Detection or Penalties? The Non-Equivalence of Enforcement Instruments

Evidence from Minimum-Wage Compliance in Colombia

Mohamed Coulibaly* Feraud Tchuisseu Seuyong†

August 3, 2026

Abstract

Standard enforcement theory treats detection and penalties as substitutes that matter only through the expected fine. We show they are not interchangeable when violations are durable states: detection clears violations, while penalties only price them. We study Colombia's 2013 Law 1610, which sharply raised penalties for labor violations. In departments more exposed to pre-reform minimum-wage underpayment, event-study estimates show that the depth of violations fell after the reform while their incidence did not. We estimate a search-and-matching model with imperfect enforcement and compare the two instruments: matched to the same reduction in aggregate underpayment, penalties reduce both the number and the depth of violations without job losses, whereas detection cuts incidence by 11 percentage points, more than twice the reduction under penalties, while leaving a negatively selected pool of deeper violations and raising unemployment. Equal expected fines need not produce equal enforcement outcomes.

Keywords: Minimum wage; non-compliance; enforcement; inspections; penalties; Colombia.
JEL Classification: J38, J46, J31.

*HEC Montréal. Email: mohamed.coulibaly@hec.ca.
†University of Montréal. Email: feraud.tchuisseu@umontreal.ca.

1 Introduction

Governments that regulate firms enforce the law with two levers: the probability that a violation is detected and the fine (penalty) imposed if it is. A long tradition in the economics of enforcement, beginning with Becker (1968), treats these levers as substitutes that affect behavior only through their product, the expected fine, and draws a sharp prescription from that equivalence: because detection is costly and fines are near-costless transfers, governments should detect rarely and punish heavily. This prescription rests on a particular view of what a violation is, a discrete act that a firm either commits or does not. However, many of the violations that labor, tax, and environmental regulators care about are not discrete acts but durable states that a firm occupies over time, and minimum-wage underpayment is a leading case: a firm that pays below the legal floor does so month after month, inside an ongoing employment relationship. When a violation is a state rather than an act, an inspection can do something a fine cannot: it can end the violation, not merely price it. Detection and fine may then stop being substitutes.

In this paper, we study whether raising detection and raising penalties are interchangeable ways to reduce minimum-wage non-compliance, and why the answer turns on treating underpayment as a persistent employment relationship rather than a one-time offense. The question matters in practice because the two levers are not equally available to the governments that most need them. Detection requires inspectors, information, and administrative capacity that are scarce in many low- and middle-income labor markets, whereas a higher fine can be legislated at little direct cost. If the instruments were substitutes, a government short on inspectors could raise fines and reach the same outcome. We show that it cannot.

Detection does something penalties cannot. A fine raises the price of underpayment, so a firm that keeps violating moves its wage closer to the legal floor, but the job goes on operating below it. Detection is different because a detected violation cannot simply continue as the same arrangement: the job is brought into compliance when its productivity can support the legal wage, and it ends when it cannot. Detection therefore governs the rate at which jobs leave the violating state, while a fine only changes the wage paid inside a job that goes on violating. The two instruments reshape non-compliance through different channels even when they deliver the same expected fine. The distinction is clearest once non-compliance is split into its two margins, the incidence of violations, meaning the share of workers paid below the floor, and their depth, meaning how far below the legal floor underpaid wages fall on average (Bhorat et al., 2012): penalties compress depth, while detection reduces incidence. This non-equivalence differs from the one already in the enforcement literature, where Polinsky and Shavell (1979) show that risk aversion breaks the Becker equivalence through the risk-bearing cost of a higher fine. That mechanism acts on the firm's decision to violate, whereas ours acts on the composition of the stock of violations and holds even when firms are risk neutral, as they are in the model we estimate.

Colombia is an unusual laboratory for separating the two levers. Minimum-wage underpayment is pervasive: in 2012, the year before the reform, 53.4 percent of private-sector wage workers with at most a secondary education earned less than the statutory minimum. In 2013, Law 1610 raised the statutory ceiling on penalties for labor violations by a factor of fifty, moving the fine lever in an environment where the detection lever was weak and changed little. The reform was national, but its effect was not uniform, because departments entered the period with very different levels of underpayment, and a national increase in the expected cost of violating has more room to operate where violations were more common. The design therefore assumes that, absent the reform, more- and less-exposed departments would not have diverged in underpayment for reasons correlated with exposure, an assumption we probe with pre-trend, placebo, and auxiliary-outcome checks.

We bring three connected pieces of evidence to the question. First, we estimate reduced-form event studies that compare departments by their pre-reform exposure to underpayment, measured as the average departmental shortfall in 2010 and 2011, around the 2013 reform. Second, we estimate a search-and-matching model with imperfect enforcement on the pre-reform cross-section. We use this class of model because it generates wage dispersion across firms, which is what lets it speak to the depth of violations and explain why detection can cut the incidence of violations while leaving the deepest ones in place. In the model, firms and workers bargain over wages, firms may pay below the legal floor, and the enforcement environment is summarized by a detection probability and a fine severity. Third, we use the estimated model to run enforcement counterfactuals, with the reduced-form estimates disciplining the size of the structural enforcement shock. The reduced form tells us what the reform did; the model tells us why the two instruments would do it differently.

The reform produced a clear pattern in the data. In departments more exposed to pre-reform underpayment, the depth of violations fell after 2013, with the conditional shortfall declining by 2.5 percentage points at mean exposure. High- and low-exposure departments were on similar trends in non-compliance before the reform, and this decline appears only afterward. The incidence of violations shows no comparable change: the point estimates are negative but not statistically distinguishable from zero. The reform compressed how far below the floor underpaid workers were paid, without a measurable change in how many workers were underpaid.

The response on the depth but not on the incidence margin appears robust and does not seem to be driven by a concurrent reform. The depth response holds under an incidence-based definition of exposure, over an extended 2008 to 2018 window, and against a placebo sample of tertiary-educated workers who were not bound by the minimum wage and show no effect. Colombia's 2012 payroll-tax reform raised formal employment over the same period (Kugler et al., 2017a), but the auxiliary outcomes we examine, including formality, sectoral allocation, and hours, show none of the broad reallocation a formalization shock would produce. The margin that moves is the depth of underpayment, which is what a higher fine should compress. To interpret this pattern and compare the two instruments, we estimate the model on the 2012 pre-reform cross-section. The estimation matches the main pre-reform objects. The estimates describe a weak baseline enforcement environment: violations are detected infrequently, and the implied fine severity is modest, so that the expected cost of underpaying is small relative to the savings. The model also performs well on additional features of the wage distribution that are not used to choose the parameters. Finally, when we discipline the fine change using the observed decline in the depth of violations, the model predicts a mean-wage response of 4.6 percent, close to the reduced-form estimate of 4.5 percent. We therefore treat the exercise as a disciplined mapping of the depth response into the model rather than as a complete account of every margin of the reform. The comparison of the two instruments is the paper's central quantitative result. We hold each lever to the same target, the aggregate reduction in underpayment, and ask what it must do to reach it. The fine policy reaches the target by raising severity by a factor of 5.1, lowering both the incidence (by 4.2 percentage points) and the depth of violations, with no rise in unemployment. The detection policy reaches the same target by raising the detection probability by a factor of 2.5, working through clearance: it cuts the incidence of violations by 11 percentage points but raises the conditional shortfall by 4.1 percentage points, because the remaining violators are negatively selected toward low-productivity matches, and raises unemployment by 0.4 percentage point. The two policies deliver the same reduction in aggregate underpayment and two different labor markets, which is the non-equivalence result made quantitative.

This paper contributes to three literatures. The first is the economics of public enforcement, in which detection and penalties are the two instruments and the expected sanction is the central object (Becker, 1968; Polinsky and Shavell, 1979, 2000; Mookherjee and Png, 1994). That literature already knows the instruments are imperfect substitutes, because with risk-averse offenders a higher fine carries risk-bearing costs and changes deterrence per unit of expected fine. That non-equivalence operates on the decision to commit a one-time act. We contribute a different one, which operates on the composition of the stock of violations when the violation is a durable state and which holds under risk neutrality: detection clears violations, while penalties only price them.

The second literature studies minimum-wage non-compliance, beginning with the classic compliance model of Ashenfelter and Smith (1979) and extending to developing-country measures of incidence and depth (Bhorat et al., 2012; Rani et al., 2013; Ham, 2015). Its central measurement insight is that non-compliance has two margins, incidence and depth, both needed to describe how far paid wages fall short of the legal floor. We use that decomposition not only to measure underpayment but to identify a mechanism, since the margin that responds to a reform reveals which instrument is operating. To our knowledge, we are the first to use an enforcement reform to estimate the incidence and depth responses separately and to map them into a structural comparison of detection and penalties. A related strand treats non-compliance as an adjustment margin that regulators may tolerate when full enforcement would threaten employment in low-productivity firms (Basu et al., 2010; Garnero and Lucifora, 2022; Badaoui and Walsh, 2022; Marshall, 2025); we take the productivity constraint seriously but ask the prior question of which instrument changes which margin. Recent evidence also shows that minimum-wage effects and compliance depend on the enforcement regime, in both developing and high-income settings (Soundararajan, 2019; Clemens and Strain, 2022).

The third literature is structural, taking the minimum wage and informality to equilibrium labor-market models (Flinn, 2006; Meghir et al., 2015; Ulyssea, 2018). Where this work treats the minimum wage as a binding floor or informality as a firm-level registration decision, we model non-compliance within the employment relationship, so that both the incidence and the depth of underpayment are equilibrium objects, and we identify the detection probability separately from fine severity. That separation is what makes the instrument comparison possible. Our results also connect to the empirical literature on labor-law enforcement in Latin America (Ronconi, 2010; Almeida and Carneiro, 2012; Pignatti, 2026), which establishes that labor inspection raises compliance but varies the number of inspectors and so moves the detection margin alone, and to Stansbury (2024), who documents that the expected costs faced by detected violators are often close to the back wages owed. That weak-deterrence environment is the precondition our model estimates and the margin Law 1610 moved.

The paper proceeds as follows. Section 2 presents the model and the non-equivalence result. Section 3 describes the institutional setting and Law 1610. Section 4 presents the data and empirical design. Section 5 reports the reduced-form evidence. Section 6 presents the structural estimates and model fit. Section 7 reports the enforcement counterfactuals. Section 8 concludes.

2 A Model of Enforcement with Non-Compliance

We build a model in which minimum-wage enforcement has two instruments. The government can raise the probability that a violating job is detected, or it can raise the penalty paid conditional on detection. In a static Becker-style enforcement problem, these instruments are close substitutes because the private cost of violation depends on the expected fine (Becker, 1968). In a labor market, this equivalence breaks down. Detection does not only change the expected price of violating the law. It also changes the state of the job. Once a violation is detected, the same non-compliant arrangement cannot simply continue: it must either be brought into compliance or dissolved if the match is not viable at the legal minimum wage. The framework follows the search-and-bargaining approach to minimum wages in Flinn (2006), but shifts attention to imperfect enforcement: job-contact rates are taken as exogenous, and subminimum wages arise as equilibrium outcomes.

2.1 Environment

Time is continuous. A unit mass of risk-neutral workers and a continuum of firms discount future payoffs at rate r. Each period, workers are either unemployed or employed. Unemployed workers receive flow value b and meet firms at Poisson rate λ. A meeting draws productivity z from distribution F, with density f. Productivity is observed by both parties and remains fixed for the duration of the match. Jobs separate exogenously at rate ρ. Wages are determined by Nash bargaining, with worker bargaining power η.

A statutory minimum wage $\omega_M > 0$ applies to all jobs. A match can comply by paying $w \geq \omega_M$, or it can violate by paying $w < \omega_M$. Violations are detected at Poisson rate π. Conditional on detection, the firm pays a penalty that is increasing and convex in the wage shortfall:

$$\Phi(\omega_M - w) = \kappa(\omega_M - w)[1 + (\omega_M - w)]. \tag{1}$$

The parameter κ indexes penalty severity. The specification has a positive marginal penalty at the legal threshold, $\Phi'(0) = \kappa$, and makes deeper violations increasingly costly. Since inspections arrive at rate π, the expected marginal cost of a small shortfall is governed by the composite

$$\chi \equiv \pi\kappa. \tag{2}$$

We refer to χ as the expected enforcement wedge. The expected flow penalty for an undetected violating job is therefore

$$\pi\Phi(\omega_M - w) = \chi(\omega_M - w)[1 + (\omega_M - w)]. \tag{3}$$

Detection has two possible consequences. Once a violation is detected, the firm pays the penalty and cannot continue violating: it must either comply with the statutory minimum wage or terminate the match. The match therefore survives detection only if the firm's continuation value under compliance is non-negative. Otherwise, separation occurs. This creates two violation regions. In low-productivity violating matches, detection destroys the job because the firm cannot profitably continue at the minimum wage. In high-productivity violating matches, detection converts the job into compliant employment at wage ω_M. Section 2.3 shows that the productivity threshold separating these two regions is exactly ω_M: violations with $z < \omega_M$ are destroyed upon detection, while violations with $z \geq \omega_M$ survive and are forced into compliance. Thus, enforcement affects non-compliant jobs through two channels: penalties raise the private cost of underpayment, and detection changes the continuation state of the match.

2.2 Value Functions

Let U denote the value of unemployment, and let rU denote its flow-equivalent value. Firm values are measured relative to an unmatched firm, whose value is normalized to zero.

Consider first a match that pays a compliant wage $w_C(z) \geq \omega_M$. Let $W_C(z)$ be the worker's value in this state. Since no enforcement action is relevant, the worker receives the wage until the match separates exogenously:

$$rW_C(z) = w_C(z) + \rho(U - W_C(z)). \tag{4}$$

The firm's value in the same compliant match is denoted by $J_C(z)$. The firm receives output net of the wage and the match destroys at rate ρ:

$$rJ_C(z) = z - w_C(z) - \rho J_C(z), \tag{5}$$

Detection of a violation forces the firm to comply with the statutory minimum wage whenever the match remains viable. The firm's continuation value under compliance at wage ω_M is therefore

$$J_M(z) = \frac{z - \omega_M}{r + \rho}.$$

Thus a detected violation continues as compliant employment if $z \geq \omega_M$, and separates otherwise.

Now consider a match that pays a subminimum wage $w < \omega_M$. Inspections arrive at rate π, and detection triggers the penalty $\Phi(\omega_M - w)$. The continuation state after detection depends on whether the match can survive under compliance.

For low-productivity violations, $z < \omega_M$, compliance at the statutory wage is not viable. Detection therefore destroys the match. Denote the worker's value in this state by $W_V^\ell(z)$. It satisfies

$$rW_V^\ell(z) = w_V^\ell(z) + (\rho + \pi)(U - W_V^\ell(z)). \tag{6}$$

The corresponding firm value, $J_V^\ell(z)$, reflects three forces: current output net of the subminimum wage, the expected inspection penalty, and the loss of the match through either exogenous separation or detection:

$$rJ_V^\ell(z) = z - w_V^\ell(z) - \pi\Phi(\omega_M - w_V^\ell(z)) - (\rho + \pi)J_V^\ell(z). \tag{7}$$

For high-productivity violating matches, $z \geq \omega_M$, detection converts the job into compliant employment at wage ω_M. The worker's value in that post-detection compliant state satisfies:

$$rW_M(z) = \omega_M + \rho(U - W_M(z)).$$

While the job remains undetected, the worker's value is $W_V^h(z)$. The worker receives the subminimum wage, faces exogenous separation, and moves to $W_M(z)$ if the violation is detected:

$$rW_V^h(z) = w_V^h(z) + \rho(U - W_V^h(z)) + \pi(W_M(z) - W_V^h(z)). \tag{8}$$

The firm's value in a high-productivity violating match is $J_V^h(z)$. Detection imposes the penalty and moves the firm to the compliant continuation value $J_M(z)$:

$$rJ_V^h(z) = z - w_V^h(z) - \pi\Phi(\omega_M - w_V^h(z)) - \rho J_V^h(z) + \pi(J_M(z) - J_V^h(z)). \tag{9}$$

The value of unemployment closes the system of worker values. Unemployed workers receive flow payoff b and meet firms at rate λ. A job opportunity is accepted only when the value of employment exceeds the value of unemployment.

$$rU = b + \lambda \int \max\{W(z) - U, 0\}\, dF(z). \tag{10}$$

These equations make explicit the two roles of enforcement. Penalties reduce the current value of underpayment, while detection changes the continuation state of the match. Detection destroys non-viable violations and converts viable violations into compliant jobs. The next section uses these values to characterize the Nash-bargained wage schedule.

2.3 Equilibrium wage regimes

Wages are determined by bilateral Nash bargaining after match productivity z is realized, with workers bargaining power $\eta \in (0,1)$. In each wage-setting regime, the bargained wage maximizes the Nash product between the worker's surplus over unemployment and the firm's continuation value. Appendix A.1 derives the regime-specific Nash problems and their first-order conditions. The equilibrium wage schedule is characterized by three productivity thresholds:

$$z^* = rU + \chi(\omega_M - rU)\left[1 + (\omega_M - rU)\right], \qquad z^c = \frac{\omega_M - (1-\eta)rU}{\eta}, \qquad z^v = \chi\omega_M + (1-\chi)z^c.$$

The threshold z^* is the acceptance threshold: matches below it are not formed. The statutory minimum ω_M separates detected violations that are destroyed from those that can survive in compliance. The threshold z^v marks the point at which the high-productivity violating wage reaches the statutory minimum, so that the violating shortfall closes. Finally, z^c is the point at which the unconstrained compliant Nash wage reaches the statutory minimum; for higher productivities, the legal floor no longer binds. Under the maintained regularity condition

$$\chi\left[1 + 2(\omega_M - rU)\right] < 1,$$

the thresholds satisfy the following condition whenever enforcement is active, and the non-compliance region is non-empty.

$$z^* < \omega_M < z^v < z^c$$

The equilibrium wage schedule is

$$w^*(z) = \begin{cases} \text{no match}, & z < z^*, \\ w_V^\ell(z) = \omega_M - x_V^\ell(z), & z^* \le z < \omega_M, \\ w_V^h(z) = \omega_M - x_V^h(z), & \omega_M \le z < z^v, \\ \omega_M, & z^v \le z < z^c, \\ w_C(z) = (1-\eta)rU + \eta z, & z \ge z^c, \end{cases} \tag{11}$$

where $x_V^\ell(z) > 0$ and $x_V^h(z) > 0$ are the equilibrium wage shortfalls in the low- and high-productivity violation regions, respectively.

The schedule has four employment regions. Low-productivity accepted matches, $z \in [z^*, \omega_M)$, violate the minimum wage and are destroyed if detected. High-productivity matches, $z \in [\omega_M, z^v)$, also violate, but detection converts them into compliant employment at wage ω_M. Matches in $[z^v, z^c)$ bunch at the statutory minimum. These matches are productive enough that the violating shortfall has closed, so non-compliance is no longer optimal. Yet they are not productive enough for the unconstrained compliant Nash wage, $w_C(z) = (1-\eta)rU + \eta z$, to clear the floor, since $w_C(z) < \omega_M$ for $z < z^c$. The minimum wage therefore binds, and both parties settle at $w^*(z) = \omega_M$. Finally, matches above z^c pay the unconstrained compliant Nash wage. An important implication is the downward jump at $z = \omega_M$. Detection has different continuation values on the two sides of the threshold. Below ω_M, detection destroys the match. Above ω_M, detection converts the match into compliant employment at wage ω_M. This conversion option raises the worker's continuation value in high-productivity violating matches and therefore substitutes for current compensation in Nash bargaining. Finally, bunching at the statutory minimum arises because enforcement gives the minimum wage an additional role beyond a lower bound on compliant wages. When $\chi > 0$, firms with productivities in $[z^v, z^c)$ prefer to comply at ω_M rather than pay the unconstrained Nash wage below the statutory minimum. When $\chi = 0$, $z^v = z^c$, and the bunching interval collapses.

Proposition 1 (Equilibrium wage schedule)**.** *Given $rU < \omega_M$, $\chi[1 + 2(\omega_M - rU)] < 1$, and the penalty schedule $\Phi(\omega_M - w) = \kappa(\omega_M - w)[1 + (\omega_M - w)]$, the equilibrium wage schedule is given by equation (11), with thresholds z^*, z^v, and z^c defined above. The violating shortfalls $x_V^\ell(z)$ and $x_V^h(z)$ are the closed-form roots of the regime-specific Nash first-order conditions derived in Appendix A.1.*

Proof. See Appendix A.1.

Throughout the remainder of the paper, we maintain $rU < \omega_M$ and $\chi[1 + 2(\omega_M - rU)] < 1$. In the quantitative implementation, we verify that both conditions are satisfied.

2.4 Steady-State equilibrium, Incidence, and Shortfall

In steady state, all worker stocks are invariant: for each employment state, inflows equal outflows. Proposition 1 partitions accepted matches into four regions. Let

$$q_L = F(\omega_M) - F(z^*), \qquad q_H = F(z^v) - F(\omega_M), \qquad q_M = F(z^c) - F(z^v), \qquad q_A = 1 - F(z^c).$$

The regions q_L and q_H correspond to violating jobs. Low-productivity violations are destroyed upon detection, whereas high-productivity violations are converted into compliant minimum-wage jobs. The regions q_M and q_A correspond to compliant jobs, respectively at the statutory minimum and above it. Let

$$q_V = q_L + q_H$$

denote the total measure of new matches that enter violation.

The employment stocks follow from flow balance. Violating jobs are created by accepted offers in q_V and leave violation either through exogenous separation or detection:

$$e_V = \frac{u\lambda q_V}{\rho + \pi}. \tag{12}$$

Minimum-wage jobs are created either directly, through accepted offers in q_M, or indirectly, through the detection and conversion of high-productivity violations:

$$e_M = \frac{u\lambda q_M}{\rho} + \frac{u\lambda\pi q_H}{\rho(\rho+\pi)}. \tag{13}$$

Above-minimum jobs are created through accepted offers in q_A and leave only through exogenous separation:

$$e_A = \frac{u\lambda q_A}{\rho}. \tag{14}$$

The unemployment rate follows directly from flow balance. Workers leave unemployment at rate $\lambda[1-F(z^*)]$ when they receive an acceptable offer. They return to unemployment through exogenous separations at rate ρ from all jobs, and through detection of low-productivity violations, which destroys the match. Detection of high-productivity violations converts the job into compliant employment and does not return the worker to unemployment. The flow-balance condition for unemployment is therefore

$$u\lambda[1-F(z^*)] = \rho(1-u) + \pi e_{V,L}, \qquad e_{V,L} \equiv \frac{u\lambda q_L}{\rho+\pi}.$$

Substituting $e_{V,L}$ and solving for u:

$$u = \frac{\rho}{\rho + \lambda\left[\dfrac{\rho\, q_L}{\rho+\pi} + q_H + q_M + q_A\right]}. \tag{15}$$

The term $\rho q_L/(\rho+\pi) < q_L$ discounts low-productivity violating matches by their higher exit rate $\rho+\pi$: these jobs are shorter-lived and contribute less to employment than their share of new matches suggests.

Non-compliance is summarized by two moments. Incidence is the share of employed workers currently paid below the statutory minimum,

$$I = \frac{e_V}{1-u} = \frac{e_V}{e_V + e_M + e_A}, \tag{16}$$

Depth is the average proportional shortfall among violating jobs,

$$G \equiv \mathbb{E}\left[\frac{\omega_M - w(z)}{\omega_M} \,\middle|\, w(z) < \omega_M\right]. \tag{17}$$

Thus I captures the stock of violations, while G captures their conditional severity. Detection affects incidence directly by shortening violation spells and converting detected high-productivity violations into compliant jobs. Penalty severity affects both incidence and depth indirectly through wages and the thresholds that determine the composition of jobs across regimes.
The mass of workers at the minimum wage is

$$B = \frac{e_M}{1-u}. \tag{18}$$

In the model, this mass is informative about both direct bunching at the legal floor and conversion of high-productivity violations into compliant jobs.

Definition 1 (Stationary equilibrium with imperfect enforcement)**.** Given the model primitives and the maintained regularity conditions, a stationary equilibrium is a tuple

$$(w^*(\cdot), z^*, z^v, z^c, u, e_V, e_M, e_A)$$

such that:

1. For each match productivity z, the wage $w^*(z)$ solves the relevant regime-specific Nash bargaining problem, subject to the minimum-wage enforcement rule. Proposition 1 characterizes the equilibrium wage schedule.

2. The thresholds (z^*, z^v, z^c) are consistent with $w^*(\cdot)$ and partition accepted matches into low-productivity violations, high-productivity violations, minimum-wage jobs, and above-minimum jobs.

3. The worker stocks (u, e_V, e_M, e_A) are stationary: inflows and outflows balance in each employment state, as characterized by equations (12)–(15).

The aggregate non-compliance moments (I, G, B) are then induced by the equilibrium wage schedule, the regime partition, and the stationary worker stocks.

2.5 Non-equivalence of detection and penalty severity

The central implication of the model is that inspection probability and penalty severity are not equivalent enforcement instruments. In a static enforcement problem, the firm cares only about the expected penalty. Holding $\chi = \pi\kappa$ fixed would therefore leave behavior unchanged. In the present environment, however, detection has an additional role: it changes the state of the match. A detected violation exits the violation state, either because the match is destroyed or because the firm is forced into compliance. Penalty severity changes the price of underpayment, but it does not by itself change the transition rate out of non-compliance.

The following proposition isolates this stock-flow channel. Consider a change in enforcement that raises π and lowers κ so that the expected enforcement wedge $\chi = \pi\kappa$ remains fixed. Since the regime partition is held fixed, the masses q_V, q_M, q_H, and q_A are unchanged. Detection nevertheless changes the duration of violation spells.

Proposition 2 (Non-equivalence of enforcement instruments)**.** *Fix the regime partition and consider two enforcement regimes with the same expected enforcement wedge $\chi = \pi\kappa$ but different inspection probabilities π. The steady-state incidence of non-compliance is*

$$I(\pi) = \frac{\rho q_V}{\rho q_V + (\rho + \pi)(q_M + q_A) + \pi q_H}. \tag{19}$$

Therefore,

$$\frac{\partial I(\pi)}{\partial \pi} = -\frac{\rho q_V(q_M + q_A + q_H)}{[\rho q_V + (\rho + \pi)(q_M + q_A) + \pi q_H]^2} < 0 \tag{20}$$

whenever $q_V > 0$ *and* $q_M + q_A + q_H > 0$. *Hence detection and penalty severity are not interchangeable: even holding the expected enforcement wedge fixed, a higher inspection probability lowers the stock of violations by increasing the exit rate from non-compliance.*

Proof. See Appendix A.2.

The result is not driven by a change in the negotiated wage schedule. It is a duration effect. A higher penalty makes violation more expensive while the match remains undetected. A higher inspection probability both raises expected penalty and accelerates the arrival of the state change that ends the violation. Detection is therefore both a price instrument and a clearance technology, whereas penalty severity is only a price instrument. This is the source of the non-equivalence.

2.6 Detection and the depth of violations

The second implication concerns the depth of violations. Detection can reduce the incidence of non-compliance while increasing the conditional shortfall among the violations that remain. The mechanism combines selection and continuation values. When detection becomes more likely, some high-productivity violations are converted into compliant minimum-wage jobs. The remaining violating jobs are therefore more negatively selected on productivity. In addition, for high-productivity violators, detection is not only a punishment: it also moves the worker into compliant employment rather than unemployment. This conversion option affects the wage bargain and can lower the wage paid in the violating state.

Lemma 1 (Detection and the depth of high-productivity violations)**.** *In the limiting case* $\chi = 0$*, the high-productivity violating wage can be written as*

$$w_V^h(z) = w_C(z) + \Delta(\pi)\eta(z - z^c), \tag{21}$$

where $w_C(z) = (1 - \eta)rU + \eta z$ *is the unconstrained Nash wage,* $\Delta(\pi) = \pi/(r + \rho)$*, and* z^c *is the productivity level at which* $w_C(z) = \omega_M$*. For high-productivity violators below the unconstrained compliance threshold,* $z < z^c$*,*

$$\frac{\partial w_V^h(z)}{\partial \pi} < 0. \tag{22}$$

Thus, a higher detection probability can lower the wage paid by the high-productivity violations that remain. Detection removes some violations, but it can make the remaining violations deeper.

Proof. See Appendix A.3.

Penalty severity works differently. A higher κ, holding π fixed, raises the marginal cost of deep underpayment. This compresses the lower tail of subminimum wages and reduces the conditional shortfall. It need not generate a large decline in the incidence of violations, especially when many violating matches remain privately profitable after a small wage adjustment. The model therefore predicts two distinct enforcement patterns. Detection-based enforcement should primarily reduce the incidence of violations. Penalty-based enforcement should reduce the depth of violations, with a weaker effect on their incidence.

This distinction guides the rest of the paper. The reduced-form design asks whether Colombia's enforcement reform generated the incidence response or the depth response. The structural estimation uses the pre-reform wage distribution, unemployment, and the mass around the minimum wage to discipline the baseline economy. The counterfactuals then compare policies that raise π with policies that raise κ, holding fixed the model's equilibrium logic. The object is not only whether enforcement matters. It is which enforcement instrument matters, through which margin, and with what labor-market consequences.

3 Institutional Setting and Detection as a Clearance Technology

3.1 Minimum-wage coverage and labor inspection

Colombia has a national statutory minimum wage, the *Salario Mínimo Mensual Legal Vigente*. The minimum wage is set annually through a tripartite process involving the government, employers, and workers, and it applies nationally (Congreso de Colombia, 1996). This national wage floor is useful for our design because regional differences in non-compliance do not reflect different legal wage floors. They reflect differences in local wage distributions, enforcement conditions, and the extent to which employers pay below the national floor.

The legal floor, however, is not self-enforcing. Enforcement is delegated to the labor inspection system under the Ministry of Labor. Labor inspectors have authority to monitor compliance with labor norms, investigate violations, mediate disputes, and impose administrative sanctions (Congreso de Colombia, 2013). Their mandate is not only punitive. The ILO framework defines labor inspection around securing the enforcement of legal provisions relating to working conditions and worker protection (International Labour Organization, 1947). Colombia's inspection system follows this logic: inspectors can verify compliance with the labor code, including minimum wages, working time, social-security contributions, and occupational safety obligations (Pignatti, 2026).

This institutional point matters for the model. When a subminimum wage is detected, the employer cannot continue the same illegal arrangement as if nothing happened. The violation must be corrected, regularized, sanctioned, or ended. In the model, we summarize this feature by assuming that detection clears the violating state. If the job can operate at the statutory minimum, detection converts it into a compliant minimum-wage job. If it cannot, detection destroys the match. This is a simplification, but it captures the relevant enforcement technology: detection changes the duration and state of non-compliant jobs.

The Colombian setting is therefore not only an empirical application. It is a useful laboratory for the theoretical distinction in Section 2. Penalties affect the price of underpayment. Inspections affect both the price and the probability that an ongoing violation is forced out of the violating state.

3.2 Law 1610 and the enforcement environment

The main reform we study is Law 1610, enacted on January 2, 2013. The law regulated labor inspections and labor formalization agreements. Its most direct sanctioning provision modified the enforcement authority of labor inspectors and increased the maximum administrative fine for labor violations to 5,000 monthly minimum wages (Congreso de Colombia, 2013). Relative to the previous ceiling of 100 monthly minimum wages, this was a large increase in the potential penalty faced by employers (Ministerio de la Protección Social, 2011).

The law did not change the statutory minimum wage. It changed the consequences of violating existing labor obligations. This distinction is central. A minimum wage can be high, salient, and legally universal, while still being weakly binding if the expected cost of violating it is small. Law 1610 raised that expected cost by increasing the severity of sanctions and by strengthening the institutional apparatus through which labor inspectors could induce compliance.

Law 1610 also had broader enforcement provisions. It clarified inspection functions, regulated labor formalization agreements, and formed part of a wider strengthening of labor-law enforcement. These features are important for interpretation. The reform should not be described as a pure fine reform. At the same time, the increase in the fine ceiling provides the clearest mapping between the statutory reform and the penalty-severity margin in the model.

This distinction shapes how we interpret the reduced-form evidence. If the relevant change primarily increased detection, the model predicts a strong incidence response: violating jobs should leave the violating state more quickly. If the relevant change primarily increased penalty severity, the model predicts a depth response: firms can remain in violation but reduce the size of the underpayment. The empirical section therefore does not rely only on an average decline in non-compliance. It asks which margin moves.

3.3 Inspector expansion and identification

Law 1610 occurred during a period in which Colombia was also expanding labor-inspection capacity. As part of the Colombian Action Plan Related to Labor Rights, the government committed to hire additional labor inspectors. Pignatti (2026) studies this inspector expansion and reports that the number of inspectors increased from about two to four per 100,000 individuals between 2011 and 2014. He exploits the allocation rule used to distribute inspector vacancies across departments and years. The rule assigned hiring targets using a department-level score designed to capture the risk of labor-law non-compliance; within an integer value of that score, whether the decimal component crossed a 0.5 cutoff generated discontinuous variation in the number of inspectors allocated to otherwise similar departments. Using this variation, Pignatti (2026) finds that additional inspectors increased labor-law compliance in the formal sector, with weaker or no effects in the informal sector.

The inspector expansion helps discipline the interpretation of our design. It confirms that detection was an active enforcement margin in Colombia, while also ruling out a literal interpretation of the post-2013 period as a one-dimensional penalty change. Colombia was changing enforcement on more than one dimension.

At the same time, the inspector expansion is not automatically a threat to our design. Our variation is not Pignatti's inspector-allocation discontinuity. We use pre-reform departmental exposure to minimum-wage non-compliance, measured before Law 1610, to predict which departments should have been more affected by a stronger enforcement environment. A contemporaneous policy threatens this design only if it generates differential post-reform changes across high- and low-exposure departments in a way that is correlated with our exposure measure and that mimics the timing and margin of our estimated effects.

The model disciplines that interpretation. A relative increase in detection should primarily reduce the incidence of violations. A relative increase in penalty severity or expected liability should primarily reduce the depth of violations. We therefore estimate these margins separately and examine employment, formality, contracts, hours, and sectoral allocation to assess whether the exposure design is capturing broader regional labor-market shifts rather than minimum-wage enforcement.

A second contemporaneous reform is Law 1607 of December 2012, which reduced non-wage labor costs for eligible employers. Existing studies show that this payroll-tax reform affected formal employment and related labor-market outcomes (Kugler et al., 2017a; Morales and Medina, 2017; Kugler et al., 2017b). Again, the presence of this reform is not by itself sufficient to invalidate the design. It matters if the payroll-tax reform generated differential changes across departments with different pre-reform minimum-wage non-compliance exposure. The economic margin is also different. Law 1607 reduced the cost of formal employment; Law 1610 raised the cost of violating labor obligations. The reduced-form section therefore examines whether the main response is concentrated in minimum-wage shortfalls rather than in broad formalization or employment margins.

The institutional setting leaves us with a disciplined empirical question. Colombia experienced a national strengthening of labor enforcement, including a large increase in potential sanctions and an expansion of inspection capacity. We use cross-departmental pre-reform exposure to measure where the enforcement change had more bite, and we use the model's margin-specific predictions to interpret the response. The key test is not whether enforcement changed. It did. The key test is whether the observed response looks like detection or like penalty severity.

4 Data and Empirical Design

4.1 Data, sample, and measurement

We use individual-level data from Colombia's *Gran Encuesta Integrada de Hogares*, harmonized through the World Bank Global Labor Database. The GEIH is Colombia's main household labor-force survey. It reports employment status, labor income, hours, sector of employment, demographic characteristics, and department of residence. Our cleaned data cover 2008 to 2018. The baseline reduced-form analysis focuses on 2010 to 2015, with the additional years used for robustness.

The main empirical sample includes private-sector paid workers between ages 15 and 64 with at most secondary education. The private-sector restriction follows the enforcement problem studied in the paper: whether private employers comply with the statutory wage floor. The education restriction focuses the analysis on workers for whom the minimum wage is most likely to bind. Workers with tertiary education are used later as a falsification group. The structural estimation uses the same demographic and education restrictions, and adds unemployed workers when constructing the unemployment moment.

We construct hourly wages using gross labor income reported for the worker's main paid job. For salaried workers, the underlying GEIH question records earnings before deductions. We combine this information with payment frequency and usual weekly hours. We then merge the wage data with Colombia's gross statutory monthly minimum wage and convert the legal floor into an hourly value using the ordinary workweek in effect during our sample period. Appendix B gives the full data-construction procedure.

All wage variables are expressed relative to the statutory hourly minimum wage. We classify a worker as paid below the minimum wage when the reported hourly wage is below 99.5 percent of the statutory hourly minimum. This tolerance prevents small reporting, rounding, and conversion differences around the minimum wage from being classified as economically meaningful violations.

We measure non-compliance with three objects, matching the notation of the model. The first is the incidence of subminimum wages,

$$I_{idt} = \mathbf{1}\left\{w_{idt} < 0.995\,\omega_{M,t}\right\}. \tag{23}$$

This is the extensive margin: whether worker i in department d and year t is paid below the legal floor. The second is the conditional shortfall among workers paid below the floor,

$$G_{idt} = \frac{\omega_{M,t} - w_{idt}}{\omega_{M,t}} \qquad \text{if } w_{idt} < 0.995\,\omega_{M,t}. \tag{24}$$

This is the depth of the violation. The third is the unconditional shortfall,

$$UGap_{idt} = I_{idt} \times G_{idt}. \tag{25}$$

This object combines incidence and depth. It measures total underpayment relative to the statutory minimum wage among employed workers. In the theory, this is the empirical counterpart of $I \times G$.

We also construct wages, hours, social-security coverage, contract status, and broad sectoral allocation. These outcomes are not the central objects of the paper. They help discipline whether changes in non-compliance are accompanied by broader labor-market shifts.

4.2 Exposure and event-study design

Law 1610 was a national reform. It changed the enforcement environment for Colombian employers simultaneously, with no untreated department and no staggered geographic rollout. We therefore use cross-departmental differences in pre-reform exposure to non-compliance. Departments with greater underpayment before the reform had more employment relationships for which a stronger enforcement environment could bind.

Our baseline exposure measure is the department-level unconditional shortfall averaged over 2010 and 2011,

$$Exposure_d = \frac{1}{2}\left(\overline{UGap}_{d,2010} + \overline{UGap}_{d,2011}\right). \tag{26}$$

The measure is high in departments where many workers were paid below the statutory minimum and where violations were relatively deep. It therefore captures the pre-reform intensity of underpayment, not only the fraction of workers below the floor. In robustness checks, we reconstruct exposure using the pre-reform incidence of non-compliance.

Appendix Figure 6 and Appendix Table 7 report the department-level exposure values used in the design. They show that exposure varies substantially across departments before the reform, which is the regional variation that identifies the event-study coefficients.

Exposure is defined for the 24 department-level units observed in both 2010 and 2011. Nine additional units appear in the harmonized data only from 2012 onward. Since their pre-reform exposure cannot be measured consistently, we do not impute exposure values for them in the exposure-based specifications.

We estimate a continuous-treatment event-study specification,

$$y_{idt} = \alpha_d + \gamma_t + \sum_{s \neq 2012} \beta_s \left(Exposure_d \times \mathbf{1}\{t = s\}\right) + X'_{idt}\theta + \varepsilon_{idt}. \tag{27}$$

The outcome y_{idt} is a non-compliance measure or another labor-market outcome. Department fixed effects absorb time-invariant differences across departments, including the level of pre-reform underpayment. Year fixed effects absorb common shocks, including annual changes in the statutory minimum wage and national labor-market conditions. The vector X_{idt} includes age, age squared, and gender. The omitted year is 2012, the year immediately before Law 1610.

The coefficients β_s measure how outcomes evolved in more exposed departments relative to less exposed departments. The 2010 and 2011 coefficients are pre-trend checks. The 2013 to 2015 coefficients measure the dynamic post-reform response. The identifying assumption is that, absent the change in the enforcement environment, departments with different pre-reform underpayment exposure would have followed parallel trends in the outcomes of interest.

We cluster standard errors at the department level. Because the number of clusters is limited, we report confidence intervals based on a wild cluster bootstrap with Webb weights. This inference choice is maintained across the main reduced-form outcomes and the auxiliary labor-market outcomes.

4.3 Descriptive evidence and design support

Before the reform, minimum-wage non-compliance was widespread. Appendix Table 6 reports the pre-reform descriptive statistics. Between 2010 and 2012, approximately 55 percent of private-sector paid workers in the baseline sample earned below 99.5 percent of the statutory minimum wage. Among underpaid workers, the average shortfall was approximately 36 percent of the legal floor.

The unconditional shortfall was approximately 20 percent. The reform therefore occurred in a labor market where the statutory floor was salient but imperfectly enforced.

Pre-reform underpayment also varied substantially across departments. Appendix Figure 6 and Appendix Table 7 show the regional variation. The average unconditional shortfall was approximately 10 percent in Bogotá, compared with more than 40 percent in Sucre, Córdoba, and Nariño. This dispersion is the cross-sectional variation used in the design. The exposure measure is not an arbitrary regional index. It summarizes how much underpayment was present before the national reform.

High-exposure departments also differed from low-exposure departments along observable labor-market dimensions. Appendix Table 6 reports these pre-reform differences. Non-compliance affected approximately 74 percent of workers in high-exposure departments, compared with 48 percent in low-exposure departments. Social-security coverage was lower, and agriculture accounted for a larger share of employment, in high-exposure departments. These differences are not surprising: pre-reform exposure is intended to capture where enforcement had more potential bite. But they also motivate the auxiliary outcomes in the reduced-form analysis. If the post-reform estimates simply reflect broader regional labor-market shifts, those shifts should appear outside the minimum-wage shortfall outcomes.

The descriptive figures and tables in Appendix D show the national evolution of non-compliance, the pre-reform wage distribution, and the regional variation in exposure. The wage distribution displays both a mass near the statutory minimum and substantial employment below it. This pattern is important for the structural model: there are workers at the legal floor, workers below it, and workers above it. The model therefore needs both a compliance margin and a non-compliance margin. These descriptive facts are not causal evidence. They show that the objects in the theory—incidence, depth, and the mass around the minimum wage—are empirically relevant in Colombia.

5 Reduced-Form Evidence on the Two Margins

5.1 Effects on non-compliance margins

The reduced-form evidence is organized around the two margins highlighted by the model. Incidence, I, measures whether a worker is paid below the statutory minimum. Depth, G, measures how far below the minimum an underpaid worker is paid. A detection shock should operate mainly through I, by clearing violations from the stock of employment. A penalty-severity shock should operate mainly through G, by making deeper underpayment more expensive.

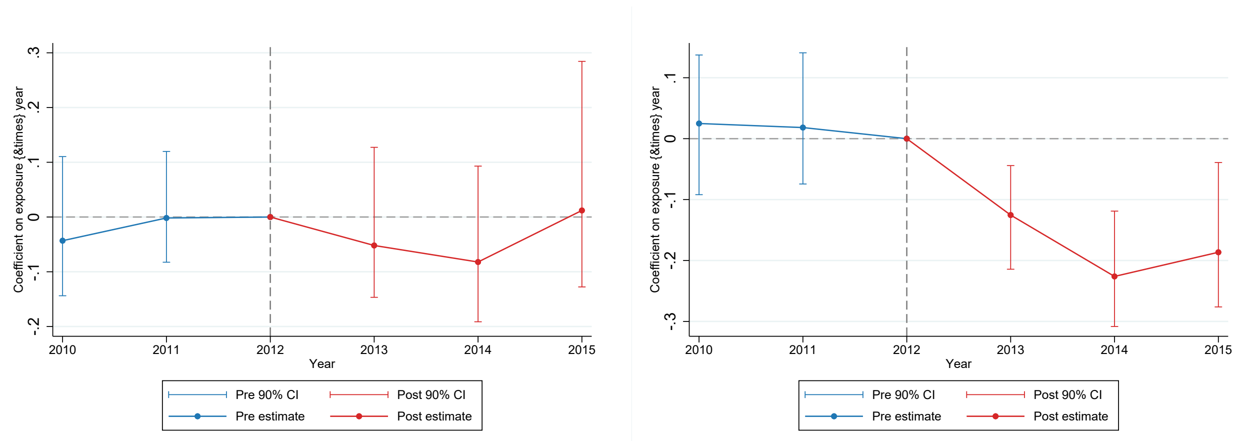
Figure 1 reports the event-study estimates for these two margins. Panel a shows the incidence of subminimum wages, I. Panel b shows the conditional shortfall among workers paid below the statutory floor, G. The omitted year is 2012, immediately before Law 1610. Coefficients measure how outcomes evolved in departments with greater pre-reform underpayment relative to less exposed departments.

The incidence estimates in Figure 1 do not show a clear post-reform decline. The coefficients are negative in 2013 and 2014, but the confidence intervals include zero. By 2015, the point estimate is close to zero. The event study therefore does not reveal the large extensive-margin response that the model associates with a strong relative increase in detection.

The depth response in Figure 1 is different. The conditional shortfall falls sharply after the reform. The coefficient is -0.125 in 2013, reaches -0.226 in 2014, and remains negative at -0.186 in 2015. The corresponding pre-reform coefficients for 2010 and 2011 are small and statistically indistinguishable from zero. Departments with greater pre-reform underpayment were therefore not already experiencing faster declines in the depth of violations before Law 1610.

The unconditional shortfall, $I \times G$, combines the two margins. Appendix Figure 7 reports the corresponding estimates. The coefficients are -0.111 in 2013, -0.208 in 2014, and -0.179 in 2015. The decline in total underpayment is therefore driven primarily by the decline in G, not by a clear decline in I.

The reduced-form pattern shows how the reform operated in our data. High-exposure departments experience a large decline in how much underpaid workers are underpaid, but no comparable decline in the share of workers paid below the legal floor. In the language of the model, the response looks more like an increase in the price of deep violations than a large relative increase in the clearance rate of violations. This is why the structural bridge in Section 7 uses the decline in G to discipline the change in penalty severity.



(a) Incidence of non-compliance

(b) Conditional shortfall

Figure 1: Effects of Law 1610 on minimum-wage non-compliance

Notes: The figure reports event-study coefficients from equation (27). Exposure is the department-level unconditional minimum-wage shortfall averaged over 2010 and 2011. The omitted year is 2012. The sample is private-sector paid workers aged 15–64 with at most secondary education. Confidence intervals are 90 percent wild-cluster-bootstrap intervals with Webb weights.

5.2 Other labor-market outcomes

The interpretation of the non-compliance results depends on whether they reflect enforcement of the minimum wage or broader labor-market changes. This concern is natural in Colombia because Law

1607 reduced non-wage labor costs around the same time, and because labor-inspection capacity was expanding during the period. If high-exposure departments experienced a broad labor-market shift unrelated to minimum-wage compliance, we should see it in outcomes beyond wage shortfalls.

We therefore estimate the same event-study specification for social-security coverage, contract status, usual weekly hours, log hourly wages, and broad sectoral allocation. These outcomes are not the main objects of the paper. They help discipline interpretation.

The full set of auxiliary event studies is reported in Appendix E.2. The appendix includes social-security coverage, contract status, log hourly wages, weekly hours, and employment shares in agriculture, industry, and services.

The auxiliary outcomes in Appendix E.2 do not show the same sharp pattern as the conditional shortfall. Social-security coverage does not display a clear post-reform increase in more-exposed departments. Contract status and hours also do not exhibit a systematic differential response. Sectoral allocation across agriculture, industry, and services remains broadly stable. Log hourly wages rise in more-exposed departments after the reform, which is consistent with the decline in wage shortfalls, but we interpret this as supporting evidence rather than as a separate mechanism.

These auxiliary results in Appendix E.2 do not rule out all contemporaneous changes in the Colombian labor market. They show that the main differential response across high- and low-exposure departments is concentrated on minimum-wage shortfalls, not on broad employment, formality, hours, or sectoral reallocation margins. This concentration is what we need for the structural interpretation.

5.3 Robustness and interpretation

We conduct three robustness and falsification exercises, reported in Appendix E.3. First, we reconstruct exposure using the pre-reform incidence of subminimum wages rather than the unconditional shortfall. This addresses the concern that our baseline exposure measure mechanically loads on the depth of pre-reform underpayment. The results remain similar. Conditional and unconditional shortfalls decline after the reform, while incidence estimates remain imprecise.

Second, we repeat the event-study analysis for workers with tertiary education. These workers are less likely to be directly constrained by the statutory minimum wage, so they provide a natural falsification group. Appendix Table 10 does not show a comparable post-reform decline in the unconditional shortfall for this group. This supports the interpretation that the baseline estimates operate through a minimum-wage compliance channel.

Third, we extend the window to 2008–2018. Appendix Table 11 shows that the decline in wage shortfalls persists beyond the baseline period. At the same time, the additional years introduce more institutional and measurement concerns. We therefore keep 2010–2015 as the baseline and treat the longer window as supporting evidence.

Appendix E.3 reports the alternative-exposure specification, the tertiary-education falsification, and the extended-window estimates. Appendix Table 8 reports the numerical coefficients for the baseline non-compliance event studies, and Appendix Figure 7 reports the unconditional-shortfall event study.

The reduced-form evidence in Figure 1, Appendix Figure 7, and Appendix E.2 points to a depth response. Departments with greater pre-reform underpayment experienced larger declines in the conditional and unconditional shortfalls after Law 1610. The incidence response is statistically imprecise and does not display a comparable post-reform break. Other labor-market outcomes do not show a similarly sharp differential shift. The results therefore give the empirical object that the structural model must explain: a policy environment that reduces total underpayment mainly by reducing the depth of violations.

6 Structural Estimation

6.1 Baseline economy and parameterization

Section 2 presents the theoretical framework. We estimate the pre-reform economy using the 2012 cross-section, the year immediately before Law 1610. The estimation sample mirrors the reduced-form population as closely as possible. For employed workers, we use private-sector paid workers aged 15–64 with at most secondary education. For the unemployment moment, we add unemployed individuals who satisfy the same age and education restrictions and remain in the labor force. This keeps the wage moments tied to the population most exposed to minimum-wage enforcement while allowing the model to discipline labor-market stocks.

Wages are normalized by the statutory minimum wage, so the minimum wage is equal to one in the estimated model. We use the same 0.5 percent tolerance band as in the reduced-form analysis. Workers below 99.5 percent of the minimum wage enter the subminimum region. Workers within the tolerance band around the minimum wage enter the minimum-wage mass. Workers above 100.5 percent of the minimum wage enter the clearly above-minimum region used for wage-distribution moments.

The productivity distribution is lognormal, with parameters μ_z and σ_z. The structural parameter vector is estimated by minimum distance:

$$\Theta = (\mu_z, \sigma_z, rU, \chi, \pi, \rho). \tag{28}$$

The first two parameters govern the productivity distribution. The parameter rU is the flow-equivalent value of unemployment that enters the Nash bargain and the acceptance thresholds. Following Flinn (2006), we estimate this equilibrium reservation value directly rather than separately estimating the primitive nonemployment payoff b. The unemployment Bellman equation in Section 2 shows that b and the option value of future contacts enter wage setting only through the value of unemployment. The enforcement parameters are the expected penalty wedge χ and the detection probability π, and ρ is the separation rate. Penalty severity is recovered from

$$\kappa = \frac{\chi}{\pi}. \tag{29}$$

We fix the worker bargaining power at $\eta = 0.5$ and use a quarterly discount rate $r = 0.0125$. The job-contact rate λ is calibrated externally using labor-force transition rates from Donovan et al. (2023). Because Colombia is not available in their data, we use Peru as the benchmark economy. For each candidate parameter vector, the model-implied acceptance probability maps this transition target into a job-contact rate λ. This procedure anchors the job-contact rate to an externally observed transition rate while allowing unemployment and separations to be disciplined by the Colombian cross-section. The next subsection describes the targeted moments and the joint mapping from those moments to the estimated parameters.

6.2 Moments and identification

We estimate the six parameters in Θ by matching six targeted moments:

$$\left(\overline{w}_{>1.005}, \frac{p75(w_{>1.005})}{p25(w_{>1.005})}, u, B, I, G\right), \tag{30}$$

where $\overline{w}_{>1.005}$ denotes the mean wage above the minimum-wage tolerance band, u is the unemployment rate, B is the mass at the minimum wage, I is the incidence of non-compliance, and G is the conditional shortfall among subminimum-wage workers.

The first two moments are computed strictly above the minimum wage and therefore isolate the part of the wage distribution least directly affected by the statutory floor. In this region, wages follow the compliant Nash schedule

$$w_C(z) = (1 - \eta)rU + \eta z.$$

These moments discipline the location and dispersion of the productivity distribution, (μ_z, σ_z), together with the reservation value rU. The reservation value shifts the intercept of the compliant wage schedule and also enters the productivity thresholds (z^*, z^v, z^c), which allocate matches across violation, bunching, and compliance regions.

The unemployment rate disciplines the separation rate ρ, given the externally anchored job-contact rate λ. Intuitively, for a given rate at which unemployed workers receive acceptable job opportunities, the steady-state stock of unemployment is informative about the rate at which jobs separate.

The enforcement moments jointly discipline the two enforcement parameters (χ, π). The conditional shortfall G loads on wage compression: higher expected penalties make deep underpayment more costly and move violating wages toward the statutory minimum. Incidence I loads on violation duration because detection shortens violation spells through

$$e_V = \frac{u\lambda q_V}{\rho + \pi}.$$

A higher detection probability therefore reduces the steady-state stock of violating jobs even when the wage schedule is held fixed.

The minimum-wage mass B loads on conversion and bunching. The expected penalty wedge determines the size of the interval $[z^v, z^c)$ in which matches pay the statutory minimum rather than a subminimum wage. Detection also affects this mass because high-productivity violations that are detected are converted into compliant minimum-wage jobs. The joint behavior of (G, I, B) therefore separates wage compression, violation duration, and conversion into compliance, allowing the estimation to distinguish the expected penalty wedge from the detection probability.

Taken together, these six moments provide the identifying variation used to recover the six parameters in Θ. The mapping is joint rather than one-to-one: wage moments above the minimum wage discipline productivity and rU; the unemployment rate disciplines separations conditional on the external job-contact benchmark; and the enforcement moments (G, I, B) jointly discipline the expected penalty wedge and the detection probability.

We leave lower-tail wage moments among subminimum workers and overall wage moments as validation objects. These untargeted moments evaluate the model's ability to reproduce dimensions of the wage distribution that were not directly used in estimation.

Model moments are computed using deterministic quadrature, and the minimum-distance objective employs a diagonal weighting matrix based on bootstrap estimates of the sampling variance of the targeted moments. To mitigate convergence to local minima, estimation combines a warm start with multiple random initializations. Appendix F provides additional computational details.

6.3 Estimates and model fit

Table 1 reports the estimated parameters, together with the recovered objects used to characterize the pre-reform economy. The estimated productivity distribution has location $\mu_z = 0.509$ and dispersion $\sigma_z = 0.472$. The flow value of unemployment entering the bargain is $rU = 0.021$. The quarterly separation rate is $\rho = 0.087$, while the job-contact rate recovered from the external transition target is $\lambda = 0.319$.

The enforcement parameters point to a weak pre-reform enforcement environment. The estimated quarterly detection probability is $\pi = 0.018$, implying that violating jobs face a low probability of inspection within a quarter. The expected enforcement wedge is $\chi = 0.029$, which corresponds to a recovered penalty-severity parameter of $\kappa = 1.639$. Thus, although the statutory minimum wage is binding in the model, the pre-reform combination of detection risk and penalty severity leaves substantial room for subminimum payment in equilibrium.

Table 2 reports the fit of the estimated economy. The estimation targets the above-minimum wage distribution, unemployment, and the main non-compliance moments. Specifically, the model matches the mean wage above the minimum, above-minimum wage dispersion, the unemployment rate, the mass at the minimum wage, the incidence of non-compliance, and the conditional shortfall. Because the specification uses six targeted moments for six estimated parameters, the targeted fit should be read as a discipline on the estimated economy rather than as an independent validation exercise.

Table 1: Pre-reform structural parameters

Parameter	Description	Value
μ_z	Log-productivity location	0.509
σ_z	Log-productivity dispersion	0.472
rU	Flow value of unemployment	0.021
χ	Expected enforcement wedge	0.029
π	Quarterly detection rate	0.018
ρ	Quarterly separation rate	0.087
λ	Quarterly job-contact rate	0.319
κ	Penalty severity, recovered	1.639

Notes: Wages are normalized by the statutory minimum wage, so $\omega_M = 1$ in estimation. The expected penalty term is $\chi(1-w)[1+(1-w)]$, with $\chi = \pi\kappa$. The table reports the estimated structural parameters, the job-contact rate recovered from the external unemployment-to-employment transition target, and the implied penalty severity. Fixed parameters are described in the text and kept in the replication output.

The table also reports two aggregate objects implied by targeted moments. The unconditional gap among employed workers, $I \times G$, equals 0.194 in the data and 0.193 in the model. This number means that, among employed workers, the average wage shortfall generated by non-compliance amounts to about 19 percent of the minimum wage when workers paid at or above the minimum are assigned a zero shortfall. Aggregate non-compliance, $(1-u)IG$, equals 0.152 in the data and 0.151 in the model. This scales the employed-worker gap by the employment rate and therefore measures the average shortfall in the full labor-force population. These objects are central for the counterfactual analysis. They are implied by the moments u, I, and G.

The untargeted moments show that the model also fits the subminimum wage distribution reasonably well. It reproduces the mean wage below the minimum almost exactly and comes close to the lower-tail percentile and subminimum dispersion ratio. The main miss is the variance of subminimum wages, which is lower in the model than in the data. The overall wage distribution is also well captured: the model matches the mean wage among employed workers closely and only modestly understates overall wage variance. This pattern is reassuring for the counterfactual exercise. The low-dimensional structure captures the central non-compliance margins and the broad wage distribution.

The estimated baseline in Tables 1 and 2 provides the two ingredients needed for the counterfactual analysis. First, it reproduces the pre-reform distribution of wages above, at, and below the minimum along the moments that discipline non-compliance. Second, it separates detection from penalty severity through the incidence, bunching, and shortfall margins. We can therefore compare policies that reduce underpayment through different instruments: increasing κ, which operates mainly through the depth of violations, and increasing π, which operates through detection, conversion, and the stock of violating jobs.

Table 2: Pre-reform model fit

Moment	Moment value	
	Data	Model
Panel A. Targeted moments		
Mean wage above minimum, $\overline{w}_{>1.005}$	1.427	1.426
Above-minimum wage dispersion, $p75(w_{>1.005})/p25(w_{>1.005})$	1.398	1.404
Unemployment rate, u	0.217	0.217
Mass at the minimum wage, B	0.097	0.097
Incidence of non-compliance, I	0.539	0.538
Conditional shortfall, G	0.359	0.359
Panel B. Moments implied by targeted moments		
Unconditional gap among employed workers, $I \times G$	0.194	0.193
Aggregate non-compliance, $(1-u)IG$	0.152	0.151
Panel C. Untargeted validation moments		
Squared conditional shortfall, G_2	0.185	0.165
25th percentile below minimum, $p25(w_{<0.995})$	0.459	0.483
Mean wage below minimum, $\overline{w}_{<0.995}$	0.641	0.641
Wage variance below minimum, $\mathrm{Var}(w_{<0.995})$	0.056	0.037
Subminimum wage dispersion, $p90(w_{<0.995})/p50(w_{<0.995})$	1.389	1.456
Mean wage among employed workers, $\overline{w}$	0.961	0.962
Wage variance among employed workers, $\mathrm{Var}(w)$	0.254	0.225

Notes: The table reports empirical and model-implied moments for the pre-reform baseline estimated by minimum distance with deterministic quadrature. Wages are normalized by the statutory minimum wage. Panel A lists the targeted moments. Panel B reports policy-relevant objects mechanically implied by targeted moments, including the targeted unemployment rate u. Panel C reports moments that are not targeted in estimation.

7 Effectiveness and the Anatomy of Enforcement

7.1 A structural bridge for Law 1610

The reduced-form evidence gives a clear empirical pattern for Law 1610. In departments with higher pre-reform exposure, the conditional shortfall falls sharply after the reform, while the incidence response is statistically imprecise. The model lets us translate this depth response into a change in enforcement primitives. We use the first post-reform year, 2013, for the bridge. This choice keeps the mapping close to the timing of the law and lets us compare the model-implied response of the other outcomes to the reduced-form estimates for the same year.

The bridge holds the baseline detection probability fixed and chooses the post-reform penalty severity that reproduces the 2013 reduced-form decline in the conditional shortfall. Formally, we keep π at its estimated pre-reform value and choose κ^{post} so that the model-implied change in G equals the reduced-form effect on G, scaled by the mean pre-reform exposure. The reduced-form coefficient for the conditional shortfall in 2013 is -0.125. At the mean exposure, this corresponds to a decline in G of 2.5 percentage points.

This exercise is not a claim that the reform changed only penalties. Section 3 documents that Law 1610 occurred in a broader strengthening of enforcement capacity. The bridge instead asks a narrower question: what penalty-severity change would rationalize the observed first-year decline in the depth of violations in the estimated economy? The answer is informative because the model distinguishes between the two enforcement instruments. Holding π fixed, the bridge isolates the penalty component of the reform response.

The bridge increases κ from 1.64 to 12.91, a multiplier of 7.9. The implied change is large because the pre-reform enforcement environment is weak. The multiplier should therefore not be read as the statutory increase in fines. The more useful object is the post-reform penalty severity needed to make the model reproduce the observed decline in G. This post-reform value generates a decline in the conditional shortfall of 2.5 percentage points, by construction.

The validation exercise then asks what else the same change in κ predicts. Table 4 compares model-implied changes to the first-year reduced-form estimates. The model predicts an increase in mean log wages of 4.6 percent. The reduced-form estimate for 2013, scaled by mean exposure, is 4.5 percent. This is a useful validation: the wage response is not targeted by the bridge, yet the model reproduces it almost exactly. The model also predicts that the unconditional shortfall among employed workers falls by 3.8 percentage points, compared with a reduced-form estimate of 2.2 percentage points. For incidence, the model predicts a decline of 7.5 percentage points, while the reduced-form estimate is a decline of 1.0 percentage point.

These comparisons are informative about the scope and limits of the bridge. The model matches the depth response by construction and performs well on the wage response. It is more aggressive on the extensive margin than the reduced form. In the first post-reform year, the reduced-form 90 percent confidence interval for the scaled incidence effect is approximately [-2.9, 2.5] percentage points; using a conventional 95 percent interval based on the clustered standard

Table 3: Structural bridge for Law 1610

	Mean-exposure bridge
Panel A. Targeted bridge	
Target change in conditional shortfall, ΔG	-0.025
Panel B. Enforcement instruments	
Baseline detection rate, π	0.018
Post-reform detection rate, π'	0.018
Baseline penalty severity, κ	1.639
Post-reform penalty severity, κ'	12.912
Penalty multiplier, κ'/κ	7.879
Panel C. Model-implied changes	
Change in incidence, ΔI	-0.075
Change in conditional shortfall, ΔG	-0.025
Change in aggregate underpayment, $\Delta(I \times G)$	-0.038
Change in mean log wage	0.046
Change in unemployment, Δu	0.001

Notes: The table reports the structural bridge from the 2013 reduced-form estimate to the model. The target in Panel A is the reduced-form change in the conditional shortfall, scaled by the weighted mean of pre-reform departmental exposure. Panel B reports the enforcement instrument change required to match this target: the exercise holds detection risk π fixed and increases penalty severity κ. Panel C reports the resulting model-implied changes in non-compliance, wages, and unemployment.

Table 4: First-year bridge validation

Moment	Reduced form	90% CI	Model
Change in incidence, ΔI	-0.010	[-0.029, 0.025]	-0.075
Change in conditional shortfall, ΔG	-0.025	[-0.042, -0.009]	-0.025
Change in aggregate underpayment, $\Delta(I \times G)$	-0.022	[-0.035, -0.009]	-0.038
Change in mean log wage, $\Delta\overline{\log w}$	0.045	[0.010, 0.094]	0.046

Notes: Reduced-form effects are 2013 event-study coefficients scaled by the weighted mean of pre-reform departmental exposure. Confidence intervals are the corresponding 90 percent intervals from the reduced-form estimates. The bridge targets the conditional shortfall, ΔG; the remaining rows are validation moments.

error gives approximately [-3.3, 1.3] percentage points. The model-implied incidence response lies below these intervals. This gap matters. It says that a pure penalty-severity interpretation of the first-year decline in G generates more job movement out of violation than the reduced-form incidence estimates suggest. We therefore treat the bridge as a disciplined mapping from the depth response into the model, not as a full causal accounting of all margins of Law 1610.

7.2 Instrument responses

We next use the estimated model to compare the two enforcement instruments directly. In the first experiment, we raise penalty severity κ and hold detection probability π fixed. In the second, we raise π and hold κ fixed. The exercise is not a fiscal-cost comparison. A given proportional increase in detection and a given proportional increase in penalties need not require the same administrative resources. The purpose is to isolate the economic mechanism of each instrument.

Figure 2 reports the responses. The horizontal axis is the policy multiplier relative to baseline.

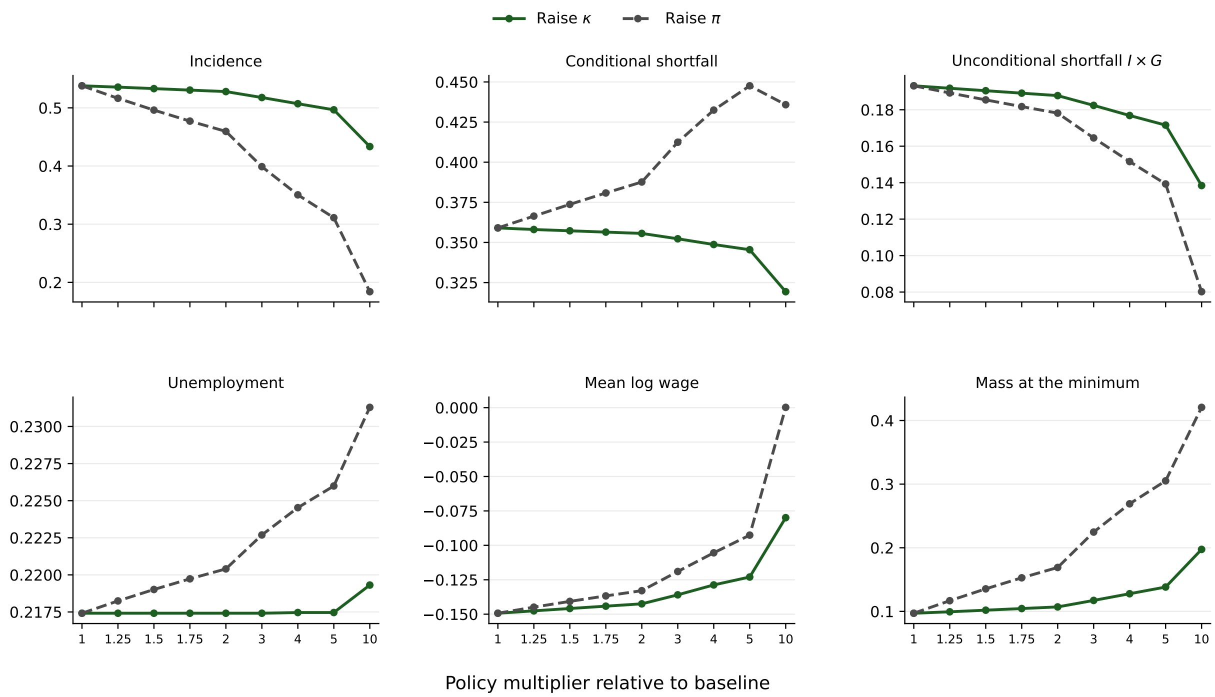


Figure 2: Responses to proportional increases in penalties and detection

Notes: The figure compares counterfactual proportional increases in penalty severity κ and the detection rate π. In each experiment, the other enforcement instrument is held fixed at its estimated baseline value. The horizontal axis reports the multiplier relative to baseline. Dashed horizontal lines mark the baseline outcome level. Only feasible multipliers are shown.

The responses show the non-equivalence result in the estimated economy. Penalty severity operates mainly through the intensive margin. Higher κ raises the marginal cost of deep underpayment, so firms that continue to violate choose wages closer to the legal floor. The conditional shortfall falls, mean wages rise, and the incidence of violations moves less over moderate policy changes. Penalties therefore reduce how far below the law underpaid workers are paid.

Detection operates differently. A higher π shortens the expected duration of a violating job. Some detected violations convert into compliant minimum-wage jobs. Other detected violations are destroyed because the match cannot survive at the legal wage. Incidence therefore falls sharply when detection increases. This is the clearance channel. Detection is not only a price in the expected penalty term; it is also the rate at which violations leave the violating state.

The conditional shortfall can move in the opposite direction under detection. When π rises, many violations disappear from the stock. The remaining violators are negatively selected toward low-productivity matches, for which paying below the minimum is still privately valuable. The average shortfall among them can therefore rise even while total underpayment falls. This is why the same aggregate objective can be achieved through very different distributions of outcomes.

The wage response follows the same logic. Stronger penalties raise wages among jobs that remain below the minimum because deep underpayment becomes more expensive. Stronger detection raises average wages by clearing violations and converting some jobs into compliance. But among the remaining violators, wages need not rise as much, because the group that remains below the minimum is negatively selected. Appendix Figure 15 reports these wage responses separately.

7.3 Equal-benefit comparison

The instrument-response exercise compares proportional policy changes. We now compare policies that deliver the same reduction in aggregate underpayment. The target is the first-year reduced-form decline in the unconditional shortfall among employed workers, $I \times G$, scaled by mean exposure. In 2013, this target is a decline of 2.2 percentage points.

We compute two counterfactuals. The first raises κ, holding π fixed. The second raises π, holding κ fixed. Each instrument is adjusted until the model generates the same decline in $I \times G$. The two policies therefore deliver the same aggregate reduction in underpayment, but they do so through different margins.

The penalty policy requires κ to increase by a factor of 5.1. It reduces incidence by 4.2 percentage points and reduces the conditional shortfall by 1.4 percentage points. The unemployment response is essentially zero at the reported precision. The detection policy requires π to increase by a factor of 2.5. It reduces incidence by 11.0 percentage points, raises the conditional shortfall among the remaining violators by 4.1 percentage points, and increases unemployment by 0.4 percentage point.

The comparison is the main policy message of the model. Penalties and detection can be calibrated to deliver the same decline in aggregate underpayment, but they do not deliver the same labor market. Penalties compress the depth of violations. Detection clears the stock of violations. The extensive-margin effect of detection is larger because detection changes the transition rates out of violation. The employment effect is also larger because some low-productivity violations cannot be converted into compliant jobs.

A simple decomposition makes the difference transparent. The change in $I \times G$ can be written

Table 5: Equal-benefit enforcement counterfactuals

	Increase penalties	Increase detection
Panel A. Common target		
Target change in aggregate underpayment, $\Delta(I \times G)$	-0.022	-0.022
Panel B. Enforcement instruments		
Post-reform detection rate, π'	0.018	0.044
Detection multiplier, π'/π	1.000	2.495
Post-reform penalty severity, κ'	8.283	1.639
Penalty multiplier, κ'/κ	5.055	1.000
Panel C. Model-implied changes		
Change in incidence, ΔI	-0.042	-0.110
Change in conditional shortfall, ΔG	-0.014	0.041
Change in aggregate underpayment, $\Delta(I \times G)$	-0.022	-0.022
Change in aggregate shortfall, ΔNC	-0.017	-0.018
Change in unemployment, Δu	0.000	0.004
Change in mean log wage, $\Delta\overline{\log w}$	0.027	0.023

Notes: Both counterfactuals are calibrated to deliver the same 2013 mean-exposure reduced-form decline in aggregate underpayment, $I \times G$. In the penalty counterfactual, detection risk π is held fixed and penalty severity κ increases. In the detection counterfactual, penalty severity κ is held fixed and detection risk π increases. Panel C reports the resulting model-implied changes in non-compliance, wages, and unemployment.

as the sum of an incidence component and a conditional-shortfall component:

$$\Delta(IG) = \frac{G_0 + G_1}{2}\Delta I + \frac{I_0 + I_1}{2}\Delta G. \tag{31}$$

Under the penalty policy, both components reduce underpayment, with the incidence component the larger of the two. Under the detection policy, the entire net reduction comes from the incidence component, while the conditional-shortfall component works in the opposite direction: clearing the shallow, high-productivity violators into compliance leaves a negatively selected pool of deeper violations. This decomposition is the empirical counterpart of the theory in Section 2: penalties change the price of a violation, while detection changes both the price and the survival of the violation.

These counterfactuals do not rank the instruments by cost-effectiveness. We do not observe the administrative cost of expanding inspection capacity, and we do not estimate the legal, fiscal, or political-economy costs of raising penalties. A full welfare comparison would also need to value worker gains, firm losses, fiscal costs, and employment effects. The result here is narrower and cleaner. Enforcement instruments are not interchangeable. A government that raises penalties is mainly choosing to reduce the depth of underpayment. A government that raises detection is mainly choosing to reduce the stock of violations, with a larger employment margin.

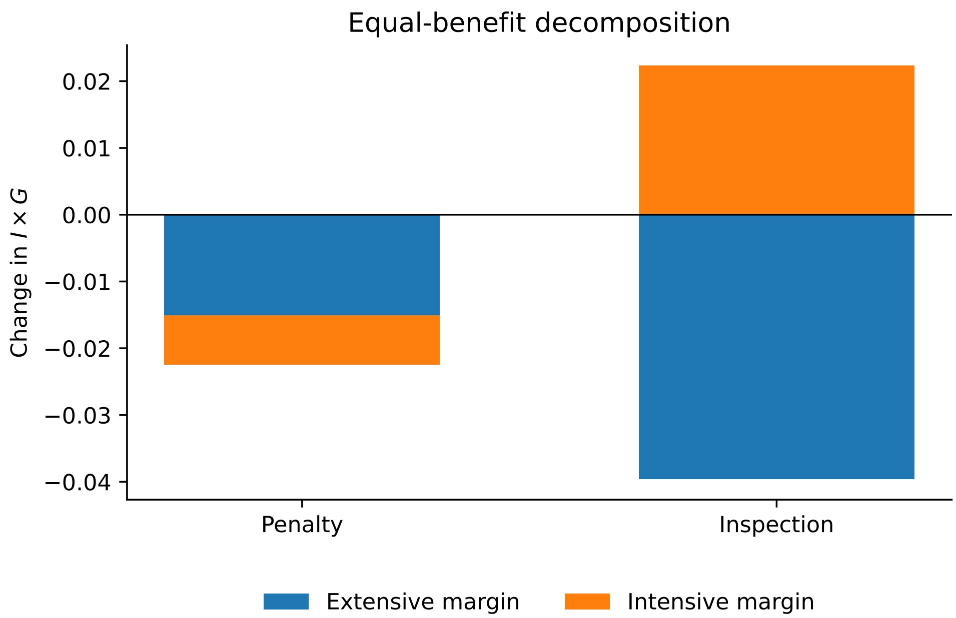


Figure 3: Decomposition of equal-benefit counterfactuals

Notes: The figure decomposes the change in the unconditional shortfall $I \times G$ into the contribution of the incidence margin and the contribution of the conditional-shortfall margin. Both policies are calibrated to deliver the same 2013 mean-exposure decline in $I \times G$.

8 Conclusion

In this paper, we study whether raising the probability of detection and raising the penalty conditional on detection are interchangeable ways to reduce minimum-wage non-compliance, and we conclude that they are not. The standard expected-fine logic treats the two as substitutes through their product, but that equivalence holds only for the static price of a single violation. Once a violation is a durable labor-market state rather than a one-time act, the two instruments part ways: a higher penalty raises the price of underpayment and compresses how far below the floor wages fall, while a higher detection probability does that and more, because it also clears the existing stock of violations by converting some underpaid jobs to compliance and ending others. Detection prices violations and clears them; penalties only price them.

Colombia's 2013 Law 1610 supplies the test, and the paper's three pieces line up behind the same conclusion. The reduced-form evidence shows that the depth of underpayment fell in more exposed departments after the reform, by 2.5 percentage points of the conditional shortfall at mean exposure, while the incidence of violations shows no comparable break. A search model with imperfect enforcement, estimated on the pre-reform cross-section, maps this depth response into a weak baseline enforcement environment and matches the wage response closely, while also making clear the limits of a pure penalty-severity interpretation on the incidence margin. The equal-benefit counterfactuals then make the divergence concrete: holding the two instruments to the same aggregate cut in underpayment, the penalty policy lowers both the number and the depth of violations without moving employment, while the detection policy reaches the same target by cutting the incidence of violations by 11 percentage points, leaving a negatively selected pool of deeper violations, and raising unemployment. Two instruments, the same headline number, two different labor markets.

The lesson for enforcement design is not that one instrument dominates, but that each one moves a different margin, so the choice should follow the margin a government wants to change. A regulator who wants to reduce the number of workers paid below the legal floor needs detection capacity, because only detection changes the rate at which jobs leave the violating state. A regulator who wants to reduce how far below the floor those workers are paid can move that margin through credible penalties, without expanding inspection. Equal expected fines are not equal policy, and writing a larger fine into law is not a substitute for the capacity to find violations.

We establish the non-equivalence of enforcement instruments, not a ranking of them. We do not price the administrative, fiscal, or political cost of raising detection against the cost of raising penalties, and supplying that cost side is the natural next step, the one that would turn non-equivalence into a cost-effectiveness frontier for enforcement policy. That step is worth taking because the underlying result is not specific to Colombia or to the minimum wage. Wherever a violation is a state that a firm occupies rather than an act it commits once, the question of how to enforce the law cannot be reduced to the size of the expected fine.

References

Almeida, Rita and Pedro Carneiro, "Enforcement of Labor Regulation and Informality," *American Economic Journal: Applied Economics*, 2012, *4* (3), 64–89.

Ashenfelter, Orley and Robert S. Smith, "Compliance with the Minimum Wage Law," *Journal of Political Economy*, 1979, *87* (2), 333–350.

Badaoui, Eliane and Frank Walsh, "Productivity, Non-Compliance and the Minimum Wage," *Journal of Development Economics*, 2022, *155*, 102778.

Basu, Arnab K., Nancy H. Chau, and Ravi Kanbur, "Turning a Blind Eye: Costly Enforcement, Credible Commitment and Minimum Wage Laws," *The Economic Journal*, 2010, *120* (543), 244–269.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 1968, *76* (2), 169–217.

Bhorat, Haroon, Ravi Kanbur, and Natasha Mayet, "Minimum Wage Violation in South Africa," *International Labour Review*, 2012, *151* (3), 277–287.

Clemens, Jeffrey and Michael R. Strain, "Understanding "Wage Theft": Evasion and Avoidance Responses to Minimum Wage Increases," *Labour Economics*, 2022, *79*, 102285.

Congreso de Colombia, "Ley 278 de 1996: Comisión Permanente de Concertación de Políticas Salariales y Laborales," Departamento Administrativo de la Función Pública, Gestor Normativo 1996. Accessed May 15, 2026.

_ , "Ley 1607 de 2012: Reforma tributaria," Departamento Administrativo de la Función Pública, Gestor Normativo 2012. Accessed May 15, 2026.

_ , "Ley 1610 de 2013: Inspecciones del trabajo y acuerdos de formalización laboral," Departamento Administrativo de la Función Pública, Gestor Normativo 2013. Accessed May 15, 2026.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman, "Labor Market Dynamics and Development," *Quarterly Journal of Economics*, 2023, *138* (4), 2287–2325.

Flinn, Christopher J., "Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates," *Econometrica*, 2006, *74* (4), 1013–1062.

Garnero, Andrea and Claudio Lucifora, "Turning a "Blind Eye"? Compliance with Minimum Wage Standards and Employment," *Economica*, 2022, *89* (356), 884–907.

Ham, Andrés, "Minimum Wage Violations in Honduras," *IZA Journal of Labor & Development*, 2015, *4* (1), 1–19.

International Labour Organization, "Labour Inspection Convention, 1947 (No. 81)," International Labour Organization, Geneva 1947. Accessed June 5, 2026.

Kugler, Adriana D., Maurice D. Kugler, and Luis O. Herrera-Prada, "Do Payroll Tax Breaks Stimulate Formality? Evidence from Colombia's Reform," *Economia*, 2017, *18* (1), 3–40.

_ , _ , and Luis Omar Herrera Prada, "Switching from Payroll Taxes to Corporate Income Taxes: Firms' Employment and Wages after the 2012 Colombian Tax Reform," *Economia*, 2017, *18* (1), 41–74.

Marshall, Samuel, "Minimum Wage Effects without Enforcement," 2025. Working paper, University of Warwick.

Meghir, Costas, Renata Narita, and Jean-Marc Robin, "Wages and Informality in Developing Countries," *American Economic Review*, 2015, *105* (4), 1509–1546.

Ministerio de la Protección Social, "Código Sustantivo del Trabajo, authorized 2011 version," Government compilation 2011. Article 486; accessed May 15, 2026.

Mookherjee, Dilip and I. P. L. Png, "Marginal Deterrence in Enforcement of Law," *Journal of Political Economy*, 1994, *102* (5), 1039–1066.

Morales, Leonardo Fabio and Carlos Medina, "Assessing the Effect of Payroll Taxes on Formal Employment: The Case of the 2012 Tax Reform in Colombia," *Economia*, 2017, *18* (1), 75–124.

Pignatti, Clemente, "Labor-Law Compliance in an Informal Labor Market," *The World Bank Economic Review*, 2026, p. lhag019. Published online May 22, 2026.

Polinsky, A. Mitchell and Steven Shavell, "The Optimal Tradeoff between the Probability and Magnitude of Fines," *American Economic Review*, 1979, *69* (5), 880–891.

_ and _ , "The Economic Theory of Public Enforcement of Law," *Journal of Economic Literature*, 2000, *38* (1), 45–76.

Rani, Uma, Patrick Belser, Martin Oelz, and Setareh Ranjbar, "Minimum Wage Coverage and Compliance in Developing Countries," *International Labour Review*, 2013, *152* (3–4), 381–410.

Ronconi, Lucas, "Enforcement and Compliance with Labor Regulations in Argentina," *ILR Review*, 2010, *63* (4), 719–736.

Soundararajan, Vidhya, "Heterogeneous Effects of Imperfectly Enforced Minimum Wages in Low-Wage Labor Markets," *Journal of Development Economics*, 2019, *141*, 102351.

Stansbury, Anna, "Incentives to Comply with the Minimum Wage in the United States and the United Kingdom," *ILR Review*, 2024, *78* (1), 190–216.

Ulyssea, Gabriel, "Firms, Informality, and Development: Theory and Evidence from Brazil," *American Economic Review*, 2018, *108* (8), 2015–2047.

A Structural Model Derivations

This appendix derives the wage schedule, the steady-state stocks, and the two results stated in Section 2. We keep the statutory minimum wage ω_M explicit. The estimation later normalizes wages by the statutory minimum and sets $\omega_M = 1$, but the economics of the model is clearer when the legal floor remains visible.

A.1 Nash Bargaining and the Equilibrium Wage Schedule

This appendix derives the wage schedule stated in Proposition 1. Throughout, the penalty paid upon detection is

$$\Phi(\omega_M - w) = \kappa(\omega_M - w)\big[1 + (\omega_M - w)\big],$$

and we define

$$\chi \equiv \pi\kappa, \qquad \Delta(\pi) \equiv \frac{\pi}{r+\rho}.$$

For compactness, when solving the violating wage equations, let

$$x \equiv \omega_M - w$$

denote the wage shortfall.

Compliant matches. Consider first a match that operates in compliance with the statutory minimum wage. The worker and firm surpluses are

$$W_C(z) - U = \frac{w - rU}{r+\rho}, \tag{32}$$

and

$$J_C(z) = \frac{z - w}{r+\rho}. \tag{33}$$

The compliant wage solves

$$w_C(z) \in \arg\max_{w \geq \omega_M} \left(\frac{w - rU}{r+\rho}\right)^{\eta} \left(\frac{z - w}{r+\rho}\right)^{1-\eta}.$$

Equivalently, since the common denominator does not affect the maximizer,

$$w_C(z) \in \arg\max_{w \geq \omega_M} (w - rU)^{\eta}(z - w)^{1-\eta}.$$

The first-order condition is

$$\eta(z - w) = (1-\eta)(w - rU),$$

so the unconstrained compliant Nash wage is

$$w_C^N(z) = (1-\eta)rU + \eta z. \tag{34}$$

This wage is feasible under compliance if and only if $w_C^N(z) \geq \omega_M$, which defines the compliant productivity threshold

$$z^c = \frac{\omega_M - (1-\eta)rU}{\eta}. \tag{35}$$

Low-productivity violating matches. For $z < \omega_M$, detection destroys the match. The worker and firm surpluses while the violation remains undetected are

$$W_V^\ell(z) - U = \frac{w - rU}{r+\rho+\pi}, \qquad J_V^\ell(z) = \frac{z - w - \pi\Phi(\omega_M - w)}{r+\rho+\pi}.$$

The low-productivity violating wage solves

$$w_V^\ell(z) \in \arg\max_{rU \leq w < \omega_M} \left(\frac{w-rU}{r+\rho+\pi}\right)^\eta \left(\frac{z - w - \pi\Phi(\omega_M - w)}{r+\rho+\pi}\right)^{1-\eta}.$$

Equivalently,

$$w_V^\ell(z) \in \arg\max_{rU \leq w < \omega_M} (w-rU)^\eta \left[z - w - \pi\Phi(\omega_M - w)\right]^{1-\eta}.$$

The first-order condition is

$$\eta\left[z - w_V^\ell(z) - \pi\Phi(\omega_M - w_V^\ell(z))\right] = (1-\eta)\left[1 - \pi\Phi'(\omega_M - w_V^\ell(z))\right]\left[w_V^\ell(z) - rU\right]. \tag{36}$$

Using

$$\pi\Phi(\omega_M - w) = \chi(\omega_M - w)[1 + (\omega_M - w)], \qquad \pi\Phi'(\omega_M - w) = \chi[1 + 2(\omega_M - w)],$$

this condition becomes, in shortfall form,

$$\chi(2-\eta)x^2 - [(1-\chi) + 2(1-\eta)\chi(\omega_M - rU)]\,x + (1-\eta)(1-\chi)(\omega_M - rU) - \eta(z - \omega_M) = 0. \tag{37}$$

The economically relevant shortfall is the smaller root,

$$x_V^\ell(z) = \frac{(1-\chi) + 2(1-\eta)\chi(\omega_M - rU) - \sqrt{D_\ell(z)}}{2\chi(2-\eta)}, \tag{38}$$

where

$$D_\ell(z) \equiv [(1-\chi) + 2(1-\eta)\chi(\omega_M - rU)]^2 - 4\chi(2-\eta)\left[(1-\eta)(1-\chi)(\omega_M - rU) - \eta(z - \omega_M)\right].$$

Thus

$$w_V^\ell(z) = \omega_M - x_V^\ell(z).$$

The acceptance threshold is pinned down by the worker's participation constraint in the low-productivity violation region. Since $w_V^\ell(z) = \omega_M - x_V^\ell(z)$, the marginal accepted match satisfies $w_V^\ell(z^*) = rU$, or equivalently $x_V^\ell(z^*) = \omega_M - rU$. Substituting this condition into (37) gives

$$z^* = rU + \chi(\omega_M - rU)\big[1 + (\omega_M - rU)\big]. \tag{39}$$

Thus matches with $z < z^*$ are not formed.

High-productivity violating matches. For $z \geq \omega_M$, detection converts the match into compliance at wage ω_M. While the violation remains undetected, the worker and firm surpluses are

$$W_V^h(z) - U = \frac{w - rU + \Delta(\pi)(\omega_M - rU)}{r + \rho + \pi},$$

and

$$J_V^h(z) = \frac{z - w - \pi\Phi(\omega_M - w) + \Delta(\pi)(z - \omega_M)}{r + \rho + \pi}.$$

The high-productivity violating wage solves

$$w_V^h(z) \in \arg\max_{w<\omega_M} \left[w - rU + \Delta(\pi)(\omega_M - rU)\right]^\eta \left[z - w - \pi\Phi(\omega_M - w) + \Delta(\pi)(z - \omega_M)\right]^{1-\eta}.$$

The first-order condition is

$$\begin{aligned} &\eta \Big[z - w_V^h(z) - \pi\Phi\big(\omega_M - w_V^h(z)\big) + \Delta(\pi)(z - \omega_M)\Big] \\ &\quad = (1 - \eta) \Big[1 - \pi\Phi'\big(\omega_M - w_V^h(z)\big)\Big] \Big[w_V^h(z) - rU + \Delta(\pi)(\omega_M - rU)\Big]. \end{aligned} \tag{40}$$

In shortfall form, this condition becomes

$$\begin{aligned} &\chi(2 - \eta)x^2 - \big[(1 - \chi) + 2(1 - \eta)\chi\big[1 + \Delta(\pi)\big](\omega_M - rU)\big]\, x \\ &\quad + \big[1 + \Delta(\pi)\big]\left[(1 - \eta)(1 - \chi)(\omega_M - rU) - \eta(z - \omega_M)\right] = 0. \end{aligned} \tag{41}$$

The economically relevant root is

$$x_V^h(z) = \frac{(1 - \chi) + 2(1 - \eta)\chi\big[1 + \Delta(\pi)\big](\omega_M - rU) - \sqrt{D_h(z)}}{2\chi(2 - \eta)}, \tag{42}$$

where

$$\begin{aligned} D_h(z) \equiv\ &\big[(1 - \chi) + 2(1 - \eta)\chi\big[1 + \Delta(\pi)\big](\omega_M - rU)\big]^2 \\ &- 4\chi(2 - \eta)\big[1 + \Delta(\pi)\big]\left[(1 - \eta)(1 - \chi)(\omega_M - rU) - \eta(z - \omega_M)\right]. \end{aligned}$$

Hence

$$w_V^h(z) = \omega_M - x_V^h(z).$$

For reference, the corresponding linear-penalty expression is

$$w_V^h(z) = (1-\eta)\left[rU - \Delta(\pi)(\omega_M - rU)\right] + \frac{\eta}{1-\chi}\left[z - \chi\omega_M + \Delta(\pi)(z-\omega_M)\right]. \tag{43}$$

Bunching. The high-productivity violating region ends when the shortfall reaches zero. Setting $x = 0$ in the high-productivity first-order condition gives

$$z^v = \omega_M + \frac{(1-\eta)(1-\chi)}{\eta}(\omega_M - rU) = \chi\omega_M + (1-\chi)z^c. \tag{44}$$

Thus high-productivity violations occur on $[\omega_M, z^v)$, while matches in $[z^v, z^c)$ bunch at the statutory minimum.

Equilibrium wage schedule. Combining the three regimes,

$$w^*(z) = \begin{cases} \text{no match}, & z < z^*, \\ \omega_M - x_V^\ell(z), & z \in [z^*, \omega_M), \\ \omega_M - x_V^h(z), & z \in [\omega_M, z^v), \\ \omega_M, & z \in [z^v, z^c), \\ (1-\eta)rU + \eta z, & z \geq z^c. \end{cases} \tag{45}$$

The thresholds have transparent interpretations. The threshold z^* is the acceptance threshold. The threshold z^c is the productivity level at which the unconstrained compliant Nash wage reaches the statutory minimum. The threshold z^v is the productivity level at which the high-productivity violating wage reaches the statutory minimum.

Second-order condition and real roots. The maintained regularity condition

$$\chi\big[1 + 2(\omega_M - rU)\big] < 1$$

ensures that

$$1 - \pi\Phi'(\omega_M - w) > 0$$

for all $w \in [rU, \omega_M]$. Hence the firm's continuation value is decreasing in the wage throughout the relevant violation region, and the second-order condition for the Nash objective is satisfied at the selected interior solutions. Under the same condition, the discriminants $D_\ell(z)$ and $D_h(z)$ are strictly positive on their respective domains. The roots selected above are the economically relevant smaller roots. They imply positive shortfalls,

$$0 < x_V^\ell(z) \leq \omega_M - rU, \qquad 0 < x_V^h(z) < \omega_M - rU,$$

so the implied wages lie in the intended violating regions.

A.2 Proof of Proposition 2: non-equivalence of detection and penalties

We prove the result in the main text by isolating the stock-flow effect of detection. Hold fixed the wage schedule and the measures q_V, q_M, q_H, and q_A. Also hold fixed the expected penalty wedge $\chi = \pi\kappa$. Any remaining difference across regimes comes from the detection rate π.

From the steady-state equations,

$$e_V = \frac{u\lambda q_V}{\rho + \pi}, \tag{46}$$

$$e_M = \frac{u\lambda q_M}{\rho} + \frac{u\lambda \pi q_H}{\rho(\rho + \pi)}, \tag{47}$$

and

$$e_A = \frac{u\lambda q_A}{\rho}. \tag{48}$$

The incidence of non-compliance is

$$I(\pi) = \frac{e_V}{e_V + e_M + e_A}. \tag{49}$$

Substituting the stock equations and cancelling the common term $u\lambda$, we obtain

$$I(\pi) = \frac{q_V/(\rho + \pi)}{q_V/(\rho + \pi) + q_M/\rho + q_A/\rho + \pi q_H/[\rho(\rho + \pi)]}. \tag{50}$$

Multiplying numerator and denominator by $\rho(\rho + \pi)$ gives

$$I(\pi) = \frac{\rho q_V}{\rho q_V + (\rho + \pi)(q_M + q_A) + \pi q_H}. \tag{51}$$

Differentiating with respect to π,

$$\frac{\partial I(\pi)}{\partial \pi} = -\frac{\rho q_V (q_M + q_A + q_H)}{[\rho q_V + (\rho + \pi)(q_M + q_A) + \pi q_H]^2}. \tag{52}$$

This derivative is strictly negative whenever $q_V > 0$ and $q_M + q_A + q_H > 0$. Hence, even holding χ fixed, a higher detection probability reduces the stock share of violations. Detection and penalty severity are not interchangeable because detection clears violation spells.

A.3 Proof of Lemma 1: detection and the depth of high-productivity violations

The lemma isolates the conversion-value channel. Consider the limiting case $\chi = 0$. The high-productivity violating Nash problem becomes

$$\max_{w < \omega_M} \left[w - rU + \Delta(\pi)(\omega_M - rU)\right]^{\eta} \left[z - w + \Delta(\pi)(z - \omega_M)\right]^{1-\eta}. \tag{53}$$

The first-order condition is

$$\eta \left[z - w + \Delta(\pi)(z - \omega_M)\right] = (1 - \eta) \left[w - rU + \Delta(\pi)(\omega_M - rU)\right]. \tag{54}$$

Solving for w gives

$$w_V^h(z) = \eta z + (1 - \eta)rU + \Delta(\pi) \left[\eta(z - \omega_M) - (1 - \eta)(\omega_M - rU)\right]. \tag{55}$$

The unconstrained compliant wage is

$$w_C(z) = \eta z + (1 - \eta)rU. \tag{56}$$

The productivity threshold z^c satisfies

$$\omega_M = \eta z^c + (1 - \eta)rU. \tag{57}$$

Therefore,

$$z^c = \frac{\omega_M - (1 - \eta)rU}{\eta}. \tag{58}$$

Using this expression,

$$\eta(z - \omega_M) - (1 - \eta)(\omega_M - rU) = \eta(z - z^c). \tag{59}$$

Thus,

$$w_V^h(z) = w_C(z) + \Delta(\pi)\eta(z - z^c). \tag{60}$$

Since $\Delta(\pi) = \pi/(r + \rho)$, we have

$$\frac{\partial w_V^h(z)}{\partial \pi} = \frac{\eta(z - z^c)}{r + \rho}. \tag{61}$$

For high-productivity violations below the unconstrained compliance threshold, $z < z^c$. Hence

$$\frac{\partial w_V^h(z)}{\partial \pi} < 0. \tag{62}$$

A higher detection probability can therefore lower the current wage paid by high-productivity violations that remain below the minimum wage. The reason is the conversion option. Detection no longer sends the worker to unemployment. It moves the worker into a legal minimum-wage job. That continuation value changes the current bargain and can deepen the remaining violations, even as detection reduces the stock of violations.

A.4 Normalization for estimation

The model is written with a general statutory minimum wage ω_M. In estimation, we normalize all wages by the statutory minimum wage and set $\omega_M = 1$. Under this normalization,

$$\Phi(1-w) = \kappa(1-w)[1+(1-w)], \tag{63}$$

and the wage schedule, thresholds, and stock equations above are the equations used by the structural solver. The normalization changes units. It does not change the economic distinction between detection and penalty severity.

B Data Construction and Sample Restrictions

This appendix describes how we construct the samples and outcomes used in the reduced-form and structural analyses.

B.1 Raw data and statutory minimum wage

We use the harmonized Colombia GEIH files from the World Bank Global Labor Database. The replication file first appends the annual Colombia files for 2008–2018. We keep this wider window in the cleaned file so that robustness exercises can be run without rebuilding the raw data. The baseline reduced-form sample uses 2010–2015. The structural estimation uses the 2012 pre-reform cross-section.

We construct a project-specific statutory minimum-wage file from the ILOSTAT nominal minimum-wage series. The Colombia series is identified by `ref_area == "COL"` and the local-currency classification `classif1 == "CUR_TYPE_LCU"`. The source reports the gross nominal monthly minimum wage. We convert it to an hourly minimum wage by dividing the monthly value by 48×4.33. The 48-hour workweek is the relevant statutory workweek during our sample period. The later reduction in Colombia's legal workweek occurs after the period used in the paper and therefore does not affect the baseline construction.

Let $\omega_{M,t}$ denote the hourly statutory minimum wage in year t. We express each worker's hourly wage relative to the minimum wage as $w_{idt}/\omega_{M,t}$, where i indexes workers, d departments, and t years. In the model and structural tables, the statutory minimum wage is normalized to one.

The wage variable in the GEIH/GLD is constructed from reported labor income and usual hours. When the last-payment wage is available, we convert it to a weekly amount using the reported pay period and divide by weekly hours. When that source is missing, we use total labor income as a fallback and convert it to an hourly wage using months worked and hours. Observations with zero reported wages among paid employees are treated as non-paid before constructing the wage measure.

The household survey wage concept is gross labor income. This matters because the statutory minimum wage is gross. We therefore compare the reported wage to the gross statutory minimum wage without applying a tax or contribution adjustment.

B.2 Reduced-form sample

The reduced-form analysis focuses on workers for whom the minimum wage is most likely to bind. The baseline sample consists of paid private-sector workers aged 15–64 with at most secondary education. Formally, we keep workers satisfying four restrictions: paid employment, private-sector job, working age, and at most secondary schooling. We use the same restrictions to construct the main non-compliance outcomes and the department-level exposure measure.

The baseline reduced-form window is 2010–2015. We omit 2012 in the event-study specification because it is the last pre-reform year and serves as the reference period. The exposure measure is constructed before the reform and is fixed at the department level. For department d, exposure is the weighted average unconditional shortfall in 2010 and 2011. Departments that do not appear before 2012 cannot be assigned a pre-reform exposure value. These departments are therefore not used to identify the baseline exposure design.

The cleaned file also contains tertiary-education workers in the private paid sector. We use this group only for falsification and robustness checks because minimum-wage non-compliance is less likely to be binding for them.

B.3 Wage normalization and non-compliance outcomes

The main empirical objects are constructed from the normalized wage

$$w = \frac{\text{hourly wage}}{\text{hourly minimum wage}}.$$

We use a tolerance band of 0.005 around the statutory minimum wage throughout the reduced-form and structural analysis. This common tolerance avoids classifying workers extremely close to the minimum wage as violators because of rounding, reporting, or conversion error.

A worker is classified as non-compliant if

$$w < 0.995.$$

The incidence of non-compliance is the weighted share of workers in the baseline sample whose normalized wage is below 0.995. The unconditional shortfall is

$$(1 - w)\mathbf{1}\{w < 0.995\}.$$

The conditional shortfall is the same wage gap, but computed only among workers classified as non-compliant. The incidence captures the extensive margin of non-compliance, the conditional shortfall captures the depth of violations among affected workers, and the unconditional shortfall combines both margins.

We also define the mass around the minimum wage as

$$\mathbf{1}\{|w-1| \leq 0.005\}.$$

This object is not a main reduced-form outcome. It is used in the structural estimation because the model predicts that enforcement can move some violating jobs into compliance at the minimum wage.

Before constructing the final outcomes, normalized wages are winsorized within year-by-department cells at the unweighted 1st and 99th percentiles. This choice limits the influence of extreme wage reports while preserving the lower tail that is central to minimum-wage non-compliance. The winsorization is unweighted because it is used as a cleaning rule for extreme reported wage ratios rather than as an estimand. After this cleaning step, descriptive statistics, reduced-form outcomes, and structural moments are computed using survey weights. The same winsorization rule is used in the structural baseline sample so that reduced-form and structural objects are computed from comparable wage distributions.

B.4 Structural baseline sample

The structural model is estimated on the 2012 pre-reform baseline. The structural sample keeps the same age and education restrictions as the reduced-form analysis: individuals aged 15–64 with at most secondary education. Among the employed, we keep private-sector paid workers with a valid normalized wage. Among the non-employed, we keep unemployed individuals. Individuals outside the labor force are excluded.

This construction aligns the unemployment stock with the wage-job market in the model. The model studies private paid jobs subject to the statutory minimum wage. The unemployment rate used in estimation is therefore computed over the population composed of private-sector paid workers and unemployed individuals satisfying the same age and education restrictions. Wage-distribution moments are computed only among employed private-sector paid workers.

The structural moments use the same tolerance band as the reduced-form analysis. Wages below 0.995 define subminimum employment. Wages above 1.005 define the part of the wage distribution clearly above the minimum wage. Wages within the interval $[0.995, 1.005]$ define the near-minimum mass.

The preferred minimum-distance target vector contains six moments from three parts of the labor market. Above-minimum moments discipline the productivity distribution away from the statutory constraint. The unemployment rate disciplines the steady-state stock of nonemployment together with externally calibrated transition information. The near-minimum mass B helps discipline detection and conversion because enforcement can move some violating matches into compliance at the minimum wage. The incidence of non-compliance I disciplines the extensive margin of violations, while the conditional shortfall G disciplines the depth of underpayment and the effective penalty wedge. We leave the squared conditional shortfall G_2, the lower-tail wage moment $p25(w < 0.995)$, and additional untargeted wage-distribution statistics as validation moments.

B.5 Survey weights

All descriptive statistics, reduced-form outcomes, and structural data moments are computed using the GLD survey weights. The only exception is the construction of the wage-ratio winsorization cutoffs: these cutoffs are computed within year-by-department cells using unweighted 1st and 99th percentiles. Once wages are winsorized, the resulting moments and regressions use survey weights. The reduced-form regressions also cluster inference at the department level.

B.6 Relationship between samples

The reduced-form and structural samples are intentionally close, but not identical. The reduced-form sample is a panel of repeated cross-sections of private-sector paid workers between 2010 and 2015. It is designed to estimate the causal effect of Law 1610 using pre-reform department exposure.

The structural sample is the 2012 pre-reform cross-section. It adds unemployed individuals satisfying the same age and education restrictions because unemployment is an equilibrium object in the model. The structural wage moments, however, remain based on the same type of employed private-sector paid workers used in the reduced-form analysis.

This distinction is useful. The reduced-form design estimates the empirical effect of the reform. The structural model uses the pre-reform labor market to recover the parameters governing wage setting, violations, detection, and penalties. The counterfactual section then uses the reduced-form estimates to discipline the post-reform change in enforcement severity.

C Institutional Details

This appendix provides additional institutional background for the setting and for the interpretation of the empirical design. The main text emphasizes three features. First, Colombia has a national statutory minimum wage, so regional variation in non-compliance does not reflect regional variation in the legal floor. Second, labor inspection is not only a fine technology. Detection can force an illegal arrangement to be corrected, regularized, or ended. Third, Law 1610 strengthened enforcement in a period when inspection capacity was also expanding. The empirical design and the structural bridge should therefore be read with these institutional facts in mind.

C.1 Minimum-wage setting

Colombia has a national statutory minimum wage, the *Salario Mínimo Mensual Legal Vigente*. The wage floor is set annually and applies nationally. The institutional process is organized around a tripartite negotiation between the government, employers, and workers (Congreso de Colombia, 1996). When the parties reach an agreement, the government adopts the agreed increase. When no agreement is reached, the executive sets the minimum wage by decree.

This national structure is useful for the empirical design. During the period we study, departments do not set their own statutory wage floors. Differences in measured non-compliance across departments therefore do not come from different legal minimum wages. They come from differences in local wage distributions, firm productivity, labor-market structure, and enforcement conditions.

The national wage floor also gives the structural model a natural normalization. Wages are expressed relative to the statutory minimum wage, and the minimum wage is normalized to one. A subminimum wage is therefore a wage below one in the model. In the empirical implementation, we use a 0.5 percent tolerance band around the legal floor to avoid classifying workers extremely close to the minimum as violators because of reporting or conversion error.

C.2 Labor inspection as detection and clearance

Minimum-wage compliance is enforced through the labor inspection system under the Ministry of Labor. Labor inspectors can monitor compliance with labor norms, investigate violations, mediate disputes, require corrective action, and impose administrative sanctions. The ILO convention on labor inspection defines the function of inspection around securing the enforcement of legal provisions on working conditions and worker protection (International Labour Organization, 1947). Colombia's inspection system follows this logic: inspectors are not only collectors of fines; they are also agents who can require an employer to correct non-compliance.

This distinction is central to the paper. For an employer, enforcement has a price component and a state component. The price component is the expected penalty: the probability of detection times the penalty conditional on detection. The state component is what happens to an ongoing violation when it is detected. A detected subminimum wage cannot continue indefinitely as the same illegal arrangement. It must be regularized, sanctioned, corrected through a formalization agreement, or ended.

The model summarizes this institutional feature by assuming that detection clears the violating state. If productivity can support the legal wage, detection converts the job into a compliant minimum-wage job. If productivity cannot support the legal wage, detection destroys the match. This is not a literal claim that every detected establishment closes. It is a parsimonious representation of the idea that detection changes the continuation value and duration of non-compliant jobs. Penalty severity, by contrast, changes the cost of underpayment inside a violating job but does not itself create a transition out of the violating state.

C.3 Law 1610 of 2013

Law 1610 was enacted on January 2, 2013. The reform regulated labor inspections and labor formalization agreements. Its most direct sanctioning provision increased the maximum administrative fine for labor violations to 5,000 monthly minimum wages (Congreso de Colombia, 2013). This was a large increase relative to the previous sanction ceiling of 100 monthly minimum wages (Ministerio de la Protección Social, 2011).

The reform did not change the statutory minimum wage. It changed the consequences of violating existing labor obligations. This is the institutional reason the paper studies enforcement design rather than the effect of a higher wage floor. A legal minimum wage may be national and salient, but its bite depends on the expected cost of paying below it.

Law 1610 should not be described as a one-dimensional penalty reform. It clarified inspection functions, regulated labor formalization agreements, and formed part of a broader effort to strengthen labor-law enforcement. The statutory increase in the fine ceiling gives the cleanest mapping to the penalty-severity margin in the model. The broader enforcement provisions matter because they may also affect the credibility of enforcement and the ability of inspectors to induce correction.

This is why the structural bridge is deliberately narrow. In the bridge, we hold the baseline detection rate fixed and ask how much penalty severity must rise to reproduce the reduced-form decline in the conditional shortfall. This does not mean that detection capacity was fixed in Colombia. It means that the bridge isolates the penalty-severity change consistent with the empirical depth response.

C.4 Labor formalization agreements

Law 1610 also regulated labor formalization agreements between employers and the Ministry of Labor. These agreements were designed to correct labor violations and promote compliance. They are relevant for the model because they reinforce the idea that detection is not only an ex post fine. Enforcement can create a path from violation to compliance.

In the model, we do not introduce formalization agreements as a separate policy instrument. We instead treat them as part of the enforcement environment created by Law 1610. When the reduced-form evidence shows that the depth of underpayment declines in more exposed departments, the structural bridge interprets that change as a policy-relevant increase in the cost of deep violations. Formalization agreements are one institutional channel through which that cost may become credible.

C.5 Inspection-capacity expansion

Law 1610 occurred during a period when Colombia was also expanding labor-inspection capacity. As part of the Colombian Action Plan Related to Labor Rights, the government committed to hire additional labor inspectors. Pignatti (2026) studies this expansion and reports that the number of inspectors increased from about two to four per 100,000 individuals between 2011 and 2014.

Pignatti's design is different from ours. He exploits an allocation rule used to distribute inspector vacancies across departments and years. The rule assigned hiring targets using a department-level score intended to capture the risk of labor-law non-compliance. Within an integer value of that score, whether the decimal component crossed a 0.5 threshold generated discontinuous variation in the number of inspectors allocated to otherwise similar departments. Using this variation, Pignatti (2026) finds that additional inspectors increased labor-law compliance in the formal sector, with weaker or no effects in the informal sector.

This evidence confirms that detection was an active enforcement margin in Colombia. It also rules out a literal interpretation of the post-2013 period as a one-dimensional fine change. Our empirical variation is different, however. We do not use the inspector-allocation discontinuity. We use pre-reform departmental exposure to minimum-wage underpayment to measure where the national enforcement change should have more bite.

A contemporaneous expansion in inspections threatens our design only if it differentially changed high- and low-exposure departments in a way correlated with our exposure measure and if it generated the same timing and margin as our reduced-form estimates. The model helps discipline this concern. A relative increase in detection should primarily reduce the incidence of violations. A relative increase in penalty severity or expected liability should primarily reduce the depth of violations. The reduced-form estimates display a pattern closer to the latter: the conditional shortfall falls sharply, while incidence does not display a comparable break.

C.6 Law 1607 and other concurrent reforms

The main concurrent-policy concern is Law 1607 of December 2012. Law 1607 reduced non-wage labor costs for eligible employers by changing payroll contributions and related employer obligations (Congreso de Colombia, 2012). Existing studies show that this payroll-tax reform affected formal employment and related labor-market outcomes in Colombia (Kugler et al., 2017a; Morales and Medina, 2017; Kugler et al., 2017b).

This reform matters because it occurred close to Law 1610. It does not automatically invalidate the design. Our empirical comparison is between departments with different pre-reform minimum-wage underpayment exposure. Law 1607 is a threat if it produced differential changes across those exposure groups that mimic the timing and margin of the Law 1610 response.

The economic margins are also different. Law 1607 reduced the cost of formal employment. Law 1610 raised the expected cost of violating labor obligations. If the exposure-based estimates were mainly driven by a broad formalization or labor-cost channel, we would expect differential responses in social-security coverage, contracts, employment, hours, or sectoral allocation. The reduced-form robustness section examines these outcomes. Their role is not to prove that concurrent reforms had no effect. Their role is to assess whether the exposure design is capturing a broad regional labor-market shift rather than a minimum-wage compliance response.

C.7 Why exposure varies across departments

Law 1610 applied nationally. There is no untreated department and no staggered implementation. The empirical design therefore uses variation in pre-reform exposure to non-compliance. Departments with larger pre-reform unconditional shortfalls had more wage relationships potentially affected by a stronger enforcement environment.

This exposure design relies on the idea that a national enforcement reform has more bite where violations were initially more prevalent or more severe. The identifying comparison is not between treated and untreated regions. It is between departments with different pre-reform exposure to the same national policy change. The design requires that high- and low-exposure departments would not have experienced differential changes in minimum-wage non-compliance after 2013 absent the reform.

The event-study dynamics, auxiliary labor-market outcomes, and falsification exercises assess this requirement. The appendix evidence supports an interpretation in which the reform changed the depth of minimum-wage underpayment in more exposed departments, rather than generating a broad regional reallocation of employment. This is the institutional and empirical reason the structural counterfactuals use the reduced-form decline in the conditional shortfall to discipline the penalty-severity bridge.

D Additional Descriptive Evidence

This appendix reports additional descriptive evidence on minimum-wage non-compliance in Colombia. The purpose is not to estimate the effect of Law 1610. Instead, the appendix documents the pre-reform environment, the cross-department variation used in the empirical design, and the aggregate evolution of non-compliance outcomes.

D.1 Descriptive statistics by exposure

Appendix Table 6 compares departments with above- and below-median pre-reform exposure. Exposure is the department-level unconditional shortfall averaged over 2010 and 2011. The table shows that high-exposure departments had substantially more non-compliance before the reform. They also differed in observable labor-market structure.

The differences are informative for two reasons. First, they show that the reform had more potential bite in high-exposure departments. In these departments, a larger share of employment relationships involved subminimum wages, and violations were deeper. Second, they clarify why the reduced-form analysis must examine outcomes beyond minimum-wage compliance. High-exposure departments had lower social-security coverage and a different sectoral composition before the reform. These differences do not invalidate the design by themselves, because department fixed effects absorb time-invariant level differences. They do, however, motivate the event-study checks for differential pre-trends and the additional labor-market outcomes reported in the reduced-form section.

D.2 Aggregate evolution of non-compliance

Appendix Figure 4 plots the national time series of the three non-compliance outcomes: incidence, conditional shortfall, and unconditional shortfall. The figure is descriptive and pools departments using the survey weights.

Two facts are useful for interpreting the paper. First, non-compliance was not a rare event in Colombia. The incidence of subminimum wages was high throughout the pre-reform period. Second, the depth of violations declined after the reform period. This aggregate decline is consistent with the reduced-form evidence, although it is not by itself causal because it may also reflect national shocks, changes in the minimum wage, or other reforms.

Table 6: Pre-reform descriptive statistics by exposure tercile

	Low	Middle	High	Overall
Non-compliance incidence	0.484	0.665	0.744	0.549
Unconditional shortfall	0.158	0.248	0.347	0.198
Conditional shortfall	0.326	0.372	0.467	0.361
Wage/minimum-wage ratio	1.030	0.874	0.756	0.968
Formal/social security	0.496	0.322	0.241	0.434
Age	35.315	35.139	34.453	35.172
Female	0.376	0.313	0.345	0.362
Agriculture	0.152	0.291	0.278	0.192
Industry	0.292	0.220	0.184	0.266
Services	0.384	0.313	0.318	0.364
Observations	121,531	76,262	74,589	272,382

Notes: The table reports weighted means for private-sector paid workers aged 15–64 with at most secondary education in 2010–2012. Exposure terciles are based on the department-level unconditional minimum-wage shortfall averaged over 2010 and 2011. Conditional shortfall is computed among workers paid below the statutory minimum wage.

The reduced-form design therefore does not rely on the aggregate time series. It uses the fact that departments entered the reform with different degrees of exposure to underpayment. The aggregate figure is useful as background: it shows that the reform occurred in a labor market where violations were economically meaningful and measurable.

D.3 Wage distribution around the minimum wage

Appendix Figure 5 reports the distribution of hourly wages relative to the statutory minimum wage in the pre-reform period. The figure shows two features that motivate the structural model. There is a visible mass near the minimum wage, but there is also substantial employment below the statutory floor. The legal minimum wage is therefore salient, but not fully enforced.

This pattern is important for the model. A standard minimum-wage model with perfect enforcement would not generate a large stock of subminimum wages among covered workers. The model in this paper allows firms and workers to bargain over subminimum wages when expected enforcement costs are low. At the same time, the mass near the minimum wage helps identify the enforcement channel because detection can move some violating matches into compliance at the statutory floor.

The figure is descriptive rather than causal. It is not used as a reduced-form outcome. It documents the wage distribution and supports the structural moment that captures the mass around the minimum wage.

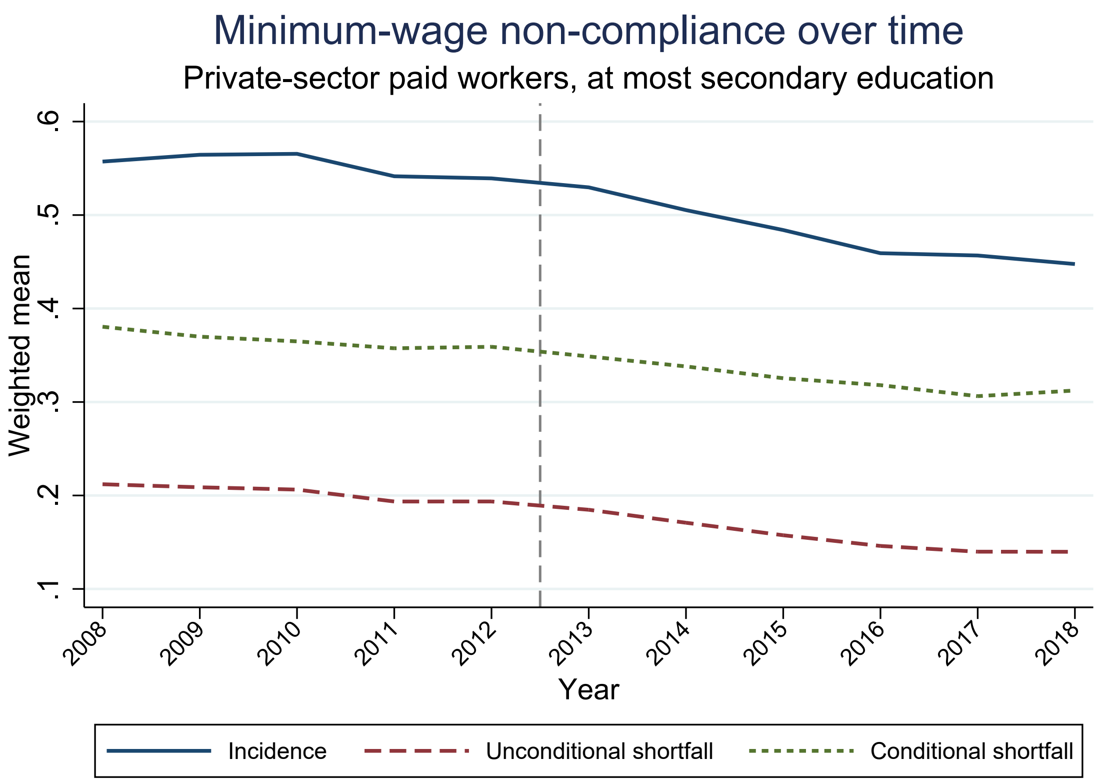


Figure 4: Minimum-wage non-compliance over time

Notes: The figure reports weighted national means for private-sector paid workers aged 15–64 with at most secondary education. The vertical dashed line marks the period immediately before Law 1610.

Figure 5: Pre-reform wage distribution around the minimum wage

Notes: The figure reports the distribution of hourly wages relative to the hourly minimum wage in 2010–2012 for private-sector paid workers aged 15–64 with at most secondary education.

D.4 Exposure and geographic heterogeneity

The department-level exposure measure summarizes both the prevalence and depth of underpayment. Departments with high exposure are not simply departments with slightly more workers below the minimum wage. They are departments where total underpayment relative to the statutory minimum wage is larger.

This heterogeneity is central to the empirical design. Because Law 1610 applied nationally, identification cannot come from untreated regions. Instead, the design compares departments with different pre-reform exposure to the same national enforcement reform. The descriptive evidence shows that this exposure variation is economically large enough to make the design informative.

Appendix Figure 6 plots the exposure measure by department. Appendix Table 7 reports the same values. The figure and table are included here because the exposure measure is constructed before the reform and supports the empirical design, but is not itself a reduced-form estimate.

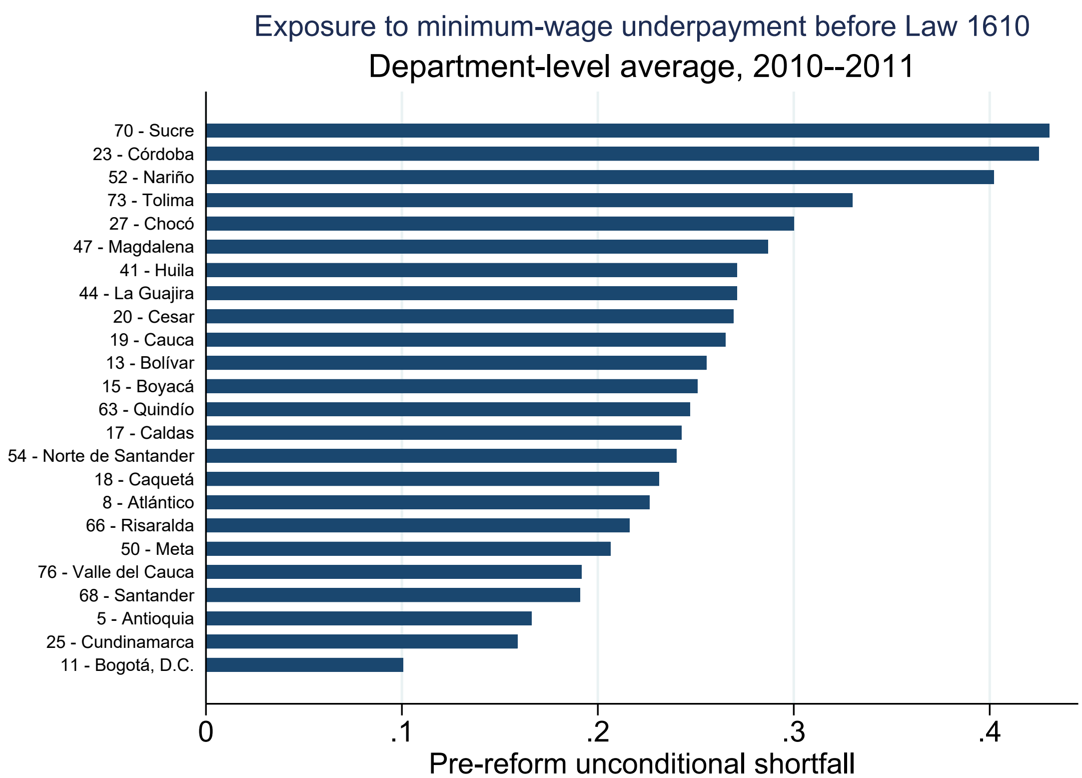


Figure 6: Pre-reform exposure by department

Notes: The figure reports the department-level unconditional minimum-wage shortfall averaged over 2010 and 2011. The sample is private-sector paid workers aged 15–64 with at most secondary education. Higher values indicate greater pre-reform exposure to minimum-wage underpayment.

Table 7: Pre-reform exposure by department

Department	Unconditional shortfall	Incidence
Sucre	0.430	0.822
Córdoba	0.425	0.822
Nariño	0.402	0.820
Tolima	0.330	0.727
Chocó	0.300	0.703
Magdalena	0.287	0.678
Huila	0.271	0.656
La Guajira	0.271	0.637
Cesar	0.269	0.690
Cauca	0.265	0.636
Bolívar	0.255	0.618
Boyacá	0.251	0.646
Quindío	0.247	0.734
Caldas	0.243	0.730
Norte de Santander	0.240	0.657
Caquetá	0.231	0.688
Atlántico	0.226	0.577
Risaralda	0.216	0.644
Meta	0.206	0.588
Valle del Cauca	0.192	0.555
Santander	0.191	0.548
Antioquia	0.166	0.505
Cundinamarca	0.159	0.469
Bogotá, D.C.	0.101	0.376

Notes: Exposure is computed among private-sector paid workers aged 15–64 with at most secondary education. The baseline exposure measure is the department-level unconditional shortfall averaged over 2010 and 2011.

E Reduced-Form Robustness

This appendix reports additional reduced-form results supporting Section 5. All specifications follow the event-study design in equation (27), unless otherwise noted. The omitted year is 2012, the last pre-reform year. Confidence intervals are computed using the wild cluster bootstrap with Webb weights.

E.1 Baseline coefficients

Appendix Table 8 reports the event-study coefficients for the three non-compliance outcomes: incidence, conditional shortfall, and unconditional shortfall. The table gives the numerical counterpart to the figures in the main text and Appendix Figure 7.

The table confirms the main pattern. The incidence estimates are imprecise. By contrast, the conditional shortfall falls sharply after the reform in more-exposed departments. The unconditional shortfall also declines, reflecting the fact that the reduction in the depth of violations dominates the noisier incidence response.

The pre-reform coefficients are small relative to the post-reform effects. This is important because the exposure design compares departments with different levels of pre-reform underpayment. The credibility of the design requires that high-exposure departments were not already on differential trends before Law 1610. The evidence in the pre-period is consistent with this requirement.

Table 8: Reduced-form effects on minimum-wage non-compliance

Year	Incidence	Conditional shortfall	Unconditional shortfall	Bunching
2010	-0.043	0.025	0.030	-0.067
2011	-0.002	0.018	0.011	0.336
2012	0.000	0.000	0.000	0.000
2013	-0.052	-0.125	-0.111	0.007
2014	-0.082	-0.226	-0.208	-0.136
2015	0.012	-0.186	-0.179	-0.212

Notes: The table reports coefficients from event-study regressions interacting year indicators with department-level pre-reform unconditional shortfall exposure. The omitted year is 2012. The sample is private-sector paid workers aged 15–64 with at most secondary education. Regressions include department and year fixed effects and controls for age, age squared, and gender. Confidence intervals are reported in the corresponding figure using wild cluster bootstrap with Webb weights.

Appendix Figure 7 reports the event-study for the unconditional shortfall. This outcome combines the incidence and conditional shortfall margins. The figure shows a clear decline after Law 1610, reinforcing the interpretation that the reform reduced total underpayment even though the incidence response is less precisely estimated.

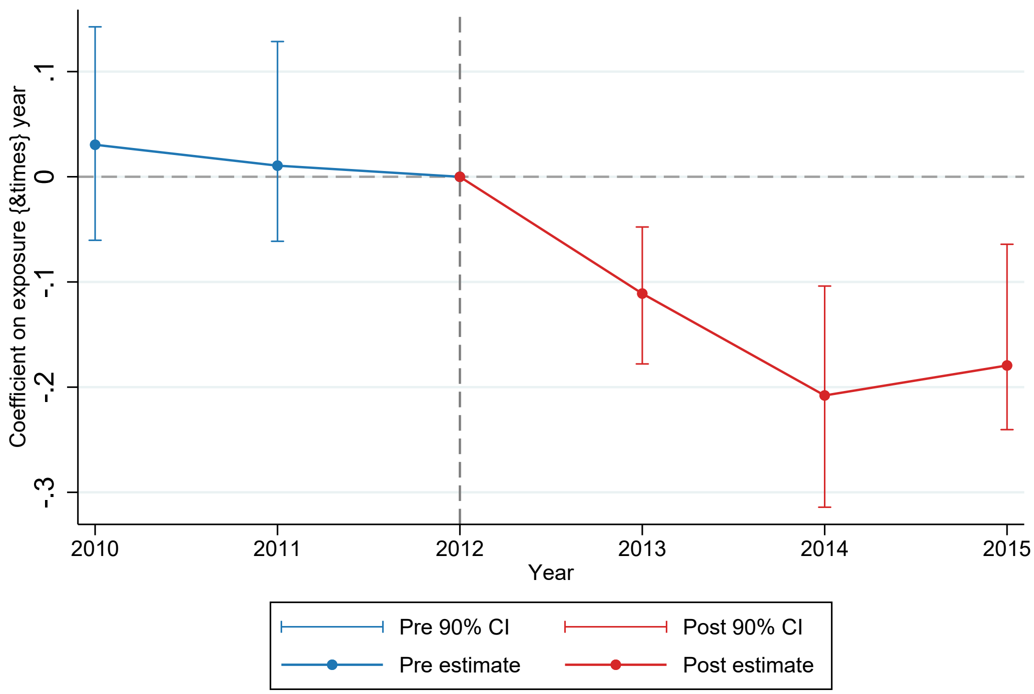


Figure 7: Effect on the unconditional shortfall

Notes: The figure reports event-study coefficients for the unconditional shortfall, $I \times G$. Exposure is the department-level unconditional minimum-wage shortfall averaged over 2010 and 2011. The omitted year is 2012. The sample is private-sector paid workers aged 15–64 with at most secondary education. Confidence intervals are 90 percent wild-cluster-bootstrap intervals with Webb weights.

E.2 Other labor-market outcomes

A central identification concern is that Law 1610 was enacted close to Law 1607, which reduced non-wage labor costs. If the reduced-form estimates were driven by a broad formalization shock rather than by enforcement of the wage floor, high-exposure departments should display clear differential changes in formality, contractual employment, hours, wages, or sectoral composition.

Appendix Figures 8–14 report event-study estimates for these additional outcomes. We first examine social-security coverage and the presence of an employment contract. These outcomes capture formalization margins. We then examine log hourly wages and usual weekly hours. Finally, we examine broad sectoral allocation across agriculture, industry, and services.

The evidence does not show a broad reallocation pattern comparable to the shortfall response. Social-security coverage and employment contracts do not display a sharp post-reform increase in more-exposed departments. Hours and sectoral shares also do not show a systematic break after 2012. Log hourly wages rise after the reform, which is consistent with the decline in wage shortfalls, but the wage result should be interpreted as supporting evidence rather than as an independent mechanism.

These results help separate the minimum-wage compliance channel from a broad formalization or composition channel. They do not prove that Law 1607 had no effects. Rather, they show that the main exposure-based pattern is concentrated in the depth of minimum-wage violations.

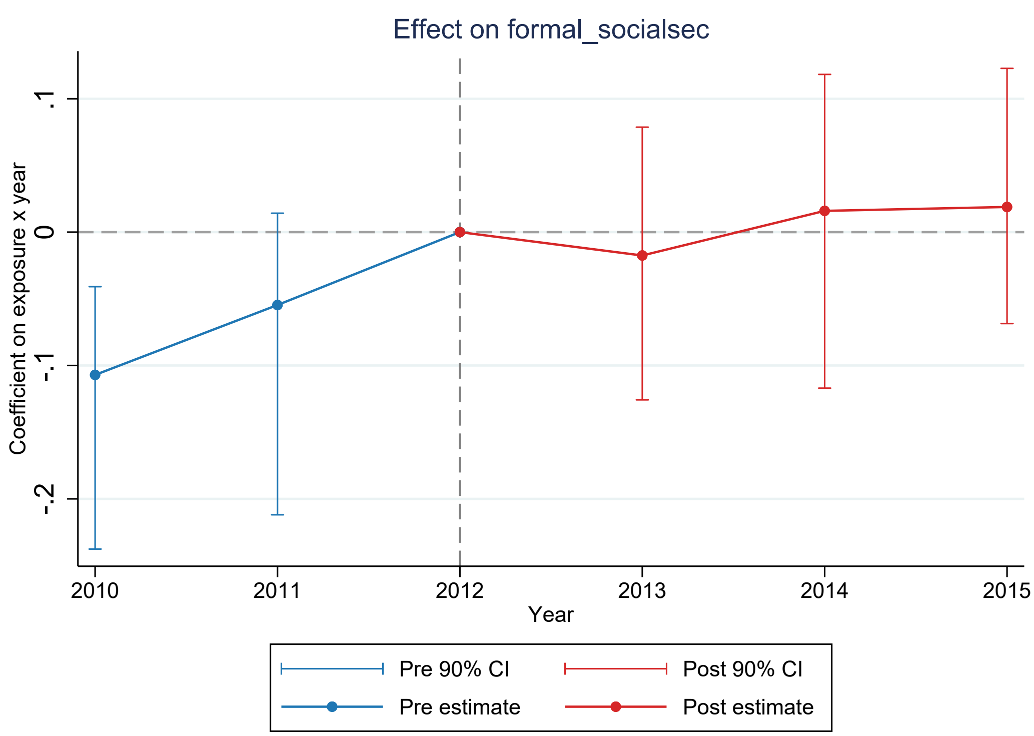


Figure 8: Effect on social-security coverage

Notes: The figure reports event-study coefficients for social-security coverage. The sample is private-sector paid workers aged 15–64 with at most secondary education.

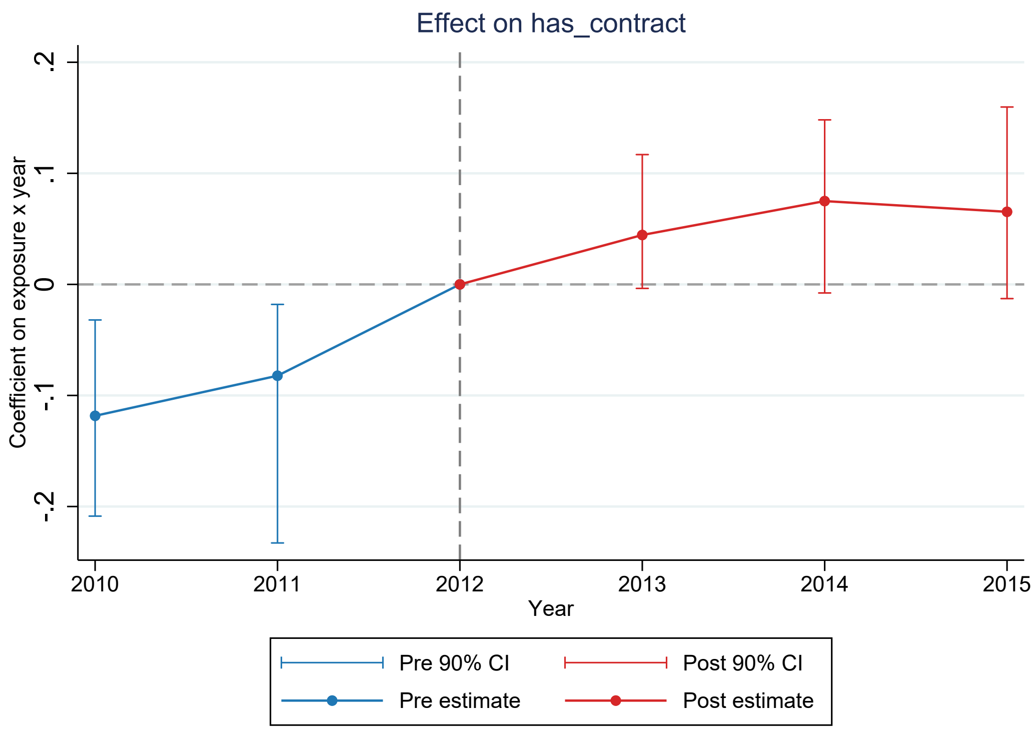


Figure 9: Effect on having an employment contract

Notes: The figure reports event-study coefficients for having an employment contract.

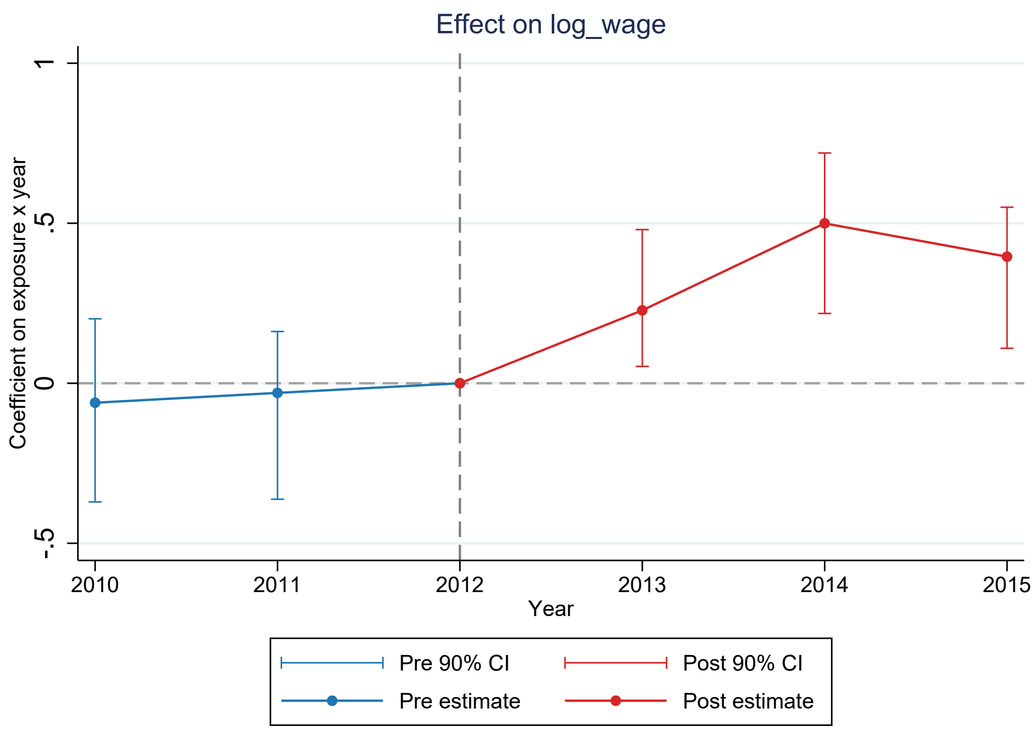


Figure 10: Effect on log hourly wages

Notes: The figure reports event-study coefficients for log hourly wages.

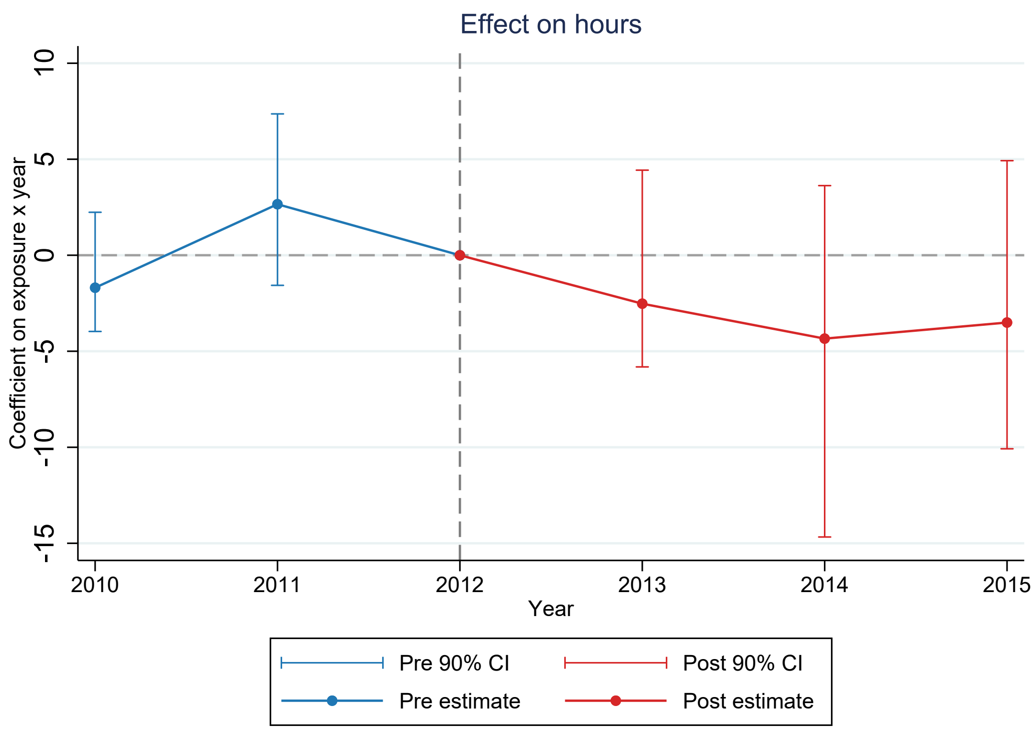


Figure 11: Effect on weekly hours

Notes: The figure reports event-study coefficients for usual weekly hours.

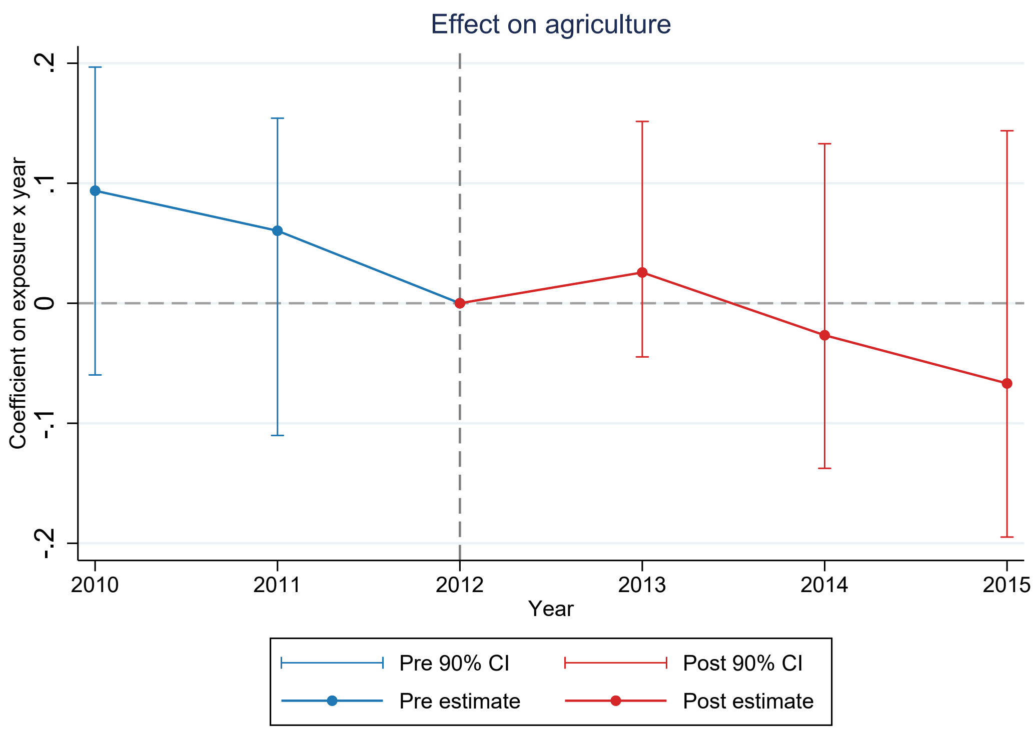


Figure 12: Effect on agricultural employment

Notes: The figure reports event-study coefficients for employment in agriculture.

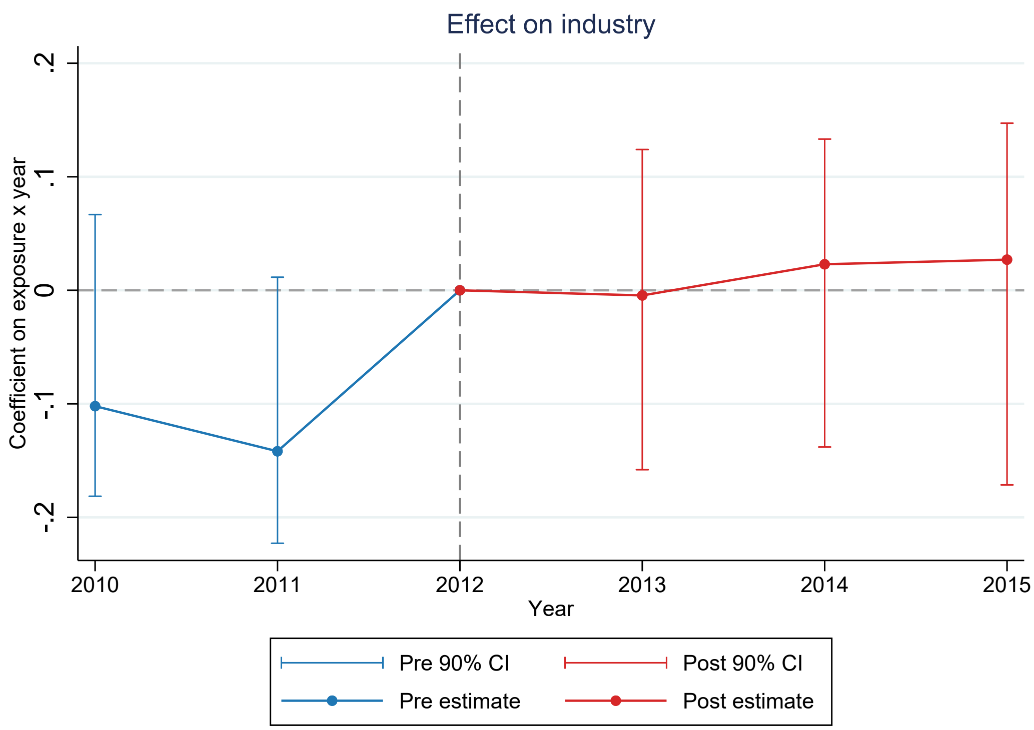


Figure 13: Effect on industrial employment

Notes: The figure reports event-study coefficients for employment in industry.

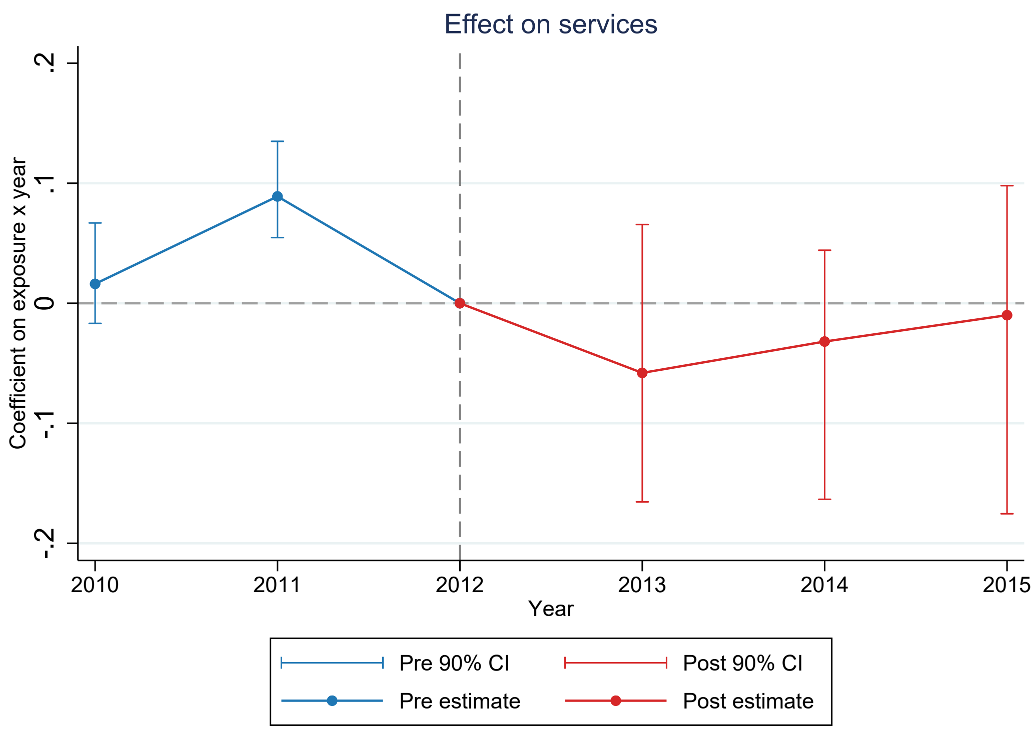


Figure 14: Effect on services employment

Notes: The figure reports event-study coefficients for employment in services.

E.3 Alternative exposure and falsification checks

The baseline exposure measure is the pre-reform unconditional shortfall. This measure is natural because it combines the share of workers paid below the legal floor and the depth of underpayment. A possible concern is that the post-reform decline in the conditional shortfall is mechanically related to using pre-reform shortfalls to construct exposure.

We therefore reconstruct exposure using the pre-reform incidence of subminimum wages. This alternative exposure measure ranks departments by the share of workers paid below the minimum wage in 2010–2011, rather than by total underpayment. Appendix Table 9 reports the results. The conditional and unconditional shortfall estimates remain negative after the reform, while the incidence estimates remain less precise. This exercise supports the main interpretation. Departments more exposed to non-compliance experience larger reductions in the depth of violations even when exposure is measured using the extensive margin.

Table 9: Robustness: exposure based on pre-reform non-compliance incidence

Year	Incidence	Conditional shortfall	Unconditional shortfall
2010	-0.103 [-0.192, 0.078]	-0.011 [-0.075, 0.091]	-0.007 [-0.080, 0.126]
2011	0.037 [-0.024, 0.137]	-0.006 [-0.064, 0.110]	0.007 [-0.057, 0.134]
2012	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]
2013	0.049 [-0.035, 0.138]	-0.112 [-0.158, -0.032]	-0.088 [-0.145, 0.000]
2014	-0.076 [-0.149, 0.092]	-0.171 [-0.279, -0.036]	-0.185 [-0.297, -0.039]
2015	0.028 [-0.033, 0.172]	-0.176 [-0.255, 0.007]	-0.172 [-0.251, 0.002]
2016	-0.052 [-0.158, 0.220]	-0.174 [-0.236, -0.014]	-0.193 [-0.256, -0.028]

Notes: Cells report event-study coefficients with 90 percent wild-cluster-bootstrap confidence intervals in brackets. Exposure is the department-level incidence of minimum-wage non-compliance averaged over 2010 and 2011. The omitted year is 2012. The sample is private-sector paid workers aged 15–64 with at most secondary education.

The main sample focuses on workers with at most secondary education because this is the part of the labor market where the minimum wage is most likely to bind. Workers with tertiary education provide a useful falsification group. They are employed in the same national economy and are exposed to many of the same macroeconomic and institutional shocks, but minimum-wage non-compliance should be less relevant for them.

Appendix Table 10 repeats the reduced-form design for private-sector paid workers with tertiary education. We do not find a comparable post-reform decline in the unconditional shortfall for this group. This result supports the interpretation that the baseline effects are tied to minimum-wage enforcement rather than to a generic department-level labor-market shock.

The falsification has limits. Tertiary workers differ from the main sample in wages, occupations, and sectors. The exercise is therefore not a substitute for the event-study design. It is useful because it asks whether the same exposure measure predicts similar changes for a group less likely to be directly constrained by the minimum wage.

The baseline reduced-form window is 2010–2015. We choose this window to focus on the years

Table 10: Falsification: workers with tertiary education

Year	Incidence	Unconditional shortfall
2010	-0.084 [-0.167, 0.074]	0.019 [-0.037, 0.062]
2011	-0.059 [-0.147, 0.005]	0.009 [-0.063, 0.053]
2012	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]
2013	-0.001 [-0.094, 0.188]	0.026 [-0.009, 0.072]
2014	-0.020 [-0.175, 0.184]	0.003 [-0.029, 0.049]
2015	-0.101 [-0.196, 0.017]	-0.019 [-0.060, 0.038]
2016	0.086 [-0.076, 0.329]	0.015 [-0.021, 0.057]

Notes: Cells report event-study coefficients with 90 percent wild-cluster-bootstrap confidence intervals in brackets. The sample is private-sector paid workers aged 15–64 with some tertiary or post-secondary education. Exposure is the baseline department-level unconditional shortfall among the main lower-education sample.

closest to Law 1610 while avoiding later institutional changes and longer-run composition shifts. The cleaned data cover 2008–2018, which allows us to examine a longer window.

Appendix Table 11 reports the extended-window estimates. The post-reform decline in wage shortfalls remains visible in the longer sample. At the same time, the additional years are less clean for identification because they increase exposure to other policy changes and macroeconomic shocks. For this reason, we retain the shorter 2010–2015 window as the baseline and interpret the extended-window estimates as supporting evidence.

Table 11: Robustness: extended 2008–2018 event-study window

Year	Incidence	Conditional shortfall	Unconditional shortfall
2008	-0.096 [-0.169, 0.021]	0.029 [-0.060, 0.146]	0.015 [-0.082, 0.138]
2009	-0.061 [-0.177, 0.184]	0.089 [0.029, 0.169]	0.080 [0.015, 0.198]
2010	-0.088 [-0.182, 0.053]	0.012 [-0.056, 0.109]	0.011 [-0.072, 0.134]
2011	0.028 [-0.021, 0.091]	0.010 [-0.054, 0.116]	0.016 [-0.048, 0.126]
2012	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]
2013	0.025 [-0.068, 0.123]	-0.105 [-0.165, -0.043]	-0.091 [-0.153, -0.023]
2014	-0.069 [-0.132, 0.064]	-0.175 [-0.269, -0.073]	-0.189 [-0.284, -0.076]
2015	0.023 [-0.034, 0.127]	-0.157 [-0.223, -0.017]	-0.158 [-0.219, -0.028]
2016	-0.029 [-0.132, 0.174]	-0.152 [-0.208, -0.013]	-0.166 [-0.225, -0.025]
2017	0.001 [-0.154, 0.321]	-0.161 [-0.214, -0.052]	-0.173 [-0.250, -0.014]
2018	-0.080 [-0.220, 0.204]	-0.180 [-0.322, 0.063]	-0.210 [-0.336, 0.034]

Notes: Cells report event-study coefficients with 90 percent wild-cluster-bootstrap confidence intervals in brackets. The sample is the core private-sector paid-worker sample with at most secondary education, extended to 2008–2018. The omitted year is 2012. This window is not the baseline because the additional years raise separate pre-period and post-period institutional concerns.

E.4 Summary

The robustness evidence supports the main reduced-form interpretation. The post-reform response is concentrated in the depth of minimum-wage violations. The incidence estimates are less precise. We do not observe similarly sharp changes in social-security coverage, contracts, hours, or sectoral allocation. The results are also consistent when exposure is measured using pre-reform incidence and are not replicated among tertiary-educated workers.

These checks suggest that Law 1610 reduced underpayment primarily through an intensive-margin compliance response.

F Minimum-Distance Estimation and Identification

This appendix describes the estimation procedure used in Section 6. The model derivations are in Appendix A. We focus here on the objects that enter the minimum-distance estimator: the estimated parameters, calibrated objects, targeted moments, weighting matrix, numerical search, and identification logic.

F.1 Implementation details

The estimation uses the parameter vector, normalizations, and fixed parameters described in Section 6. This appendix records implementation details that are useful for replication but not central to the main text.

First, we impose a small positive lower bound on rU. This rules out estimates in which the bargaining-relevant value of unemployment is mechanically pushed to zero or below. Because the estimate lies close to this lower bound, we interpret rU as the flow value of unemployment relevant for the wage bargain, not as a primitive measure of home production or social-insurance generosity. The primitive flow value of unemployment, denoted b in the code output, is recovered as a diagnostic object and is not a central parameter in the paper.

Second, the job-contact rate λ is recovered at each candidate parameter vector from the external unemployment-to-employment transition target and the model-implied acceptance probability. The separation rate ρ is estimated jointly with the other structural parameters, while the unemployment rate enters the targeted moment vector. Thus the external transition target anchors job contacts, and the Colombian unemployment rate disciplines separations in the estimated steady state.

F.2 Targeted, implied, and validation moments

The preferred specification targets the six moments listed in equation (30) in the main text. The empirical moments are constructed using the same normalized wage regions throughout the paper: workers with $w < 0.995$ are classified as subminimum-wage workers, workers with $0.995 \leq w \leq 1.005$ enter the minimum-wage mass, and workers with $w > 1.005$ enter the above-minimum wage distribution. The corresponding model moments are computed from the stationary wage distribution generated by Θ.

We separately report two objects that are implied by targeted moments. The unconditional gap among employed workers is

$$I \times G,$$

and aggregate non-compliance in the labor force is

$$(1-u)IG.$$

The first combines the incidence of subminimum work with the conditional shortfall among subminimum workers. The second additionally scales this gap by the employment rate. Because u, I, and G are targeted moments, these objects are useful for interpretation but should not be read as independent validation moments.

Validation moments are computed after estimation and do not enter the minimum-distance objective. They include moments of the subminimum wage distribution, such as G_2, lower-tail percentiles, and subminimum wage dispersion, as well as moments of the overall employed wage distribution. These moments are used to assess whether the estimated economy reproduces features of the wage distribution beyond the objects directly matched in estimation.

F.3 Objective function and weighting

Let m^{data} denote the vector of targeted empirical moments and $m^{model}(\Theta)$ the corresponding model-implied moments. The minimum-distance estimator solves

$$\widehat{\Theta} = \arg\min_{\Theta} \left[m^{data} - m^{model}(\Theta)\right]' W \left[m^{data} - m^{model}(\Theta)\right]. \tag{64}$$

The baseline weighting matrix is diagonal. Each diagonal element is the inverse of a bootstrap estimate of the sampling variance of the corresponding targeted data moment:

$$W_{kk} = \frac{1}{\widehat{\mathrm{Var}}(m_k^{data})}. \tag{65}$$

The bootstrap resamples households when household identifiers are available, so that workers in the same household remain clustered in the resampled data. The purpose of this weighting is to put moments measured with different precision on comparable footing. Because the preferred specification has the same number of targeted moments as estimated parameters, the economic interpretation of the estimate comes from the moment restrictions themselves rather than from a high-dimensional choice of weights.

Model moments are computed by deterministic quadrature over the productivity distribution. The same 0.5 percent minimum-wage tolerance used in the empirical analysis defines the below-minimum, near-minimum, and above-minimum regions when data moments are constructed. The theoretical model uses the normalized cutoff $w = 1$, while the empirical implementation uses the tolerance band to avoid classifying workers extremely close to the legal floor as violators because of rounding or reporting error.

F.4 Numerical search

The minimum-distance objective is nonlinear. We therefore combine a warm start with multiple random starting values. Each candidate parameter vector is transformed from an unconstrained space into the admissible parameter space. This ensures that $\sigma_z > 0$, rU lies in the admissible interval, $\chi \geq 0$, $\pi \geq 0$, and $\rho \geq 0$.

For the preferred estimates, the main estimation script uses 200 random starts and a productivity grid with 3,000 points. The reported estimate is the parameter vector with the lowest objective value across all starts. The optimization log is saved as part of the replication output.

The search is especially important for the enforcement parameters. The conditional shortfall gives information about χ, because the penalty schedule makes deeper violations more costly. The incidence moment disciplines the stock of subminimum jobs and therefore loads on the duration of violation spells. The mass around the minimum wage helps discipline conversion and bunching because detection can move high-productivity violators into compliant minimum-wage employment. The data discipline χ and π jointly, and penalty severity κ is recovered from their ratio.

F.5 Identification diagnostics

Section 6.2 describes the economic mapping from targeted moments to parameters. The implementation uses that mapping in two diagnostic ways. First, each candidate vector is required to generate a valid stationary equilibrium: the wage thresholds are ordered, the regularity condition in Proposition 1 holds, and the implied worker stocks are non-negative and sum to one. Second, after estimating (χ, π), we recover penalty severity as $\kappa = \chi/\pi$ and verify that the separate roles of χ and π are reflected in the fitted moments: shortfalls load on wage compression, while incidence and bunching load on violation duration and conversion into compliance.

External inspection benchmarks are used only to interpret the magnitude of the estimated detection probability. They are not imposed in the preferred minimum-distance objective.

G Additional Counterfactual Results

This appendix reports additional counterfactual results supporting Section 7. The main text presents the first-year structural bridge for Law 1610, the instrument-response figure, and the equal-benefit comparison between penalties and detection. Here, we add two pieces of information. First, we describe how the wage response helps interpret the bridge and the instrument-response exercise. Second, we clarify why the policy-multiplier figures report only feasible multipliers.

G.1 Additional diagnostics for the Law 1610 bridge

The main text bridges the 2013 reduced-form decline in the conditional shortfall into the structural model by holding the baseline detection probability fixed and choosing the penalty severity that reproduces the observed decline in G. The target decline is 2.5 percentage points at mean exposure. This mapping increases κ from 1.64 to 12.91, a multiplier of 7.9.

The bridge is disciplined by the depth response, but it also has implications for other moments. The model predicts a 4.6 percent increase in mean log wages, almost identical to the scaled 2013 reduced-form estimate of 4.5 percent. This is useful because wages are not targeted by the bridge. The model also predicts a 3.8 percentage-point decline in the unconditional shortfall $I \times G$, compared with a reduced-form estimate of 2.2 percentage points. The difference does not overturn the mechanism, but it shows that the penalty bridge produces a somewhat stronger aggregate underpayment response than the reduced form.

The main discrepancy is on incidence. The model predicts a 7.5 percentage-point decline in I, while the scaled reduced-form estimate is a 1.0 percentage-point decline. The reduced-form confidence interval is wide, but the model-implied decline is outside both the 90 percent interval reported in the reduced-form output and the conventional 95 percent interval based on the clustered standard error. This is why the main text treats the bridge as a disciplined mapping from the depth response, not as a complete accounting of every margin of Law 1610. The model's economic mechanism is useful, but the first-year reduced form does not show the same extensive-margin response implied by a pure penalty-severity interpretation.

G.2 Wage responses to enforcement instruments

Appendix Figure 15 reports wage responses to proportional increases in penalty severity and detection. The figure complements the main instrument-response figure. The main text focuses on non-compliance, unemployment, and the mass at the minimum. The wage-response figure shows how those mechanisms appear in the wage distribution.

The first pattern concerns penalties. Increasing κ raises wages among workers who remain below the statutory minimum. This is the direct bargaining channel of the model. A higher penalty makes deep violations more costly, so the wage bargain shifts toward the legal floor. The conditional shortfall falls and the average wage among remaining subminimum workers rises.

The second pattern concerns detection. Increasing π raises average wages in the economy because many violating jobs are either converted into compliant minimum-wage jobs or removed from the stock of violations. At the same time, the average wage among remaining violators can fall. This is a composition effect. When detection becomes more frequent, higher-productivity violations are more likely to leave the violating state through conversion. The remaining violating jobs are therefore more negatively selected and have deeper shortfalls on average.

This pattern explains why the conditional shortfall can rise under the detection counterfactual even though total underpayment falls. Detection reduces the number of violations. Penalties reduce the depth of surviving violations. The wage responses therefore give the same economic message as the main counterfactuals: the two instruments can reduce underpayment through different margins.

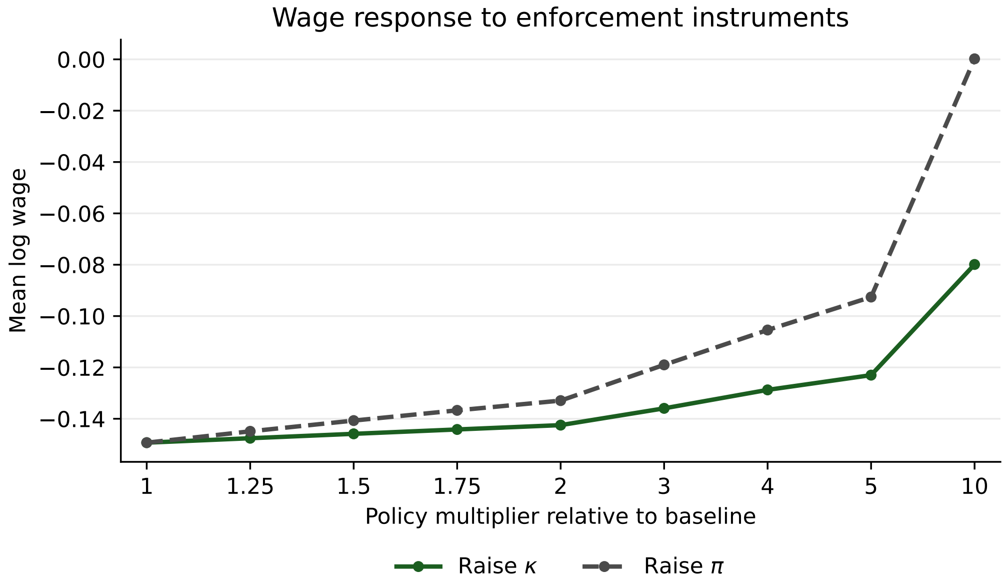


Figure 15: Wage responses to proportional increases in penalties and detection
Notes: The figure compares counterfactual proportional increases in penalty severity κ and the detection rate π. In each experiment, the other enforcement instrument is held fixed at its estimated baseline value. Wages are normalized by the statutory minimum wage. Dashed horizontal lines mark the baseline outcome level. Only feasible multipliers are shown.

G.3 Feasible policy multipliers

The policy-multiplier figures report proportional increases in κ and π relative to their estimated baseline values. In the code, the requested multiplier grid is 1, 1.25, 1.5, 1.75, 2, 3, 4, 5, 10, 15, and 20. Not all points are feasible in the estimated model. Large increases in either instrument can violate the model's regularity condition for the wage bargain. When this happens, the counterfactual is not economically admissible under the maintained equilibrium structure.

For this reason, the figures report only feasible multipliers. In the current baseline, the feasible grid runs through a multiplier of 10 for both instruments. Multipliers of 15 and 20 are not plotted because they violate the regularity condition. This choice avoids extrapolating outside the model's admissible region. It also makes the figure easier to read because the plotted points correspond to valid counterfactual equilibria.

The feasible range is still wide enough to show the mechanism. Over this range, penalties reduce the conditional shortfall and increase wages among remaining violators. Detection reduces incidence, increases the mass at the minimum, and raises unemployment more than penalties. The divergence between the two instruments appears before the boundary of the feasible region.

G.4 Equal-benefit decomposition

The equal-benefit counterfactuals in Section 7 compare a penalty increase and a detection increase that generate the same decline in the unconditional shortfall $I \times G$. The target is the 2013 reduced-form decline at mean exposure, equal to 2.2 percentage points. The penalty policy reaches the target by increasing κ by a factor of 5.1. The detection policy reaches the same target by increasing π by a factor of 2.5.

The decomposition in Section 7 separates the total change in $I \times G$ into an incidence component and a conditional-shortfall component. Under the penalty policy, the reduction comes mainly from the intensive margin. The conditional shortfall falls and incidence moves less. Under the detection policy, the reduction comes mainly from the extensive margin. Incidence falls substantially, while the conditional shortfall among remaining violators rises.

This decomposition is useful because it shows why equal aggregate reductions in underpayment need not imply similar labor-market effects. The detection policy produces a larger unemployment response because it destroys more low-productivity violating matches. The penalty policy produces a smaller employment response because it mainly changes wages inside surviving matches. Both policies reduce underpayment, but they do not produce the same allocation.

G.5 Interpretation and limits

These counterfactuals compare economic mechanisms, not fiscal costs. We do not observe the administrative resources required to increase detection, nor do we observe the legal, fiscal, or political-economy costs of increasing penalties. The model also does not evaluate welfare. It focuses on non-compliance, wages, unemployment, and the margins through which enforcement instruments operate.

The results should therefore be read as policy-design evidence. They show that equal reductions in total underpayment can come from different transition and wage-setting mechanisms. This distinction matters for policy because a reform aimed at reducing the number of violations can have different employment consequences than a reform aimed at reducing the depth of underpayment among surviving violations.