \qquad \iota(\ell,\lambda) \;:=\; \underbrace{-\mathbb{E}\big[\mathbb{1}(d=3)\,\partial_\lambda^2 \ln f(\omega \mid d=3)\big]}_{\substack{\text{information on compliance}\\ \text{in the violations data}}}, \tag{128}$$

where the sums run over the nodes in bin b, weighted by their probability mass, and $\bar{\lambda}_b$ is their mean compliance. The weight is larger for a firm that is more likely to be sanctioned, because only sanctioned firms enter the violations data. We assign simulated firms to bins by their head count, and round it only inside the information lookup, which maps a firm onto

the integer head count Block 2 can observe. The information ι depends only on the Block 2 estimates, so we tabulate it once by simulation. Simulated panels confirm that λ_b^I is close to the probability limit of the pooled estimator, while mean compliance departs from it in the largest bins.

Wage statistic. We average log wages with each firm weighted equally rather than take the median. The payroll median is taken over worker-months weighted by the inverse of head count, so it reflects wage dispersion within firms that a model with one wage per firm cannot produce. The median is also uninformative in the smallest bins, where roughly half of firms pay the minimum wage and the median sits at it.

The wage profile and two checks on it. Figure 17 plots the profile the wage moments are taken from. We build it separately in each month of 2016 and average the twelve, so the level is the calendar-year one rather than a year-end one and no single month's composition can tilt the shape. Two lines beside it check that the profile measures what the model asks of it. The first reweights each bin's log wage from its own mix of worker cells, education by tenure by age by sex, to the economy-wide mix, so it removes the part of the gradient that worker composition can account for. The index it implies stays between 0.998 and 1.056 across the fifteen bins and is not monotone in size, and the gradient at a thousand workers against the smallest bin moves only from 2.01 to 2.07: on these observables the size-wage gradient is a firm effect rather than sorting. The second keeps only the worker-months that cannot carry more than base pay, dropping the two gratuity months and any month short of its own full complement of days or above its own ordinary hours. It sits on the raw line at every bin, never further than 1.5% away, so the payroll variable is a base monthly salary and the profile is not picking up gratuities or overtime.

Profit share. Census value added is net of intermediate consumption, and census operating surplus subtracts only compensation of employees from it. In the simulated economy the delivered wage bill and the backpay transfer $\mathcal{P}^{\mathrm{bp}}$ of eq. (73) are compensation of employees. Amenity delivery is intermediate consumption, and the compliance cost $\mathsf{b}(w)\mathcal{X}(\lambda)$ and the penalty opportunity cost $\mathcal{X}_1\big(\mathcal{P}^{\mathrm{bp}}+\mathcal{P}^{\mathrm{s}}\big)$ of eq. (13) are output the firm never produces, so value added already excludes all three, while sanctions are paid out of operating surplus.

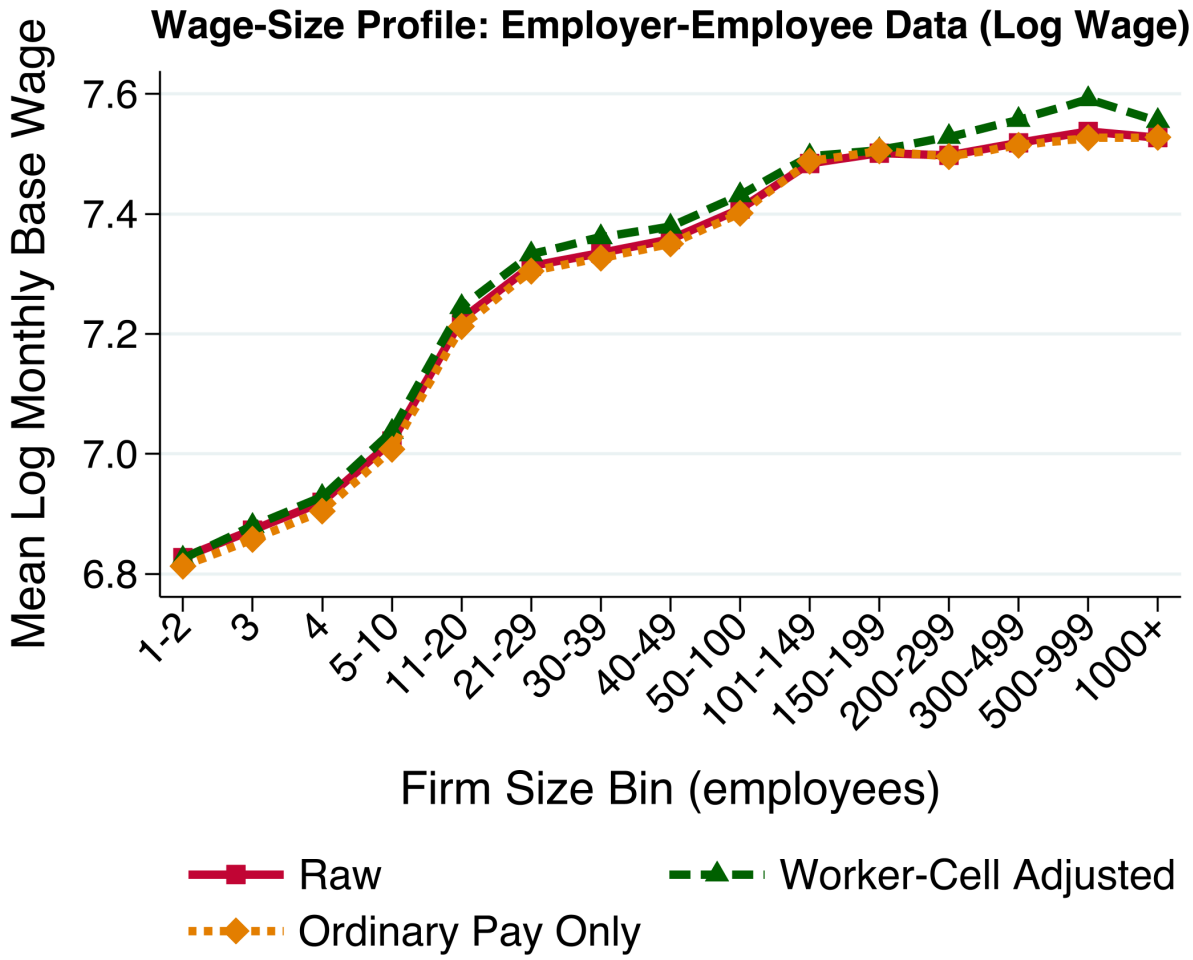


Figure 17: The size-wage profile in the payroll panel, 2016. Each line is the firm-equal mean of the log monthly base wage by size bin, averaged over the twelve monthly profiles.

The simulated share of firm j is therefore

$$\frac{\pi_j + \mathcal{P}_j^{\mathrm{s}}}{\pi_j + \mathcal{P}_j^{\mathrm{s}} + \Lambda(\lambda_j)\, w_j\, \ell_j + \mathcal{P}_j^{\mathrm{bp}}}, \tag{129}$$

and the moment is its median across simulated firms with at least 50 workers, against a census value of 0.45.

The level of this share identifies α. Write $y = z\ell^{\alpha}$ for value added (eq. (6)) and drop the subscript j. Because capital is not modeled separately, the term $(1-\alpha)\, y$ below includes the

return to capital. Euler's theorem gives $\ell\,\partial_\ell y = \alpha y$, so the numerator of eq. (129) splits as

$$\pi + \mathcal{P}^{\mathrm{s}} \;=\; (1-\alpha)\,y \;+\; \underbrace{\ell\left(\partial_\ell y - c(w,\lambda) - \partial_\ell \mathcal{P}\right)}_{\text{markdown}} \;+\; \underbrace{\ell\,\partial_\ell \mathcal{P} - (1+\mathcal{X}_1)\mathcal{P}^{\mathrm{bp}} - \mathcal{X}_1\mathcal{P}^{\mathrm{s}}}_{\substack{\text{marginal penalty net of the}\\ \text{outlay beyond the sanction}}}. \tag{130}$$

The markdown is the gap between a worker's marginal product and the full cost of hiring one. The first-order condition for ℓ in eq. (7) makes it nonnegative whenever one more hire raises the firm's labor supply by less than one worker. Value added is at most output, so when the marginal penalty covers the outlay beyond the sanction, one minus the census share is a lower bound on α.

Criterion. Writing Ψ_e for simulated moment e of table 6, $\hat{\Psi}_e$ for its value in the data and υ_e for its weight, we choose the parameters to minimize

$$10\sum_e \upsilon_e \left(\Psi_e - \hat{\Psi}_e\right)^2, \qquad \upsilon_e \;:=\; \begin{cases} 1 & \text{compliance, log wages, tail slope,} \\ 1/\hat{\Psi}_e^{\,2} & \text{formal employment share, firm share, profit share.} \end{cases} \tag{131}$$

Every moment carries the same weight on its natural scale. A compliance gap is in probability points, a wage gap and the tail slope are in log points, and the three shares enter as proportional gaps, so five points of compliance, five log points of wages and five percent of a share cost the same. The factor of ten only sets the scale of the criterion. Nothing in it is asymmetric: a model above the data costs the same as one below it, and there is no separate penalty on compliance near one.

A candidate can leave a size range with no simulated firm, so that a moment defined on that range has no value. Skipping the moment would reward emptying the range, since a dropped moment costs nothing. We therefore score such a moment at the model's value in the nearest populated size bin, and we count the firms at or above ten and a hundred workers for the tail slope with a normal kernel of width 0.05 in log size, so that the slope stays finite and keeps falling as the largest firms shrink. A candidate whose equilibrium is meaningless receives a penalty in place of the criterion: when the formal sector holds less than two percent or more than 95 percent of employment, the penalty is 10^4, above any value the criterion takes at a solved equilibrium, times one plus the distance past the bound, scaled so that a formal sector of zero or of the whole economy adds one, so that the search has a slope back toward admissible candidates. A failed solve receives the flat 10^4.

Search. The greedy algorithm starts by perturbing each parameter in turn, which shows the moments each parameter moves. Each iteration targets the category of moments that contributes most to the criterion. For that category it computes a Gauss–Newton step in each parameter separately, and it tries the steps with the largest predicted improvement, alone and jointly, at several step lengths. It accepts the trial with the lowest value of the full criterion rather than of the targeted category, which prevents cycling between categories. A category that stops improving is set aside for a few iterations, after smaller steps and larger groups of parameters have failed. When the greedy phase has used its time budget, or every category has been set aside, the local refinement runs bounded nonlinear least squares on the weighted gaps, followed by a bounded minimization of the criterion.

D Welfare Accounting

This appendix derives the welfare accounting of section 6.1 and reports its results at the estimate. Money is in units of the informal wage $w_0 = 1$, the numeraire. We index the informal sector by $j = 0$ and write $\mathcal{J}_0 := \mathcal{J} \cup \{0\}$ for the options a worker chooses among. Table 9 lists the objects.

D.1 *The Welfare Function*

Values and sorting. Workers sort on the value

$$v_j \;=\; W_j + \vartheta B_j + \mathsf{m}\, T_j \quad (j \in \mathcal{J}), \qquad\qquad v_0 \;=\; V_0 \;=\; \xi_0, \tag{132}$$

where T_j is a lump-sum transfer, in money, that the planner pays each worker at firm j. There are no transfers in a market economy ($T_j = 0$, so $v_j = V_j$) and no backpay without enforcement ($B_j = 0$). Delivered utility is $W_j = \beta \ln w_j + \lambda_j D$ (eq. (5)), and the informal worker receives $W_0 = \beta \ln w_0 = 0$ at the numeraire $w_0 = 1$, so the informal value of eq. (16) is its taste alone, $V_0 = \xi_0$. Worker i draws i.i.d. type-I extreme value tastes ε_{ij} with scale one and takes the option with the highest $v_j + \varepsilon_{ij}$. We maintain two assumptions.

1. *Free sorting.* Every worker who chooses an option works there, so employment is the logit share
$$\ell_j \;=\; L\, e^{v_j - IV(v)}, \qquad IV(v) \;:=\; \ln \sum_{j \in \mathcal{J}_0} e^{v_j}, \tag{133}$$

and the labor market clears, $\sum_{j\in\mathcal{J}_0}\ell_j = L$.

2. *Workers receive what they sort on.* A worker at option j receives $v_j + \varepsilon_{ij}$ utils. Under enforcement this means expected backpay is worth ϑB_j to her, the same weight she uses when she sorts.

Free sorting fails at a firm that rations hiring, which happens under perfect enforcement. Appendix D.5 extends the accounting to that case.

The price of a util. We convert utility into money at $1/\mathsf{m}$, where $\mathsf{m} = \beta/\bar{w}$ is the marginal utility of money at the mean wage. The mean $\bar{w} = 1.7$ is money income per worker in the status quo: the formal wages firms pay, the backpay they restore and informal earnings, divided by L. We use the same $\bar{w}$ in every regime, because a welfare change must convert both regimes at one rate.

Welfare. Welfare is profits, plus worker utility converted to money, less the planner's outlay on transfers and what the inspection agency spends net of the sanctions it collects:

$$\mathcal{W} \;=\; \sum_{j\in\mathcal{J}}\pi_j \;+\; \frac{1}{\mathsf{m}}\,U \;-\; \sum_{j\in\mathcal{J}}\ell_j\,T_j \;-\; \sum_{j\in\mathcal{J}}\big(c^{\text{insp}}\,\mathscr{I}_j - \mathcal{P}^{\text{s}}_j\big), \tag{134}$$

where $\pi_j = y_j - c_j\ell_j - \mathcal{P}_j$ is firm j's profit, with output $y_j = z_j\ell_j^{\alpha}$ and the expected penalty $\mathcal{P}_j = (1+\mathcal{X}_1)\big(\mathcal{P}^{\text{bp}}_j + \mathcal{P}^{\text{s}}_j\big)$ of eq. (9); the informal sector earns zero profit (section 3.2.6); U is total worker utility, in utils; and the last sum is the agency's spending on inspections less the sanctions it collects. The agency's cost is net of sanctions because a sanction is a transfer from the firm to the agency: firms pay it inside π_j, and the agency receives it. Its opportunity cost $\mathcal{X}_1\mathcal{P}^{\text{s}}_j$ is not a transfer and stays in the net enforcement cost.

Worker utility. This establishes the split of U into what formal workers sort on and a part that depends on the allocation alone.

Step 1: expected utility. The best of the Gumbel utilities, $\max_j(v_j + \varepsilon_{ij})$, has mean $IV(v) + \gamma_E$, with γ_E Euler's constant. By the second assumption this is each worker's utility, so $U = L\big(IV(v) + \gamma_E\big)$.

Step 2: invert the choice shares. Taking logs of eq. (133), $IV(v) = v_j - \ln(\ell_j/L)$ for every option $j \in \mathcal{J}_0$, the informal sector included.

Step 3: sum over options. Multiplying Step 2 by ℓ_j and summing over $j \in \mathcal{J}_0$, with $\sum_{j\in\mathcal{J}_0} \ell_j = L$,

$$L \cdot IV(v) \;=\; \sum_{j\in\mathcal{J}_0} \ell_j v_j \;-\; \sum_{j\in\mathcal{J}_0} \ell_j \ln \frac{\ell_j}{L}.$$

Step 4: separate the informal sector. Substituting Step 3 into Step 1 and using $v_0 = \xi_0$,

$$U \;=\; \sum_{j\in\mathcal{J}} \ell_j\, v_j \;+\; U^{\text{taste}}(\ell), \qquad U^{\text{taste}}(\ell) \;:=\; \ell_0\, \xi_0 \;-\; \sum_{j\in\mathcal{J}_0} \ell_j \ln \frac{\ell_j}{L} \;+\; L\gamma_E. \tag{135}$$

The first term is what the formal workers sort on. The second, U^{taste}, is the value of the informal option and of the taste draws, and it depends on the allocation ℓ alone. Its sum runs over every option, the informal sector included.

Surplus and the net enforcement cost. Substitute eqs. (132) and (135) into eq. (134).

Step 1: transfers cancel. The formal workers' term splits as

$$\frac{1}{\mathsf{m}} \sum_{j\in\mathcal{J}} \ell_j v_j \;=\; \frac{1}{\mathsf{m}} \sum_{j\in\mathcal{J}} \ell_j W_j \;+\; \frac{1}{\mathsf{m}} \sum_{j\in\mathcal{J}} \ell_j \vartheta B_j \;+\; \sum_{j\in\mathcal{J}} \ell_j T_j,$$

and the last sum cancels the planner's outlay in eq. (134): each worker receives T_j in money and the planner pays it.

Step 2: the sanction transfer cancels. Expanding the penalty and the agency's spending net of sanctions,

$$-\sum_{j\in\mathcal{J}} \Big[\mathcal{P}_j + c^{\text{insp}}\, \mathscr{I}_j - \mathcal{P}_j^{\text{s}}\Big] \;=\; -\sum_{j\in\mathcal{J}} \Big[(1+\mathcal{X}_1)\mathcal{P}_j^{\text{bp}} + \mathcal{X}_1 \mathcal{P}_j^{\text{s}} + c^{\text{insp}}\, \mathscr{I}_j\Big].$$

Step 3: collect. What remains is the surplus $\mathcal{F}$ of eq. (36) less the net enforcement cost NEC of eq. (37), which is eq. (35).

The surplus depends on the postings only through c_j and W_j, and on the allocation through output and U^{taste}. There is no constant. Utility is measured from the informal wage, $\beta \ln w_0 = 0$, so the level of $\mathcal{W}$ carries that origin, but every result below is a difference, which does not.

D.2 *The Planner's Problem*

The planner operates in the no-enforcement economy, so NEC $= 0$ and $B_j = 0$. It takes each firm's no-enforcement posting as given and chooses how many workers each option employs:

$$\mathcal{W}^* \;:=\; \max_{\ell \in \mathbb{A}} \; \mathcal{F}\big(\ell;\, \mathbf{p}^{NE}\big), \qquad \mathbb{A} \;:=\; \Big\{ \ell \in \mathbb{R}^{|\mathcal{J}|+1}_{++} :\; \sum_{j \in \mathcal{J}_0} \ell_j = L \Big\}. \tag{136}$$

The constraint is labor market clearing, so informal employment is $\ell_0 = L - \sum_{j \in \mathcal{J}} \ell_j$. The rest of this subsection shows that taking the postings as given costs the planner nothing, that choosing the allocation is the same as choosing transfers, and that the optimum is the unique allocation at which every marginal surplus is zero.

The mix. Holding the allocation fixed, a change in the mix (w_j, λ_j) that keeps W_j moves $\mathcal{F}$ only through $-c_j \ell_j$. The firm without enforcement already chooses its mix to minimize $c(w, \lambda)$ subject to $\beta \ln w + \lambda D \geq W_j$, the least-cost problem of eq. (46), so a planner facing the same cost function cannot deliver W_j more cheaply. Its first-order conditions set the cost of a util through the wage equal to the cost through compliance, and the envelope theorem makes both the marginal cost of a util,

$$\frac{w\, c_w}{\beta} \;=\; \frac{c_\lambda}{D} \;=\; e'(W). \tag{137}$$

The price of a util to the worker, $1/\mathsf{m}$, does not appear. We do not hold the planner to the minimum wage, because it is a constraint that only enforcement imposes, so the argument also requires that the minimum wage not bind at the cheapest mix. It never binds: across all 997,737 evaluations of $e(W)$ at the utilities delivered in the three regimes, no cheapest mix pays below the minimum wage.

Lemma 4 (Allocations and values)**.** *Let $\mathcal{V}_j(\ell) = \xi_0 + \ln(\ell_j/\ell_0)$ be the inverse labor supply of eq. (39).*

(i) The logit allocation ℓ at values v with $v_0 = \xi_0$ satisfies $v_j = \mathcal{V}_j(\ell)$ for every $j \in \mathcal{J}$.

(ii) Conversely, every $\ell \in \mathbb{A}$ is the logit allocation at the values $v_0 = \xi_0$ and $v_j = \mathcal{V}_j(\ell)$.

Proof. Step 1: values from the allocation. Dividing $\ell_j = Le^{v_j - IV}$ by $\ell_0 = Le^{\xi_0 - IV}$ gives $\ell_j/\ell_0 = e^{v_j - \xi_0}$, so $v_j = \mathcal{V}_j(\ell)$.

Step 2: the allocation from the values. Given $\ell \in \mathbb{A}$, set $v_0 = \xi_0$ and $v_j = \mathcal{V}_j(\ell)$. Then $e^{v_j} = e^{\xi_0}\,\ell_j/\ell_0$ for every option $j \in \mathcal{J}_0$, so $\sum_{j\in\mathcal{J}_0} e^{v_j} = e^{\xi_0} L/\ell_0$ and $L\,e^{v_j - IV} = \ell_j$. □

Transfers implement the allocation. With transfers, workers sort on $v_j = W_j^{NE} + \mathsf{m}\,T_j$. By lemma 4, every choice of transfers produces some $\ell \in \mathbb{A}$, and every $\ell \in \mathbb{A}$ is produced by the transfers $T_j = \big[\mathcal{V}_j(\ell) - W_j^{NE}\big]/\mathsf{m}$. Transfers do not enter $\mathcal{F}$, so a planner that chooses transfers attains the same $\mathcal{W}^*$ as one that chooses the allocation directly. We therefore state the planner's problem eq. (136) in allocations, and return to transfers only to read off the ones that implement the optimum.

Lemma 5 (Marginal surplus)**.** *Let $\ell_0 = L - \sum_{j\in\mathcal{J}} \ell_j$ absorb any change in formal employment. Then $\partial\mathcal{F}(\ell;\,\mathbf{p})/\partial\ell_j$ is $G_j(\ell;\,\mathbf{p})$ of eq. (39), with $MRPL_j = \alpha z_j \ell_j^{\alpha-1}$.*

Proof. Step 1: output and posting. Only firm j's bracket in eq. (36) depends on ℓ_j, and c_j and W_j are fixed with the postings, so its derivative is $MRPL_j - c_j + W_j/\mathsf{m}$.

Step 2: tastes. Three terms of U^{taste} in eq. (135) depend on ℓ_j. With $\partial\ell_0/\partial\ell_j = -1$,

$$\frac{\partial(\ell_0\,\xi_0)}{\partial\ell_j} = -\xi_0, \qquad \frac{\partial}{\partial\ell_j}\Big[-\ell_j \ln\frac{\ell_j}{L}\Big] = -\ln\frac{\ell_j}{L} - 1, \qquad \frac{\partial}{\partial\ell_j}\Big[-\ell_0 \ln\frac{\ell_0}{L}\Big] = \ln\frac{\ell_0}{L} + 1.$$

Their sum is $-\xi_0 - \ln(\ell_j/\ell_0) = -\mathcal{V}_j(\ell)$.

Step 3: assemble. Adding Step 1 to Step 2 divided by m gives eq. (39). □

Proposition 3 (The planner's condition)**.** *The planner's problem eq. (136) has a unique solution ℓ^*, and it is the only allocation in $\mathbb{A}$ at which $G_j\big(\ell^*;\,\mathbf{p}^{NE}\big) = 0$ for every $j \in \mathcal{J}$.*

Proof. Step 1: a maximum exists and is interior. Setting $0\ln 0 = 0$, $\mathcal{F}$ is continuous on the closure of $\mathbb{A}$, which is compact, so it attains a maximum there. Take a point of the closure with $\ell_j = 0$ for some option j, and move a mass ϵ into j from an option with positive employment. The taste term $-\ell_j \ln(\ell_j/L)$ rises from zero to $\epsilon\ln(L/\epsilon)$, output at j does not fall, and every other term of $\mathcal{F}$ changes by $O(\epsilon)$. Since $\epsilon\ln(L/\epsilon)/\epsilon \to \infty$, $\mathcal{F}$ rises for ϵ small. The maximum is therefore in $\mathbb{A}$.

Step 2: at the maximum every marginal surplus is zero. At the maximum ℓ^*, move a small mass $d\ell$ from the informal sector to firm j', holding every other firm fixed. By lemma 5, $\mathcal{F}$ changes by $G_{j'}\,d\ell$ to first order. If $G_{j'} > 0$ this move raises $\mathcal{F}$, and if $G_{j'} < 0$ the reverse

move raises it. Both moves stay in $\mathbb{A}$ because ℓ^* is interior. Hence $G_{j'}\big(\ell^*;\, \mathbf{p}^{NE}\big) = 0$ for every j'.

Step 3: only one allocation has zero marginal surplus. $\mathcal{F}$ is strictly concave in $(\ell_j)_{j\in\mathcal{J}}$ on $\mathbb{A}$. Output is strictly concave since $\alpha < 1$, and the posting terms are linear. Differentiating the result of Step 2 of the proof of lemma 5 once more, $\partial^2 U^{\text{taste}}/\partial\ell_j\partial\ell_{j'} = -\mathbb{1}(j=j')/\ell_j - 1/\ell_0$, so for any $x \neq 0$

$$\sum_{j,j'\in\mathcal{J}} x_j \,\frac{\partial^2 U^{\text{taste}}}{\partial\ell_j\,\partial\ell_{j'}}\, x_{j'} \;=\; -\sum_{j\in\mathcal{J}} \frac{x_j^2}{\ell_j} \;-\; \frac{\big(\sum_{j\in\mathcal{J}} x_j\big)^2}{\ell_0} \;<\; 0.$$

The gradient of $\mathcal{F}$ in $(\ell_j)_{j\in\mathcal{J}}$ is $(G_j)_{j\in\mathcal{J}}$, and a strictly concave function has at most one point with zero gradient. That point is ℓ^* by Step 2. □

The transfers that implement the optimum. By lemma 4, the planner reaches ℓ^* with the transfers $T_j^* = \big[\mathcal{V}_j(\ell^*) - W_j^{NE}\big]/\mathsf{m}$. Substituting this into eq. (39) and using $G_j(\ell^*) = 0$ gives $T_j^* = MRPL_j - c_j^{NE}$, with $MRPL_j$ evaluated at ℓ^*. The planner hires into j until the marginal product equals the full price of the marginal worker. The transfer goes to every worker at j, but to the inframarginal ones it is a pure transfer and costs nothing. Nothing makes the transfers sum to zero: the planner's outlay is 3.5 million w_0, or 1.9% of $\mathcal{W}^*$, which workers receive, so it nets out of welfare, and T_j^* is positive at every firm.

D.3 *The Wedge and the Welfare Change*

The wedge. At an allocation reached by free sorting on values v, lemma 4(i) gives $\mathcal{V}_j(\ell) = v_j$, and eq. (39) becomes

$$G_j \;=\; MRPL_j \;-\; c_j \;-\; \frac{v_j - W_j}{\mathsf{m}}. \tag{138}$$

The gap $v_j - W_j$ is the part of what workers sort on that the posting does not deliver. With the no-enforcement postings, $v_j^{NE} = W_j^{NE}$ gives eq. (41) and $v_j^E = W_j^E + \vartheta B_j$ gives eq. (42). Equation (40) is proposition 3.

Reallocation and the deadweight loss. The reallocation of eq. (43) is the change in surplus as the allocation moves from ℓ^{NE} to ℓ^E with the postings fixed. We compute it by

moving workers in small steps and adding up the steps. By lemma 5, moving $d\ell_j$ workers from the informal sector to firm j changes the surplus by $G_j\, d\ell_j$, so along any path of allocations from ℓ^{NE} to ℓ^E

$$\text{Reallocation} \;=\; \int_{\ell^{NE}}^{\ell^E} \sum_{j\in\mathcal{J}} G_j\big(\ell;\, \mathbf{p}^{NE}\big)\, d\ell_j. \tag{139}$$

The integral is this sum of small steps: at each point on the path, every firm contributes its marginal surplus times the workers it gains there. A worker moved from firm j' to firm j adds $G_j - G_{j'}$. Because G is the gradient of the single function $\mathcal{F}\big(\cdot\,;\, \mathbf{p}^{NE}\big)$, the steps always add up to $\mathcal{F}\big(\ell^E;\, \mathbf{p}^{NE}\big) - \mathcal{F}\big(\ell^{NE};\, \mathbf{p}^{NE}\big)$, whatever path is taken. Define the deadweight loss of an allocation as its distance from the planner's optimum,

$$\text{DWL}(\ell) \;:=\; \mathcal{F}\big(\ell^*;\, \mathbf{p}^{NE}\big) - \mathcal{F}\big(\ell;\, \mathbf{p}^{NE}\big) \;=\; \int_{\ell}^{\ell^*} \sum_{j\in\mathcal{J}} G_j\big(\ell';\, \mathbf{p}^{NE}\big)\, d\ell'_j \;\geq\; 0, \tag{140}$$

which is nonnegative by proposition 3. Along any path from ℓ to ℓ^*, the integrand starts at the wedge of ℓ and ends at zero. Subtracting, the reallocation is $\text{DWL}(\ell^{NE}) - \text{DWL}(\ell^E)$: how much closer enforcement moves the allocation to the planner's. The same terms give each regime's distance from the benchmark,

$$\mathcal{W}^* - \mathcal{W}^E \;=\; \text{DWL}(\ell^E) \;+\; \text{Posting} \;+\; \text{NEC}^E, \qquad \mathcal{W}^* - \mathcal{W}^{NE} \;=\; \text{DWL}(\ell^{NE}). \tag{141}$$

Lemma 6 (Allocations and wedges)**.** *For every $\mathbf{g} \in \mathbb{R}^{|\mathcal{J}|}$ there is exactly one $\ell(\mathbf{g}) \in \mathbb{A}$ with $G\big(\ell(\mathbf{g});\, \mathbf{p}^{NE}\big) = \mathbf{g}$.*

Proof. $G\big(\ell;\, \mathbf{p}^{NE}\big) - \mathbf{g}$ is the gradient of $\mathcal{F}\big(\ell;\, \mathbf{p}^{NE}\big) - \sum_{j\in\mathcal{J}} g_j \ell_j$, which is $\mathcal{F}$ with each posting cost c_j^{NE} replaced by $c_j^{NE} + g_j$. The proof of proposition 3 does not depend on the posting costs, so this function has exactly one point in $\mathbb{A}$ with zero gradient. □

From workers to wedges. Equation (139) adds up marginal surpluses over the workers moved. To write the reallocation in terms of the wedges instead, we pick the path along which the wedges, rather than the workers, move in a straight line, and use lemma 6 to find the allocation at each point. Two objects describe how the allocation responds along that path.

The *Hessian* $\mathbf{H}(\ell)$ is the matrix with entry $\partial G_j / \partial \ell_{j'}$: how firm j's marginal surplus changes when firm j' gains a worker from the informal sector. It is the matrix of second

derivatives of $\mathcal{F}$, since G is its gradient. Differentiating eq. (39), with the postings fixed and $\ell_0 = L - \sum_{j\in\mathcal{J}} \ell_j$,

$$\mathbf{H}(\ell) \;=\; -\big(\operatorname{diag}(\varkappa_j) + \varkappa_0\,\mathbf{1}\mathbf{1}'\big), \qquad \varkappa_j \;:=\; \alpha(1-\alpha) z_j \ell_j^{\alpha-2} + \frac{1}{\mathsf{m}\ell_j}, \qquad \varkappa_0 \;:=\; \frac{1}{\mathsf{m}\ell_0}. \tag{142}$$

The diagonal term $\varkappa_j$ is how fast firm j's own marginal surplus falls as it grows: its marginal product falls, and its marginal worker likes it less. The common term $\varkappa_0$ is how much every firm's marginal surplus falls when the informal sector shrinks, because the marginal informal worker likes informality more. The postings enter $\mathcal{F}$ linearly, so $\mathbf{H}$ does not depend on them. Its inverse translates wedges into employment: changing the allocation by $d\ell$ changes the wedges by $\mathbf{H}\,d\ell$, so a change dG in the wedges needs $d\ell = \mathbf{H}^{-1}dG = -\mathbf{Q}\,dG$, where $\mathbf{Q} := -\mathbf{H}^{-1}$.

Proposition 4 (Reallocation and the wedge change)**.** *Let $\Delta G := G^E - G^{NE}$. Along the wedges $G(s) := G^{NE} + s\,\Delta G$, $s \in [0,1]$, let $\ell(s) := \ell\big(G(s)\big)$ and $\mathbf{Q}(s) := -\mathbf{H}\big(\ell(s)\big)^{-1}$. Then the reallocation is eq. (44).*

Proof. Step 1: endpoints. $G\big(\ell^{NE};\,\mathbf{p}^{NE}\big) = G^{NE}$ and $G\big(\ell^{E};\,\mathbf{p}^{NE}\big) = G^{E}$ by eqs. (41) and (42), so lemma 6 gives $\ell(0) = \ell^{NE}$ and $\ell(1) = \ell^{E}$. The path therefore runs between the two allocations of the reallocation.

Step 2: how the allocation moves. Raising s by ds raises the wedges by $\Delta G\,ds$. The allocation must move by the $d\ell$ that produces this change, $\mathbf{H}\,d\ell = \Delta G\,ds$, so its speed along the path is

$$\dot{\ell} \;:=\; \frac{d\ell}{ds} \;=\; -\mathbf{Q}(s)\,\Delta G.$$

$\mathbf{H}$ is negative definite by Step 3 of the proof of proposition 3, so it can be inverted, and it is symmetric because it is a matrix of second derivatives, so $\mathbf{Q}(s)$ is symmetric too.

Step 3: the surplus along the path. By lemma 5, raising s by ds changes the surplus by the marginal surplus of each firm times the workers it gains,

$$\frac{d\mathcal{F}}{ds} \;=\; \sum_{j\in\mathcal{J}} G_j(s)\,\dot{\ell}_j \;=\; -\sum_{j\in\mathcal{J}} G_j(s)\big[\mathbf{Q}(s)\,\Delta G\big]_j \;=\; -\sum_{j\in\mathcal{J}} \Delta G_j\,\big[\mathbf{Q}(s)\,G(s)\big]_j,$$

where the last equality uses the symmetry of $\mathbf{Q}(s)$.

Step 4: add up the steps. Integrating over $s \in [0,1]$, an integral over a single number,

and using Step 1,

$$\mathcal{F}\big(\ell^E;\,\mathbf{p}^{NE}\big)-\mathcal{F}\big(\ell^{NE};\,\mathbf{p}^{NE}\big)\;=\;-\sum_{j\in\mathcal{J}}\Delta G_j\int_0^1\big[\mathbf{Q}(s)\,G(s)\big]_j\,ds,$$

which is eq. (44). □

The gap depends on the path, and so on the whole change ΔG, but at a given firm it is the same for every part of that change. Any split $\Delta G_j=\sum_n\Delta G_j^{(n)}$ therefore splits the reallocation into the terms $-\sum_{j\in\mathcal{J}}\Delta G_j^{(n)}\,\bar{\ell}_j^{\text{gap}}$, one for each part.

The employment gap in closed form. The vector $x:=\mathbf{Q}\,G$ solves $\big(\text{diag}(\varkappa_j)+\varkappa_0\,\mathbf{1}\mathbf{1}'\big)x=G$. Row j reads $\varkappa_j x_j+\varkappa_0 X=G_j$, where $X:=\sum_{j'\in\mathcal{J}}x_{j'}$ is the total change in formal employment, so $x_j=(G_j-\varkappa_0 X)/\varkappa_j$. Summing this over j and solving for X,

$$\big[\mathbf{Q}\,G\big]_j\;=\;\frac{G_j-\varphi\,\hat{G}}{\varkappa_j},\qquad\hat{G}\;:=\;\frac{\sum_{j'}G_{j'}/\varkappa_{j'}}{\sum_{j'}1/\varkappa_{j'}},\qquad\varphi\;:=\;\frac{\varkappa_0\sum_{j'}1/\varkappa_{j'}}{1+\varkappa_0\sum_{j'}1/\varkappa_{j'}}\in(0,1),$$

with $\varkappa_j$ and $\varkappa_0$ at $\ell(s)$. Here $\hat{G}$ averages the wedges with weights $1/\varkappa_{j'}$, how easily each firm expands, and $\varphi\hat{G}=\varkappa_0 X$ is how much every firm's marginal surplus falls as the informal sector shrinks to supply the added workers. To first order, the planner adds workers at firm j when its wedge exceeds $\varphi\hat{G}$, a fraction of the average wedge.

Posting. At the fixed allocation ℓ^E, only firm j's bracket in eq. (36) changes with its posting, by $-\ell_j^E(c_j^E-c_j^{NE})+\ell_j^E(W_j^E-W_j^{NE})/\mathsf{m}=-\ell_j^E\,\Delta c_j^{\text{net}}$. Adding and subtracting $e(W_j^E)$ and using $c_j^{NE}=e(W_j^{NE})$ gives eq. (47). The posting-transfer integrand compares $e'(W)=w\,c_w/\beta$ of eq. (137) with $1/\mathsf{m}$. No regime equates the two, and $e'(W)$ rises with the wage, so the posting transfer counts a cut in what high-wage firms deliver as a gain.

D.4 *What Moves the Wedge*

The wedge of each regime follows from the firm's first-order conditions of appendix A.10. At a firm that hires on its labor supply curve and pays above the minimum wage, the wage

condition eq. (86) with $\kappa = 0$ and the labor condition eq. (91) give

$$\eta\mu_j \;=\; \frac{w_j}{\beta}\Big(c_w + \frac{\mathcal{P}_w}{\ell_j}\Big), \qquad\qquad MRPL_j - c_j \;=\; \mu_j \;+\; \mathcal{P}_\ell \;-\; \eta\mu_j\,\vartheta\,\ell_j B_\ell, \tag{143}$$

where μ_j is the recruiting cost and $\eta\mu_j$ the cost of a util. Without enforcement, $\mathcal{P} = B = 0$, so the recruiting cost $\mu_j^{NE} = w_j c_w/(\beta\eta)$ is the markdown and $G_j^{NE} = MRPL_j - c_j^{NE} = \mu_j^{NE}$.

Proposition 5 (What moves the wedge)**.** *Under free sorting, at a firm that hires on its labor supply curve and pays above the minimum wage in both regimes, the wedge change is eq.* (45)*.*

Proof. Step 1: separate the posting change. Adding and subtracting c_j^E and W_j^E/m in eq. (42) gives

$$G_j^E \;=\; \underbrace{MRPL_j - c_j^E}_{\text{set by the firm's hiring condition}} \;+\; \Delta c_j^{\text{net}} \;-\; \frac{\vartheta B_j}{\mathsf{m}}, \tag{144}$$

$$\Delta c_j^{\text{net}} \;:=\; \big(c_j^E - c_j^{NE}\big) - \frac{W_j^E - W_j^{NE}}{\mathsf{m}}, \tag{145}$$

Step 2: the hiring condition. By eq. (143), $MRPL_j - c_j^E = \mu_j^E + \mathcal{P}_\ell - \eta\mu_j^E\vartheta\ell_j B_\ell$. The wage condition splits μ_j^E into the markdown at the enforced posting, $w_j c_w/(\beta\eta)$, and the wage penalty $w_j\mathcal{P}_w/(\beta\eta\ell_j)$, and eq. (98) splits $\mathcal{P}_\ell = (1+\mathcal{X}_1)\mathcal{P}_\ell^{\text{bp}} + \mathcal{P}_\ell^{\text{s}}$.

Step 3: subtract. Subtracting $G_j^{NE} = \mu_j^{NE}$ from Step 2 and grouping the terms by channel gives eq. (45). □

Margin and level. Each enforcement channel can enter the welfare change in two places. Its *margin*, what the firm pays when it hires one more worker, moves the allocation: it is a line of eq. (45), and proposition 4 turns it into welfare by weighting it with the firm's employment gap. Its *level*, the total paid or used up at ℓ^E, enters the posting term or the net enforcement cost. The two are different objects, so nothing is counted twice. A tax whose revenue is burned works the same way: the tax lowers the quantity traded, which is the deadweight triangle, and the burned revenue at the new quantity is a separate loss. Table 18 lists where each line of eq. (45) has its level.

Channel	**Margin: line of** (45)	**Level**
Markdown	$w_j c_w/(\beta\eta) - \mu_j^{NE}$	None of its own: it changes because the posting changes, whose level is in the posting term
Wage penalty	$w_j \mathcal{P}_w/(\beta\eta\ell_j)$	None of its own: it is the backpay penalty's margin in the wage, whose level is in the net enforcement cost
Backpay	$\mathcal{P}_\ell^{\text{bp}} - \vartheta B_j/\mathsf{m} - \eta\mu_j\vartheta\ell_j B_\ell$	Net enforcement cost: $\mathcal{P}_j^{\text{bp}} - \ell_j\vartheta B_j/\mathsf{m}$
Penalty OC	$\mathcal{X}_1\big(\mathcal{P}_\ell^{\text{bp}} + \mathcal{P}_\ell^{\text{s}}\big)$	Net enforcement cost: $\mathcal{X}_1\big(\mathcal{P}_j^{\text{bp}} + \mathcal{P}_j^{\text{s}}\big)$
Sanction	$\mathcal{P}_\ell^{\text{s}}$	None: firms pay it and the agency receives it
Agency spending	None: the firm does not pay it	Net enforcement cost: $c^{\text{insp}}\,\mathscr{I}_j$
Mix	Mix_j	Posting term: $\ell_j^E\,\text{Mix}_j$
Posting transfer	Transfer_j	Posting term: $\ell_j^E\,\text{Transfer}_j$

Table 18: Where each channel enters the welfare change.

Corollary 1 (The welfare change by channel)**.** *Under the assumptions of proposition 5 at every firm,*

$$\begin{aligned}\Delta\mathcal{W} \;=\; -\sum_{j\in\mathcal{J}} \bar{\ell}_j^{\text{gap}}\Big[&\underbrace{\frac{w_j\,c_w}{\beta\eta} - \mu_j^{NE}}_{\textit{markdown}} + \underbrace{\frac{w_j\,\mathcal{P}_w}{\beta\eta\ell_j}}_{\textit{wage penalty}} + \underbrace{\mathcal{P}_\ell^{\text{bp}} - \frac{\vartheta B_j}{\mathsf{m}} - \eta\mu_j\,\vartheta\,\ell_j B_\ell}_{\textit{backpay}} \\ &+ \underbrace{\mathcal{X}_1\big(\mathcal{P}_\ell^{\text{bp}} + \mathcal{P}_\ell^{\text{s}}\big)}_{\textit{penalty OC}} + \underbrace{\mathcal{P}_\ell^{\text{s}}}_{\textit{sanction}} + \underbrace{\text{Mix}_j}_{\textit{mix}} + \underbrace{\text{Transfer}_j}_{\textit{posting transfer}}\Big] \\ &- \sum_{j\in\mathcal{J}} \ell_j^E\,\text{Mix}_j \;-\; \sum_{j\in\mathcal{J}} \ell_j^E\,\text{Transfer}_j \\ &- \sum_{j\in\mathcal{J}} \Big[\mathcal{X}_1\big(\mathcal{P}_j^{\text{bp}} + \mathcal{P}_j^{\text{s}}\big) + \Big(\mathcal{P}_j^{\text{bp}} - \frac{\ell_j\,\vartheta B_j}{\mathsf{m}}\Big) + c^{\text{insp}}\,\mathscr{I}_j\Big].\end{aligned} \tag{146}$$

The first two lines are the reallocation, the third is the posting term, and the fourth is the net enforcement cost. Adding the mix and posting-transfer lines of the reallocation to the posting term, which weight the same Mix_j *and* Transfer_j *by* $\bar{\ell}_j^{\text{gap}}$ *and by* ℓ_j^E*, gives the grouped decomposition eq.* (48)*.*

Proof. Proposition 4 with $G_j^E - G_j^{NE}$ from eq. (45) gives the reallocation, eq. (47) gives the posting term, and eq. (37) at ℓ^E gives the net enforcement cost. They add up to $\Delta\mathcal{W}$ by eq. (43). □

D.5 *Perfect Enforcement and Rationing*

Perfect enforcement. Complaining is free, so every shorted worker complains and every shortfall surfaces: the escape probability is $\text{esc}_k = 0$, and the expected complaint cost is $H_k = 0$ because the mean complaint cost $1/\rho(\ell)$ goes to zero (eq. (72)). A firm is still uninspected when none of its $\ell|\mathcal{K}|$ worker-regulation pairs is shorted, so its inspection probability is $\mathscr{I}_j = 1 - \lambda_j^{\ell_j|\mathcal{K}|}$ (table 10), and it pays the sanction of eq. (64) on the regulations it fails to resolve in the window. The backpay value is $B_j^{PE} = (1-\lambda_j)D$, so $B_\ell = 0$, and the backpay transfer of eq. (73) is linear in employment, $\mathcal{P}_j^{\text{bp}} = \ell_j\big[w_j\big(\Lambda(1) - \Lambda(\lambda_j)\big) + c^a K^a(1-\lambda_j)\big]$.

Rationing. At the estimate 319,647 of the 332,579 firms ration hiring under perfect enforcement: they hire fewer workers than choose them, $\ell_j < N_j := Le^{V_j - IV(V)}$, where $V_j = W_j + \vartheta B_j$ is the value workers sort on. This breaks both assumptions of appendix D.1. As in section 6.1.4, applicants a firm turns away take the informal option, so informal employment is $\ell_0 = \ell_0^{\text{free}} + \sum_{j\in\mathcal{J}}(N_j - \ell_j)$, where $\ell_0^{\text{free}} := Le^{\xi_0 - IV(V)}$ is the informal employment that free sorting on V would produce. We replace the two assumptions with one:

3. *Taste-efficient rationing.* Hired workers at firm j receive $V_j + \varepsilon_{ij}$, and among all assignments of workers that produce the realized employment ℓ, the one that occurs maximizes total taste utility.

By lemma 4(ii), ℓ is the free-sorting allocation at the effective values $v_j^{\text{eff}} := \mathcal{V}_j(\ell)$, and free sorting on v^{eff} is the assignment that maximizes total taste utility given ℓ.[27] The assumption therefore says that tastes are those of free sorting at v^{eff}. Any other rule, such as random rationing, gives lower taste utility at the same allocation, and the difference would be a loss outside the three terms of eq. (43). Under random rationing, the tastes would be a further 629 thousand w_0 lower.

Lemma 7 (Welfare under rationing)**.** *Under taste-efficient rationing,* $U = \sum_{j\in\mathcal{J}} \ell_j V_j + U^{\text{taste}}(\ell)$. *The welfare function eqs.* (35) *to* (37) *and the decomposition eq.* (43) *hold unchanged.*

[27] Maximizing $\sum_i \varepsilon_{i,j(i)}$ subject to the employment counts ℓ_j, with a multiplier $-v_j$ on each count, gives the rule that each worker takes the option with the highest $v_j + \varepsilon_{ij}$. The multipliers that produce ℓ are v^{eff}, up to a constant that $v_0^{\text{eff}} = \xi_0$ fixes.

Proof. Step 1: tastes. By the assumption, taste draws and the informal option are those of free sorting at v^{eff}, so by Steps 1–4 of the derivation of eq. (135), workers would receive $L\big(IV(v^{\text{eff}})+\gamma_E\big)=\sum_j \ell_j v_j^{\text{eff}}+U^{\text{taste}}(\ell)$ if each hired worker at j received v_j^{eff}.

Step 2: what hired workers receive. Each hired worker at j receives V_j instead of v_j^{eff}, so $U=\sum_j \ell_j v_j^{\text{eff}}+U^{\text{taste}}(\ell)+\sum_j \ell_j\big(V_j-v_j^{\text{eff}}\big)=\sum_j \ell_j V_j+U^{\text{taste}}(\ell)$.

Step 3: welfare. This is eq. (135) with $v_j=V_j=W_j+\vartheta B_j$, the value used there with $T_j=0$, so the derivation of eq. (35) goes through. The decomposition uses only eq. (35). □

Rationing therefore enters welfare only through the allocation it produces, which the reallocation values like any other move. No separate rationing loss is needed: the loss to the applicants a firm turns away is part of the move to ℓ^{PE}. What changes is the wedge, because at a rationed allocation lemma 4(i) no longer gives $\mathcal{V}_j(\ell)=V_j$.

Proposition 6 (The wedge under rationing)**.** *For every firm, the gap between what workers sort on and the effective value is eq.* (49)*. The wedge at the perfect-enforcement allocation, $G_j^{PE}:=G_j\big(\ell^{PE};\mathbf{p}^{NE}\big)$, satisfies:*

(i) at a firm that hires on its labor supply curve, eq. (45) *with G_j^{PE} in place of G_j^E, plus $\frac{1}{\mathsf{m}}\ln(\ell_0/\ell_0^{\text{free}})$;*

(ii) at a firm that rations,

$$\begin{aligned} G_j^{PE}-G_j^{NE} \;=\; -\,\mu_j^{NE} \;+\; \underbrace{\mathcal{P}_\ell^{\text{bp}}-\frac{\vartheta B_j}{\mathsf{m}}}_{\textit{backpay}} \;+\; \underbrace{\mathcal{X}_1\big(\mathcal{P}_\ell^{\text{bp}}+\mathcal{P}_\ell^{\text{s}}\big)}_{\textit{penalty OC}} \;+\; \underbrace{\mathcal{P}_\ell^{\text{s}}}_{\textit{sanction}} \\ +\;\underbrace{\text{Mix}_j}_{\textit{mix}} \;+\; \underbrace{\text{Transfer}_j}_{\textit{posting transfer}} \;+\; \underbrace{\frac{1}{\mathsf{m}}\Big[\ln\frac{\ell_0}{\ell_0^{\text{free}}}+\ln\frac{N_j}{\ell_j}\Big]}_{\textit{rationing}}. \end{aligned} \tag{147}$$

Proof. Step 1: the gap. From $N_j=Le^{V_j-IV(V)}$ and $\ell_0^{\text{free}}=Le^{\xi_0-IV(V)}$, $V_j=\xi_0+\ln(N_j/\ell_0^{\text{free}})$. Subtracting $v_j^{\text{eff}}=\xi_0+\ln(\ell_j/\ell_0)$ gives eq. (49). Both terms are nonnegative, since $\ell_j\le N_j$ and turned-away applicants raise ℓ_0 above ℓ_0^{free}.

Step 2: the wedge. By lemma 5 with $\mathcal{V}_j(\ell^{PE})=v_j^{\text{eff}}=V_j-(V_j-v_j^{\text{eff}})$,

$$G_j^{PE} \;=\; MRPL_j-c_j^{NE}-\frac{V_j-W_j^{NE}}{\mathsf{m}}+\frac{V_j-v_j^{\text{eff}}}{\mathsf{m}}.$$

	Status quo		Perfect enforcement	
	w_0, 000s	wage, log pts	w_0, 000s	wage, log pts
Workers				
Wage	74.0	0.30	−1,287.6	−5.19
Regulations	−210.9	−0.85	4,303.5	17.35
Backpay	148.9	0.60	3,209.0	12.94
Rationing	–	–	−6,155.9	−24.82
Total	12.0	0.05	69.0	0.28
Firms				
Output	204.6	0.82	−14,159.7	−57.08
Costs	134.0	0.54	11,043.0	44.52
Penalties	−282.6	−1.14	−996.0	−4.02
Total	56.0	0.23	−4,112.8	−16.58
The public				
Agency, net of sanction revenue	−1.8	−0.01	−44.0	−0.18
Welfare change	66.2	0.27	−4,087.8	−16.48

Table 19: Incidence of the welfare change from no enforcement on workers, firm owners and the public, in the units of table 8.

The first three terms are eq. (42) at the perfect-enforcement allocation and postings. At a firm on its labor supply curve, the proof of proposition 5 applies to them unchanged, which gives (i).

Step 3: a firm that rations. Its labor supply constraint is slack, so its recruiting cost is zero, $\mu_j = 0$: more workers already sort to it than it hires. The labor condition then reads $MRPL_j - c_j^{PE} = \mathcal{P}_\ell$ (appendix A.10). Substituting into Step 1 of the proof of proposition 5, adding the gap from Step 2, and subtracting $G_j^{NE} = \mu_j^{NE}$ gives eq. (147). □

At a firm that rations, the markdown, the wage penalty and the backpay slope all drop out with μ_j, so the firm's recruiting terms are replaced by $-\mu_j^{NE}$. Under perfect enforcement $B_\ell = 0$, so the backpay slope is zero at every firm. Proposition 4 and corollary 1 apply unchanged, with G^{PE} in place of G^E and, at a firm that rations, the lines of eq. (147) in place of those of eq. (45).

D.6 *Incidence*

Under free sorting, worker utility $U = L\big(IV(v) + \gamma_E\big)$ changes by $\Delta U = L\,\Delta IV$. Since $\partial IV/\partial v_j = \ell_j/L$, integrating along the straight path in v from v^{NE} to v^{SQ} gives $\Delta U = \sum_{j\in\mathcal{J}} \bar{\ell}_j\,\Delta V_j$, where $\bar{\ell}_j$ is the employment of firm j averaged along the path. The change in value splits as $\Delta V_j = \beta\,\Delta \ln w_j + D\,\Delta\lambda_j + \vartheta\,\Delta B_j$ into wage, regulations and backpay, and firm j's profit change net of penalties, $\Delta\pi_j$, splits into output, costs and penalties. The agency's net cost falls on the public purse. Then, exactly,

$$\Delta\mathcal{W} \;=\; \sum_{j\in\mathcal{J}} \Delta\pi_j + \frac{\Delta U}{\mathsf{m}} - \sum_{j\in\mathcal{J}} \Delta\big(c^{\text{insp}}\,\mathscr{I}_j - \mathcal{P}^{\text{s}}_j\big), \tag{148}$$

which is the change in eq. (134) with no transfers. Under perfect enforcement, ΔU also subtracts the rationing loss,

$$\text{rationing loss} \;:=\; L\big(IV(V) + \gamma_E\big) - \Big[\sum_{j\in\mathcal{J}} \ell^{PE}_j V_j + U^{\text{taste}}(\ell^{PE})\Big] \;\geq\; 0.$$

It is the gap between the utility that free sorting on V would give and realized utility under taste-efficient rationing (lemma 7). It is nonnegative because free sorting maximizes $\sum_{j\in\mathcal{J}} \ell_j V_j + U^{\text{taste}}(\ell)$ over allocations. We need it only because the incidence formula starts from free sorting; it is not a separate term of the welfare decomposition.

D.7 *The Inspection Agency*

The model has no cost of running inspections, but SUNAFIL's accounts do. Its 2016 budget was S/ 85 million, and it collected S/ 8.7 million in sanctions. At twelve months of the informal wage, the unit in which sanctions enter the model (appendix C), the budget is 8,358 w_0. The model's status-quo sanctions are 2.2 times the 2016 collection, the same order of magnitude, with the complaint-driven inspections of formal firms that the model represents on the high side of the agency's receipts.

SUNAFIL's budget also pays for planned inspections, orientation visits and firms outside the model, so we attribute to the model only a share of it and spread that share over the model's own inspections:

$$c^{\text{insp}} \;:=\; \frac{\text{share} \times \text{budget}_{2016}}{\sum_{j\in\mathcal{J}} \mathscr{I}^{SQ}_j}, \tag{149}$$

	$\mathcal{W}$	$\mathcal{W}^* - \mathcal{W}$	DWL	Mix	Transfer	NEC
No enforcement	181.28	2.831	2.831	0.000	0.000	0.000
Status quo	181.35	2.765	2.800	0.120	−0.290	0.135
Perfect enforcement	177.19	6.918	6.022	0.117	0.288	0.492

Table 20: Welfare and its distance from the benchmark $\mathcal{W}^*$, in millions of w_0. Each row is the identity in eq. (141).

where $\sum_j \mathscr{I}_j^{SQ} = 5{,}465$ is the expected number of inspections a year in the status quo. The agency then spends $c^{\text{insp}} \sum_j \mathscr{I}_j$ in each regime: nothing without enforcement, where no firm is inspected, and $c^{\text{insp}} \sum_j \mathscr{I}_j^{PE}$ under perfect enforcement, where a firm is inspected unless none of its worker-regulation pairs is shorted (table 10) and $\sum_j \mathscr{I}_j^{PE} = 139{,}220$ expected inspections a year. We take the share as the inspectors' share of SUNAFIL's staff times the complaint-driven share of its inspection orders, $(480/788) \times (39{,}728/55{,}173) = 0.44$, which gives $c^{\text{insp}} = 0.67\ w_0$ per inspection. The agency's net cost, its spending less the sanctions it collects, is then 1.8 thousand w_0 in the status quo and 44 thousand under perfect enforcement. Attributing the whole budget to the model raises these to 6.5 and 164 thousand.

D.8 *Results at the Estimate*

The benchmark. Welfare in the no-enforcement market is 181.3 million w_0 and the benchmark is 184.1 million, so the market is 2.83 million below it, or 1.6% of $\mathcal{W}^{NE}$. Table 20 reports the distance of each regime from the benchmark and its three sources. Valued at the no-enforcement postings, the employment-weighted wedge is 1.36 w_0 per marginal worker without enforcement, 1.35 in the status quo and 2.80 under perfect enforcement.

Compliance and rationing. Employment-weighted compliance is 0.84 without enforcement, 0.83 in the status quo and 0.92 under perfect enforcement. Under perfect enforcement 9.8 million workers sort to formal firms, 1.9 million are hired and 7.9 million are turned away into the informal sector. The gain to workers from restored regulations and backpay under free sorting, 6.22 million w_0, is nearly offset by the rationing loss of 6.16 million.

By firm size. Figure 18 shows posting and the net cost of enforcement by firm size, fig. 19 shows the employment gap and the reallocation line by line, and fig. 20 shows the incidence on workers and firms. The employment gap is positive at every firm in both regimes, so along

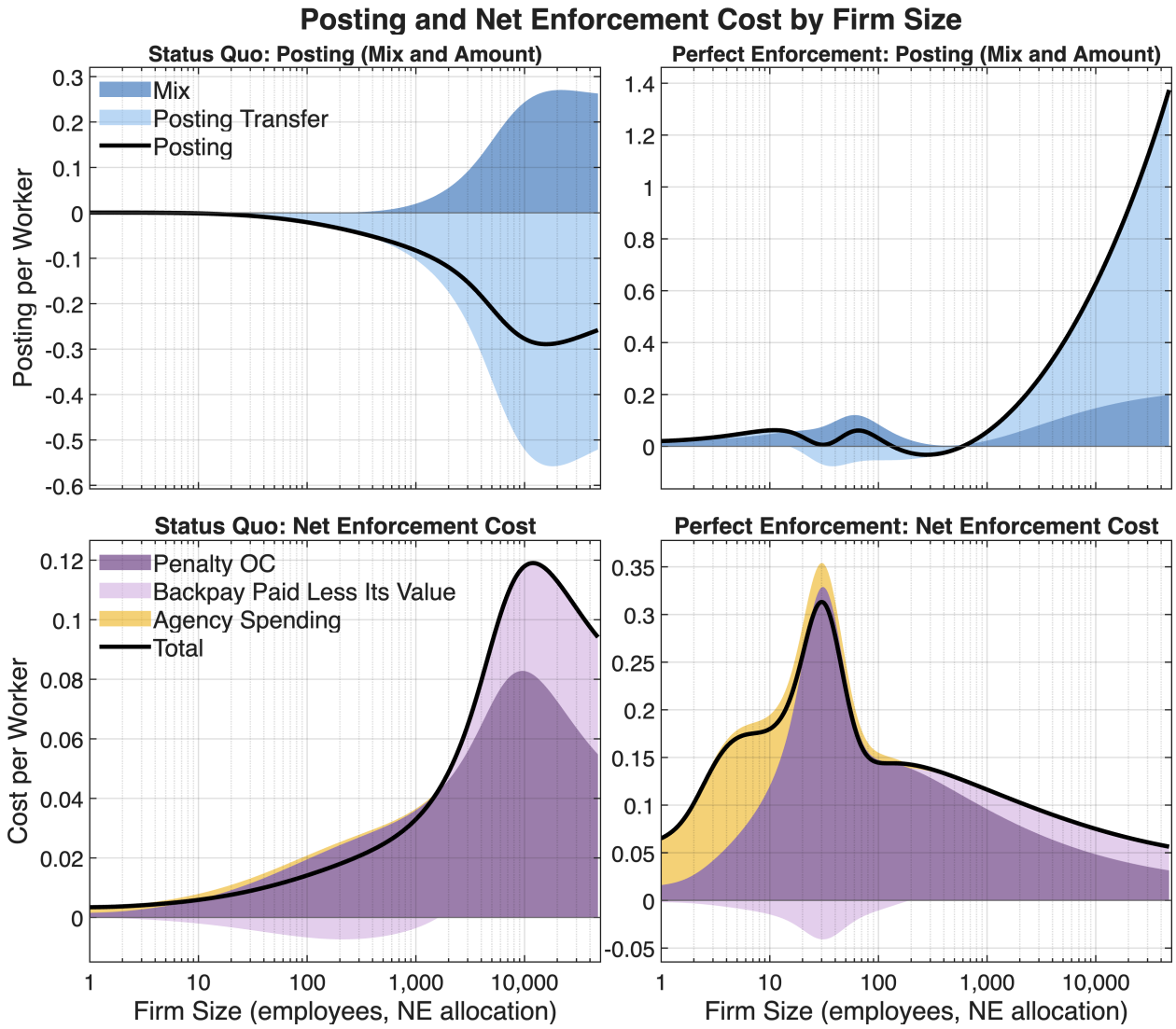


Figure 18: Posting and the net cost of enforcement against firm size in the no-enforcement allocation, per no-enforcement worker. Top: the mix and the posting transfer weighted by $\bar{\ell}_j^{\text{gap}} + \ell_j^E$, their effect through hiring and on the workers employed (eq. (48)). Bottom: the three parts of the net enforcement cost. Both rows show costs, which enter the welfare change with a minus. "Penalty OC" is the penalty opportunity cost, the cost of paying backpay and sanctions at the inspection date.

either path the planner would expand every firm, and relative to its size it would expand the largest firms most.

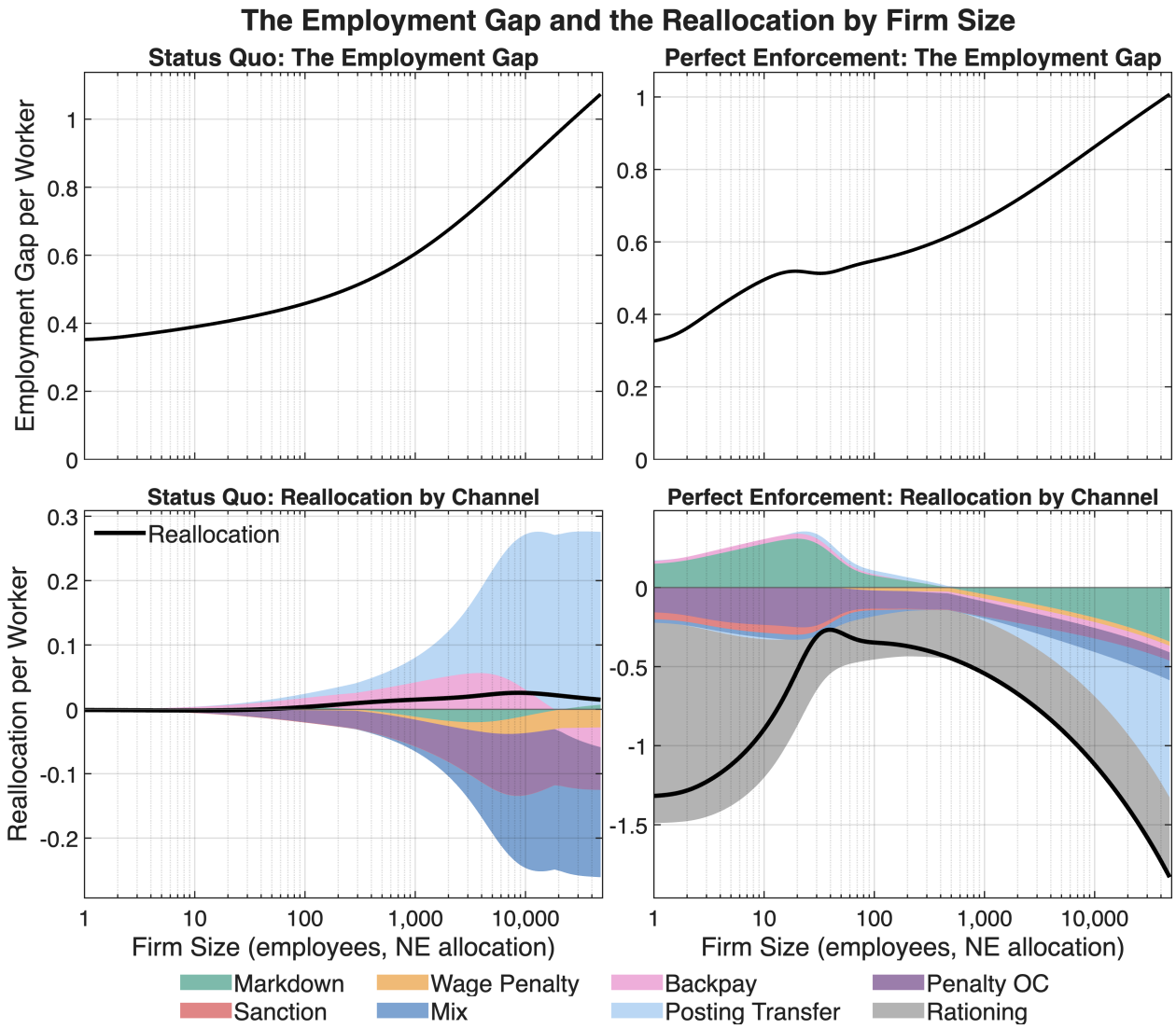


Figure 19: The employment gap and the reallocation against firm size in the no-enforcement allocation. Top: the employment gap $\bar{\ell}_j^{\text{gap}}$ of eq. (44) per no-enforcement worker; the planner would expand a firm above zero and shrink one below. Bottom: the reallocation per no-enforcement worker, $-\bar{\ell}_j^{\text{gap}}$ times each line of eqs. (45) and (147), with their sum in black. Mix and Posting Transfer are posting's effect through hiring.

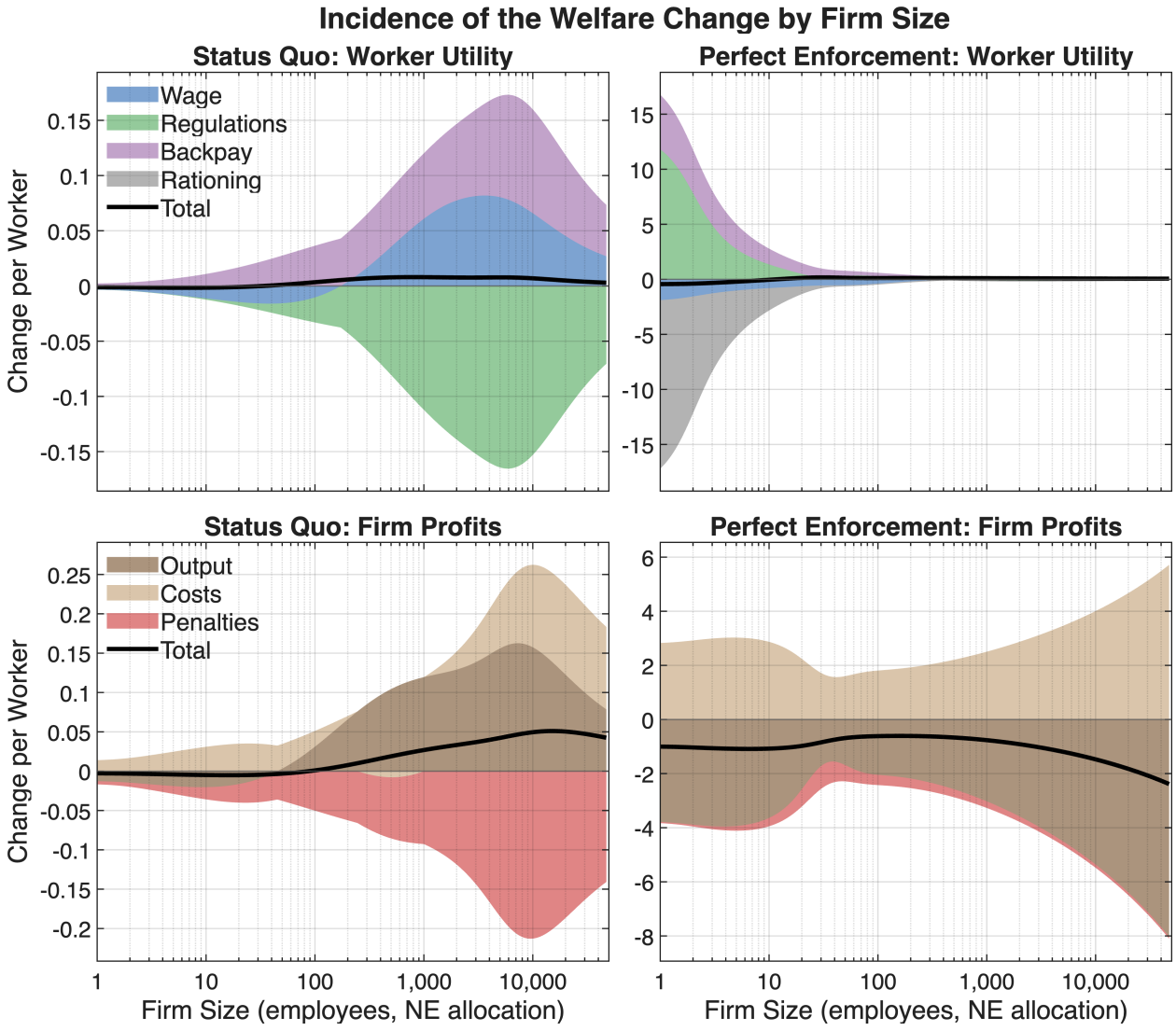


Figure 20: The change in worker utility and firm profits against firm size in the no-enforcement allocation, per no-enforcement worker. The two rows add to the welfare change at each size. "Regulations" is regulated compensation and amenities.