The Cost of Non-Compliance: Firm Adjustments to Threshold-Based Labor Policies*

Alix Naugler†

August 8, 2026

Abstract

This paper examines the unintended consequences of size-dependent formalization policies that raise the cost of informality for firms, focusing on a provision in Vietnam's Labor Code 2012. The policy sharply increases the financial penalty for firms with at least 10 formally contracted, paid employees that fail to comply with pre-existing labor regulations at this threshold. I develop a profit maximization model illustrating how this policy incentivizes firms to avoid compliance by substituting toward unregulated labor arrangements or partially formalizing. Using Vietnamese micro-, small-, and medium-enterprise panel data, I employ a difference-in-discontinuities design to estimate the causal effect of the increased financial penalty at the threshold. McCrary density tests suggest no evidence of bunching below the 10-formal paid employee threshold post-policy. Instead, firms adjusted along alternative margins: those just below the threshold increased their reliance on unpaid labor while those just above it registered with the government while employing informal workers. Firms just above the threshold also experienced profit and labor productivity gains. These findings show that threshold-based labor policies can lead to selective—rather than comprehensive—firm formalization, suggesting informality is restructured instead of reduced.

JEL codes: D22, O17, L25, O12

Keywords: Firm formalization, informal labor, size-dependent regulation, regulatory enforcement, selective compliance, difference-in-discontinuities, Vietnam

*I am grateful to Jenny Aker, Chris Barrett, Ben Leyden, Dan Maggio, Shubhalakshmi Nag, Elio Nimier-David, Anh Nguyen, Zhuan Pei, Heather Schofield, and Daniela Scur for their helpful feedback. I also thank participants at the 2026 Pacific Conference for Development Economics and the 2026 Agricultural and Applied Economics Association Annual Meeting for their valuable comments. All errors are my own. I received no external funding for this research and have no conflicts of interest to disclose.

†Cornell University. Email: ann34@cornell.edu

1 Introduction

Developing economies are characterized by a large number of small firms, the majority of which operate informally (La Porta & Shleifer, 2014; McKenzie, 2017). Informal firms play an important role in sustaining household incomes, generating employment, absorbing surplus labor, and fostering entrepreneurship; yet they face significant constraints concerning growth, productivity, and formal market integration (De Soto, 1989; La Porta & Shleifer, 2008). Following the conceptual framework of Ulyssea (2018), firm formality can be understood as a multi-dimensional and endogenous choice. Firms may comply with certain regulatory requirements—such as government registration or universal labor contracts—while avoiding others. This selective compliance reflects firms' cost-benefit tradeoffs (Fietz et al., 2026). Compared to fully formal firms that are government-registered and adhere to labor standards, firms operating along at least one dimension of informality (i.e., partially formal or fully informal firms) are generally smaller and less productive (La Porta & Shleifer, 2008, 2014; Benjamin & Mbaye, 2012; Amin & Okou, 2020). These firms typically face restricted access to credit markets and are frequently excluded from public procurement and support programs that promote firm growth and productivity (Perry, 2007). Despite these constraints, informality offers flexibility and cost advantages by allowing firms to bypass taxes and regulations; yet these private gains come at the cost of unfair competition, market fragmentation, lower tax revenues, and weaker provision of public goods and services (Perry, 2007; Prado, 2011; La Porta & Shleifer, 2014).

To address pervasive informality, governments often implement policies designed to promote firm formalization, thereby improving regulatory compliance, expanding the tax base, and stimulating economic growth. However, prior studies show that even when barriers to registration are reduced or incentives (like subsidies or free services) are provided, firm formalization rates remain extremely low (Bruhn & McKenzie, 2014). One reason is that the perceived benefits of firm formalization—such as improved access to formal credit or eligibility for government contracts—may not outweigh its costs, including higher taxes, increased labor expenses, and administrative burdens (De Soto, 1989; Perry, 2007; Bruhn & McKenzie, 2014; La Porta & Shleifer, 2014; Benhassine et al., 2018). These tradeoffs are especially salient when regulations link compliance requirements to employment thresholds. Such policies introduce discrete increases in compliance costs that create strong incentives for firms to adjust their behavior. Consequently, this regulatory design can induce size-dependent distortions that discourage firm growth and sustain the prevalence of small firms in developing economies.

This paper analyzes how Vietnamese firms respond to an updated formalization policy anchored in a pre-existing firm-size threshold. Under Article 119 of Vietnam's Labor Code 2012, firms with 10 or more formal paid (FP) employees are required to develop, register, and update internal work regulations (IWRs) with the government. The sole substantive change in this revised policy was a five- to 10-fold increase in the financial penalty for non-compliance. Using biennial panel data from the UNU-WIDER Vietnam Small- and Medium-Enterprise (SME) database—which covers private manufacturing firms from 2005 to 2015—I evaluate how the higher financial penalty affects firms at the threshold, including any adjustments in formal employment, labor composition, and formality status. Moreover, I assess spillover effects on firm revenue, profitability, productivity, and costs to

determine if regulatory avoidance imposes broader economic consequences. These data also support classifying firms as fully formal, partially formal, or fully informal based on their registration status and use of formal labor contracts. This enables an analysis of whether firms selectively comply with only *certain* dimensions of firm formality to minimize regulatory compliance costs.

To interpret firms' responses to the higher threshold-induced cost, I develop a profit maximization model in which firms jointly choose their labor composition—across FP, informal paid, and unpaid workers—and formality status subject to compliance costs. The model underscores that the policy introduces a discontinuous jump in expected costs at the threshold, generating three testable predictions. First, firms may "bunch" just below the 10-FP employee threshold when remaining below it yields a higher payoff than any crossing strategy (Theorem 1). Second, firms just below the threshold may substitute toward unregulated labor to maintain effective labor input and evade regulation (Theorem 2). Third, firms just above the threshold may adopt hybrid formality strategies: registering with the government while relying on informal workers; hence, they reduce compliance costs without constraining their size (Theorem 3). These predictions guide the empirical analysis.

Methodologically, this study employs a difference-in-discontinuities (DiDisc) design, which combines the cross-sectional precision of regression discontinuity design (RDD) with the temporal variation of difference-in-differences (DiD). This strategy isolates the causal effect of the higher financial penalty by comparing firms just above and below the regulatory threshold before and after policy implementation. Conventional RDD is inappropriate in this setting since the running variable—the number of FP employees—is discrete, highly left-skewed, and characterized by a pronounced mass point at zero. A standard DiD approach is also ill-suited as it averages across heterogeneous firms and cannot cleanly recover the policy-relevant threshold margin. The analysis therefore implements the DiDisc design using a local randomization procedure that selects optimal symmetric and asymmetric windows around the cutoff based on covariate balance tests. These data-driven bandwidths define estimation samples in which pre-policy firm characteristics are statistically indistinguishable across the threshold, allowing for credible causal inference on firms operating at the margin.

Did firms endogenously adjust their formal labor force to circumvent compliance requirements at the 10-FP employee threshold? To evaluate this prediction from Theorem 1, I apply McCrary (2008) density tests to assess whether the firm distribution around the threshold changed between the pre- and post-policy periods using the optimal windows. Results show no significant discontinuity post-policy despite the increased financial penalty. The absence of bunching behavior suggests that most firms did not find it optimal to manipulate their FP employment to locate just below the threshold. Firms instead adapted along alternative margins, consistent with Theorems 2 and 3.

Indeed, subsequent analysis leveraging the DiDisc methodology reveals distinct behavioral responses to the policy on either side of the threshold. Firms just *below* the 10-FP employee threshold increased their reliance on unpaid workers, likely reflecting an effort to maintain operational capacity while remaining outside the regulatory boundary to avoid compliance costs. This is consistent with Theorem 2 which predicts higher threshold-induced costs make threshold-excluded labor more attractive. Alternatively, firms just *above* the threshold decreased their use of unpaid workers and

experienced significant increases in gross profit and labor productivity. This suggests that, despite facing higher compliance costs, these firms experienced efficiency gains. However, they did not fully formalize. Consistent with Theorem 3, these firms transitioned to partial formality: registering with the government while relying on informal labor. This behavior reflects strategic non-compliance in which firms selectively satisfy the more observable dimensions of formality but avoid the costlier or less enforceable dimensions. Such selective non-compliance may reduce regulatory exposure while preserving operational flexibility, consistent with evidence from Fietz et al. (2026) which document selective within-firm worker formalization based on cost-benefit considerations. In the Vietnamese context, these qualitative patterns retain some support after multiple-hypothesis adjustments.

Several studies analyze the distortionary effects of size-dependent government policies on firms (Garibaldi et al., 2004; Guner et al., 2006; Schivardi & Torrini, 2008; Candela, 2013; Gourio & Roys, 2014; Garicano et al., 2016; Benedek et al., 2017; Dabla-Norris et al., 2018; Amirapu & Gechter, 2020; Mulligan, 2020; Bertrand et al., 2021; Padmakumar, 2022; Qian & Vereshchagina, 2022; Aghion et al., 2023). This literature consistently finds that such policies impact firm growth, productivity, resource allocation, innovation, and gender discrimination in developed and developing economies. Garibaldi et al. (2004) and Schivardi and Torrini (2008) show Italian firms close to a 15-employee threshold—which triggers stricter dismissal protections—reduce their growth propensity. Gourio and Roys (2014) and Garicano et al. (2016) find similar effects in France, where firms with more than 50 employees face additional labor requirements. In India, Amirapu and Gechter (2020) show a 10-employee threshold raises labor costs and induces bunching while Padmakumar (2022) finds a 100-employee threshold drives substitution toward capital and temporary labor inputs.

This research introduces three contributions to the literature on firm responses to size-dependent government regulations. First, it leverages a unique policy environment to provide evidence on the deterrence effect of an explicit financial penalty associated with a threshold-based labor policy. Prior studies demonstrate that firms tend to strategically adjust their size and labor composition to avoid crossing costly regulatory thresholds (Garibaldi et al., 2004; Schivardi & Torrini, 2008; Garicano et al., 2016; Aghion et al., 2023; Padmakumar, 2022). However, these studies focus on thresholds that generate *implicit* compliance costs—such as administrative burdens, tax obligations, or mandatory labor protections. None examine the behavioral response to an *explicit* economic consequence tied to non-compliance with a firm-size threshold. Comparatively, Vietnam's Labor Code 2012 includes a revised provision that increases the financial penalty for non-compliance five- to 10-fold for firms with 10 or more FP employees who do not submit IWRs. This setting provides a distinct empirical opportunity to evaluate how firms respond to formalization incentives when non-compliance is not merely symbolic but also economically costly. It isolates the deterrence effect of financial penalties by contrasting responses to perceived economic costs with those to general compliance burdens.

Second, this paper advances empirical strategies for studying firm-level distortions by applying a DiDisc design. Although this method has recently gained traction (Takahashi, 2024), it remains underused relative to more established empirical approaches. Several studies (Gourio & Roys, 2014; Garicano et al., 2016; Padmakumar, 2022) employ structurally calibrated models to stimulate firm

responses under counterfactual policies. While useful for capturing general equilibrium dynamics, such models depend on strong functional form assumptions and extensive data, often unavailable in developing countries. Existing studies also use reduced-form techniques but they do not integrate cross-sectional and temporal variation in a unified framework. By contrast, the DiDisc identification strategy used here enables direct estimation of the causal effects of policy changes at size-based thresholds without simulation or restrictive modeling assumptions. This is advantageous in settings like Vietnam, where localized policy variation and repeated survey rounds can be credibly exploited.

Lastly, this research enhances understanding of firm informality by adopting a multi-dimensional definition based on both government registration and labor contract compliance. Prior research uses binary classifications, typically tied to registration status alone (Perry, 2007; La Porta & Shleifer, 2008; Dabla-Norris et al., 2018; Padmakumar, 2022) which overlooks variation in *how* firms comply with formalization policies. Guided by the conceptual framework in Ulyssea (2018), I distinguish among fully formal, partially formal, and fully informal firms to identify the intermediate stages of formalization—such as formal firms that operate informally or vice versa. By capturing these hybrid formality strategies, the analysis underscores forms of strategic non-compliance that dichotomous definitions obscure. This approach provides a more nuanced and policy-relevant characterization of how firms adjust their formality status in response to threshold-based labor policies.

The remainder of this paper proceeds as follows: Section 2 introduces the country context and the threshold-based labor policy. Section 3 describes the data and variable construction. Section 4 presents the profit maximization model and its testable predictions. Section 5 outlines the empirical strategy, including the DiDisc design, the outcome variables, and the procedure for selecting optimal bandwidths. Section 6 reports the results and explores mechanisms. Section 7 concludes.

2 Country & Policy Setting

Since Vietnam's liberalization movement in 1986, the country has transitioned from a centrally planned economy to a market-oriented one characterized by rapid industrialization and sustained economic growth (Rand & Tarp, 2020). As part of the transition, the government prioritized small- and medium-enterprise (SME) development and business environment reform to catalyze structural change. In 2020, SMEs accounted for 95 percent of all firms, half of the labor force, and 40 percent of GDP in Vietnam (Rand & Tarp, 2020). Yet despite policy efforts, significant constraints—lack of a skilled labor force, limited credit access, and an excessive informal sector—continue to hinder SME entrepreneurial activity and growth (Angelino et al., 2021). Consequently, many SMEs still straddle the boundary between formality and informality, limiting their full economic potential.

A specific attempt of the Vietnamese government to directly encourage firm formalization among SMEs is Decree No. 10/2012/QH13, also known as Labor Code 2012. It is a detailed legal framework that governs labor relations including employment contracts, dispute resolution, wages, and working conditions. Passed in June 18, 2012 and then implemented in May 1, 2013, it is designed to protect workers' rights and promote fair labor practices across firms. Article 119 of Labor Code

2012—hereafter referred to as the 2012 employee threshold policy—specifically requires firms with at least 10 employees to prepare, issue, and then submit internal work regulations (IWRs) to a local Ministry of Labor, Invalids, and Social Affairs (MOLISA). According to Article 3 of Labor Code 2012, an "employee" is defined as someone "15 years or older, has the ability to work, works under a labor contract, is paid with wage and is managed and controlled by an employer." Thus based on the 2012 employee threshold policy, a firm's size is determined by the number of working-age individuals who are both formally contracted and paid. Unpaid workers and paid workers without a formal contract (i.e., informal workers) are not considered part of a firm's labor force according to the government. For simplicity, workers that satisfy the employee definition outlined in Labor Code 2012 (including the 2012 employee threshold policy) will be referred to as "formal paid (FP) workers" in this paper. While this 10-FP employee threshold is an exogenous policy parameter, a firm's decision to adjust their size in terms of its number of FP workers is endogenous.

The 2012 employee threshold policy dictates that IWRs submitted by firms must be consistent with current labor laws and include employee protocols related to workplace order, occupational safety, working hours, and disciplinary measures specific to their firm. Any firm that fails to create IWRs when it employs at least 10 FP workers, to register such IWRs with the appropriate provincial MOLISA, or to update expired IWRs is liable to pay a financial penalty contingent on inspection.[1] The economic consequence associated with the 2012 employee threshold policy ranges from 5 to 10 million Vietnamese Dong (VND) (i.e., 240-475 USD) per infraction.[2]

The previous version of this legal framework was Labor Code 1994 which included the same 10-FP employee threshold requirement in its respective Articles 82 and 83; however, the financial penalty for non-compliance was 1 million VND (i.e., 47 USD).[3] Thus, the updated provision imposes a financial penalty that is five- to 10-times larger. Notably, this provision is the *only* one in Labor Codes 1994 and 2012 that explicitly links regulatory obligations to firm size in terms of the number of employees, regardless of how "employee" is defined. This enables a clearer attribution of observed changes in firm size to this provision itself, reducing the risk of confounding with other regulations.

3 Data

This research uses panel data from the UNU-WIDER Vietnam Small- and Medium-Enterprise (SME) database which includes private formal and informal manufacturing firms surveyed biennially from 2005 to 2015. These firms are sampled from 10 Vietnamese provinces and distributed across roughly 18 manufacturing sectors.[4] Data was collected from June to August for each survey year for approximately 2,500 firms. Each survey captures firm data across two distinct time frames. The

[1]The fine associated with a firm failing to comply with the 2012 employee threshold policy is specified in Article 15 of Decree No. 95/2013/ND-CP, which was implemented on October 10, 2013.

[2]Conversion to USD used the 10/10/2013 VND to USD exchange rate, coinciding with the day that Decree No. 95/2013/ND-CP was implemented.

[3]The fine associated with a firm failing to comply with Articles 82 and 83 of Labor Code 1994 is specified in Article 20 of Decree No. 38/CP, which was implemented on July 1, 1996. Conversion to USD used the 10/9/2013 VND to USD exchange rate, aligning with the day *before* Decree No. 95/2013/ND-CP was implemented.

[4]Provinces were not selected randomly but included the country's main urban cities and specific rural areas.

first includes modules that document firm characteristics at the time of the survey, such as formality status, owner or manager attributes, sales structure, investment activity, credit history, production technology and practices, and perceived constraints. The second focuses on retrospective, year-end data, including firms' annual economic accounts (e.g., sales revenue, gross profit, labor costs, debt, etc.), employment levels, and labor force composition for the previous calendar year.

Additionally, the dataset supports classification of firms into fully formal, partially formal, and fully informal dimensions.[5] Firms that have an Enterprise Code Number (ECN) and provide all (i.e., 100 percent) of their full-time workers with a formal, written labor contract are considered fully formal.[6] These firms are both registration and labor contract compliant. By contrast, firms that do not have an ECN and fail to provide all (i.e., less than 100 percent) of their full-time workers with a formal, written labor contract are considered fully informal. These firms are neither registration nor labor contract compliant. Firms that comply with only one of the two are considered partially formal. Specifically, firms that have an ECN but fail to provide a formal, written labor contract to all their full-time workers are referred to as formal firms that operate informally while firms that do not have an ECN but do provide a formal, written labor contract to all their full-time workers are referred to as informal firms that operate formally.[7] Hence a firm's formality status is endogenous because it is determined by the firm's *own* registration and labor contract compliance decisions.

The survey samples are stratified by ownership type to capture the full range of legal structures among private SMEs (e.g., registered households, cooperatives, limited liability companies, etc.). The sampling scheme for this dataset is based on a representative sample of registered firms drawn from the General Statistics Office of Vietnam enterprise census information. However, this census only includes firms registered with the government. Unregistered firms that are a part of the sample were identified "on-site" as they operated along-side registered firms that were surveyed. A random selection of these unregistered firms were included in the sample in each survey year. Therefore, the sample of informal firms is not representative of the country's informal manufacturing sector. This limitation affects the external validity of the results. Nevertheless, it is not a major concern as the analysis only focuses on within-in firm changes over time, causal relationships, and mechanisms—not generalizations to Vietnam's informal manufacturing sector.

In reference to the 2012 employee threshold policy, the survey does not explicitly ask firms to report their exact number of formal paid (FP) workers for the previous calendar year's year-end. Instead, I estimate this metric independently using the survey's question structure and sequencing. Enumerators instructed firms to separately report their number of full-time, part-time, and casual workers, along with the percentage of their total workers who were unpaid and the percentage of their full-time workers with formal, written labor contracts. With this information, I estimate the

[5] I extend the conceptual framework from Ulyssea (2018) by operationalizing firm formality along two dimensions: government registration and labor contract compliance.

[6] An Enterprise Code Number (ECN) is a unique identifier for registered firms in Vietnam, combining each firm's Business Registration Certificate and Tax Code numbers as mandated by Decree No. 43/2010/ND-CP.

[7] Appendix Table A.1 reports the year-end distribution of firms by formality status for survey years from the UNU-WIDER Vietnam SME database where government registration and labor contract compliance data are available.

number of FP *full-time* workers for each firm i at time t using Equation 1:[8]

$$FP_{it} = C_{it} \times FT_{it} \times (1 - U_{it}) \tag{1}$$

where $C_{it} \in [0, 1]$ is the percentage of full-time workers with a formal, written labor contract; FT_{it} is the number of full-time workers; and, $U_{it} \in [0, 1]$ is the percentage of total workers that are unpaid.[9] Note that total workers is the summation of full-time, part-time, and casual workers.[10]

Article 3 of Labor Code 2012 defines "employee" without any reference to the *amount* of time worked, making no distinction between full-time, part-time, or casual workers. Because the data can only determine the percentage of *full-time* workers that have a formal, written labor contract, only the number of FP *full-time* workers can be estimated for the sample (see Equation 1). The dataset does not include survey questions that can be used to estimate the number of formally contracted part-time or casual workers. Although excluding part-time FP workers may slightly underestimate a firm's size in terms of its total FP employment—which defines the firm-size threshold in the 2012 employee threshold policy—this does not undermine the analysis.[11] This is because FP full-time workers have a higher probability of driving firm compliance decisions and performing compliance tasks compared to their part-time counterparts given their more stable and consistent roles in the firm as well as their greater responsibility and involvement in operational and regulatory processes (Mulligan, 2020; Morikawa, 2023). Hence, using the number of FP full-time workers still strongly aligns with the policy's intent to capture a firm's substantive employment obligations.

Figure 1 illustrates the timing of Labor Code 2012's passage and implementation relative to the UNU-WIDER Vietnam SME survey years. The timeline raises a potential concern about using year-end 2012 and 2014 firm data to represent the pre- and post-policy periods, respectively, due to the possible anticipatory affects during the interim period. Given the year-long gap shown in Figure 1, firms may have preemptively adjusted their behavior *before* the 2012 employee threshold policy's formal enforcement to mitigate the risk of future financial penalties associated with non-compliance.

To select the appropriate pre-policy period (i.e., year-end 2010 or 2012), I calculate the number of (i) FP, (ii) informal paid, and (iii) unpaid full-time workers for each firm within the year-specific rounded samples.[12] I then further restrict these rounded samples to firms with fewer than 50 total

[8] Appendix B provides the full mathematical derivation.

[9] Equation 1 assumes that unpaid workers are distributed proportionately across full-time, part-time, and casual work-time categories; therefore, the reported share of total workers who are unpaid is applied to full-time workers.

[10] Labor force classification guidelines referenced by enumerators during UNU-WIDER Vietnam SME data collection: "unpaid labour refers to those who do not receive wages or other remuneration directly related to the work they perform...Full-time is considered as a person working more than than 183 days per year, more than 20 days a month and more than 20 hours per week. Part-time is a person working under 20 hours per week and/or between 5 and 20 days a month. Casual labour [is] the residual that is working on average less than 5 days a month and/or few hours some weeks a month" (Institute of Labour Science and Social Affairs & the University of Copenhagen, 2009, pg. 21).

[11] As further reassurance, McCrary (2008) density tests performed in Section 6.1 provide no evidence of post-policy manipulation in firms' number of FP full-time workers at the 10-FP employee threshold, suggesting that firms report their employment levels and related data accurately.

[12] Appendix C describes the construction of the rounded samples as an intermediate step in deriving the analytical sample from the Vietnam SME dataset.

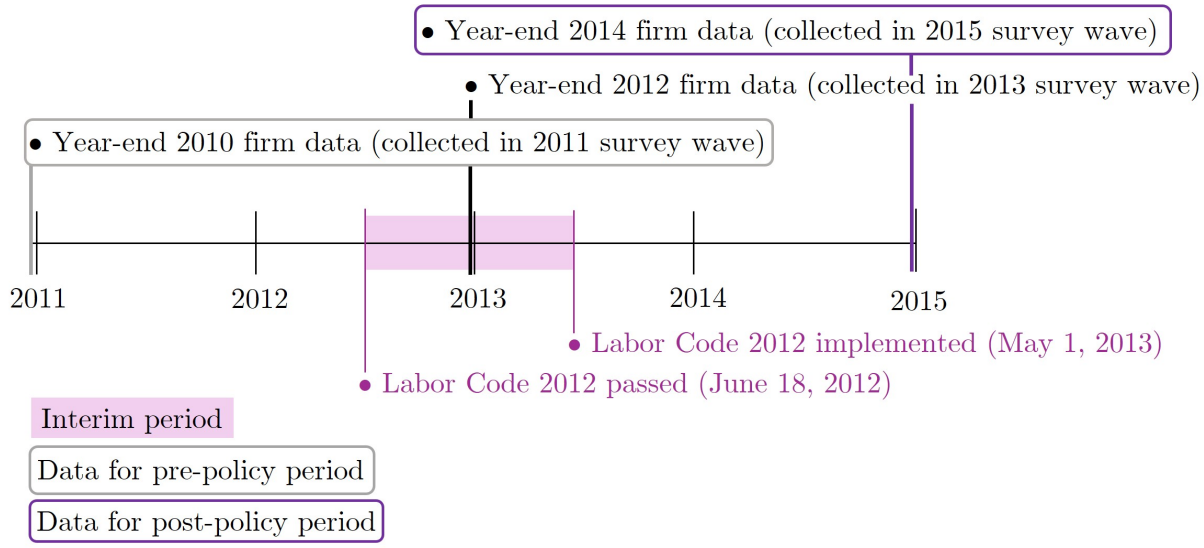


Figure 1: Policy & Data Collection Timelines

workers across years, consistent with the World Bank's definition of micro- and small- enterprises.[13] Resulting year-end averages for each of the three full-time labor categories are plotted in Figure 2.

Figure 2 displays the average number of full-time workers by formal contract and pay status from year-end 2008 to 2014 for these restricted rounded samples. The average number of FP full-time workers remains stable from year-end 2008 (i.e., 3.0) to 2010 (i.e., 3.1), declines modestly from 3.1 in year-end 2010 to 2.7 in year-end 2012, and then levels off again through year-end 2014 (i.e., 2.6). Using two-sided tests of mean differences with standard errors clustered by firm, I find no statistical difference in means between year-end 2008 and 2010 ($p = 0.9432$) or between year-end 2012 and 2014 ($p = 0.3930$). However, both the declines from year-end 2010 to 2012 ($p = 0.0284$) and from year-end 2010 to 2014 ($p = 0.0081$) are statistically significant. By comparison, the average number of informal paid full-time workers declines more sharply and steadily over time while the average number of unpaid full-time workers remains relatively constant. These trends suggest that selecting year-end 2010 as the pre-policy period is prudent, as it precedes potential behavioral adjustments firms may have made in anticipation of the 2012 employee threshold policy's implementation.[14]

To contextualize the increase in the financial penalty tied to the 2012 employee threshold policy, Table 1 presents the fine as a percentage of average year-end sales revenue, gross profit, and labor costs in the restricted rounded samples across the pre-policy, interim, and post-policy periods. For the pre-policy and interim periods, the calculations apply the 1 million VND fine specified under Labor Code 1994 because the revised fine introduced by Labor Code 2012 had not yet taken effect.

[13] See https://ieg.worldbankgroup.org/sites/default/files/Data/Evaluation/files/SME_Synthesis.pdf for details.

[14] Repeating this timing check within the optimal windows defined in Section 5.3 yields the same qualitative pattern. Paired t-tests show no significant difference in average FP full-time employment between year-end 2008 and 2010, a significant decline between year-end 2010 and 2012, and no significant change between year-end 2012 and 2014. This supports treating year-end 2012 as an interim period rather than a pre-policy period. See Appendix Figure A.1.

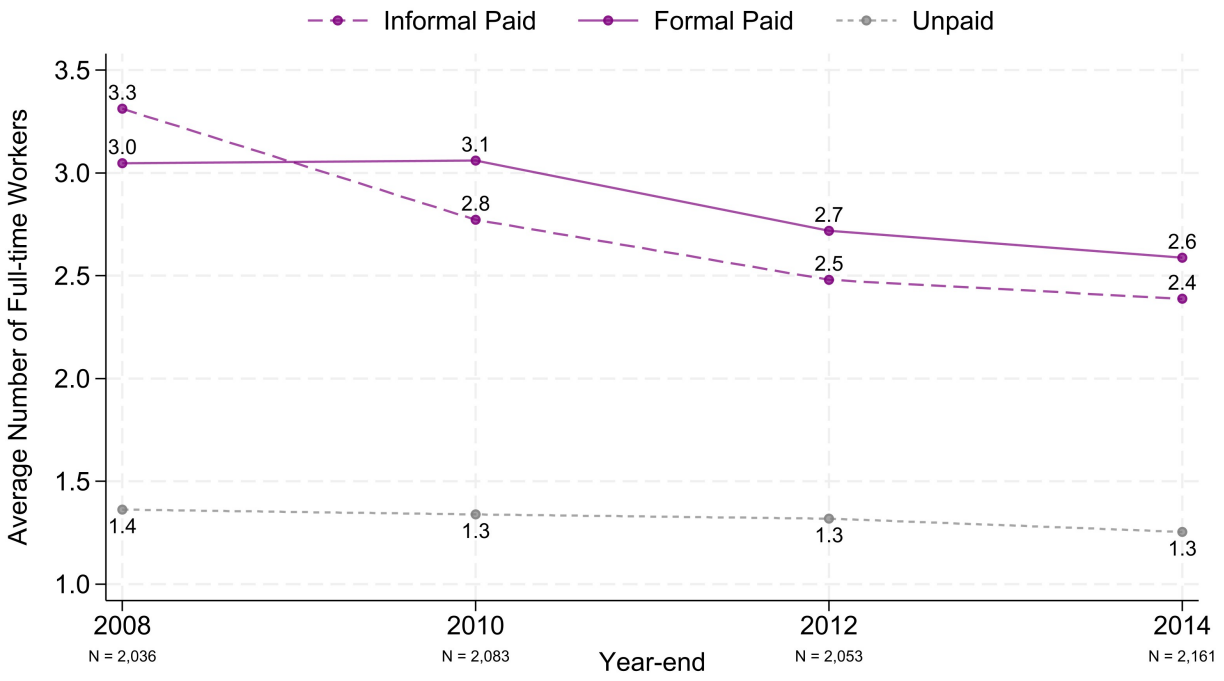


Figure 2: Average Number of Full-time Workers by Labor Category for Restricted Rounded Samples by Year-End Period

Notes: This figure shows the average number of (i) informal paid, (ii) formal paid, and (iii) unpaid full-time workers at each year-end period within the rounded samples. These samples are also restricted to firms with fewer than 50 total workers across years, consistent with the World Bank's definition of micro- and small-enterprises. Each year-end period's sample size is indicated below the corresponding year.

For the post-policy period, the percentages instead reflect the increased financial penalty of 5 to 10 million VND established under Labor Code 2012, thus the range of values shown in Table 1. The findings underscore the substantial rise in the economic burden of non-compliance, illustrating how the 2012 employee threshold policy raised the costs of informality for firms.[15]

Table 1: Financial Penalty Relative to Firm Economic Account Metrics for Restricted Rounded Samples by Year-end Period

	Pre-Policy Period	Interim Period	Post-Policy Period
Percent of Average Annual	2010	2012	2014
Sales revenue	0.05%	0.04%	0.18–0.37%
Gross profit	0.30%	0.32%	1.18–2.37%
Labor costs	0.48%	0.42%	1.83–3.66%
Firms	2,083	2,053	2,161

Notes: Data reflect year-end values (i.e., December 31st). Only firms from the rounded samples with fewer than 50 total workers across years are included, which is consistent with the World Bank's definition of micro- and small-enterprises. Gross profit is the difference between sales revenue and total costs—including raw materials, indirect expenses, and labor costs (e.g., wages, allowances, recruitment, training, as well as health, social, and unemployment insurance).

Figure 3 complements these findings by showing changes in inspection activity across the pre-policy, interim, and post-policy periods within the restricted rounded samples. Between the pre- and post-policy periods, average total inspections per firm increased significantly from 0.16 to 0.66 while average policy inspections—those targeting labor and tax law compliance—rose significantly from 0.07 to 0.30.[16] Table 1 indicates that the higher financial penalty greatly increased the cost of non-compliance while Figure 3 shows that this cost was actively enforced rather than merely theoretical. The concurrent escalation in policy inspections make the threat of detection and sanction salient and credible. Together, these shifts created a dual mechanism of deterrence: firms faced higher expected costs of informality from the increased financial penalty *and* from the heightened enforcement risk. This institutional environment made the 2012 employee threshold policy more binding for firms.

4 Firm Profit Maximization Problem

The following profit maximization model builds on frameworks with heterogeneous productivity and size-dependent regulatory costs—as developed by Garicano et al. (2016), Dabla-Norris et al. (2018), and Ulyssea (2018)—to characterize firm responses to Vietnam's 2012 employee threshold policy. Following Garicano et al. (2016) and Dabla-Norris et al. (2018), the model abstracts from endogenous firm entry and focuses on firms' size, labor composition, and formality status choices

[15] Appendix Table A.2 repeats these calculations within the optimal windows defined in Section 5.3, separately for firms just below and just above the 10-FP employee threshold. The revised financial penalty generally represents a larger share of year-end economic account metrics for firms just *below* the cutoff within these estimation samples.

[16] Policy inspections include announced and unannounced government visits to verify labor and tax law compliance. Non-policy inspections include technical compliance (i.e., workplace standards) or investigations following accidents.

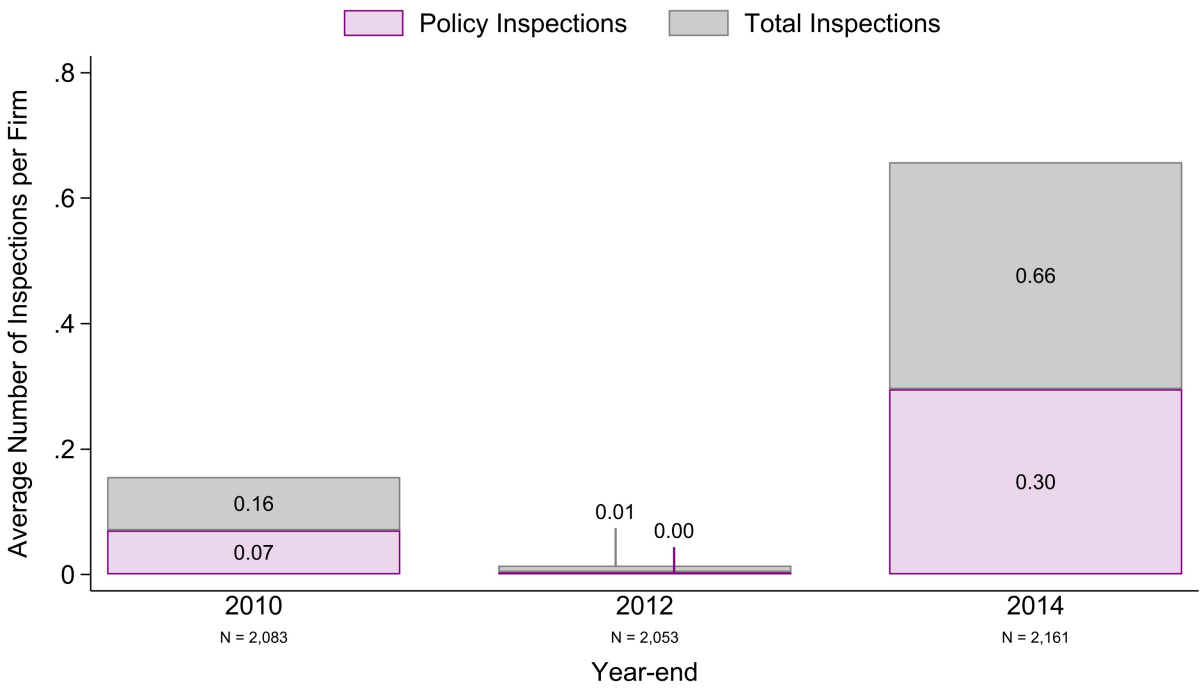


Figure 3: Average Number of Inspections per Firm for Restricted Rounded Samples by Year-End Period

Notes: The figure shows the average number of total (gray) and policy (purple) inspections per firm across multiple year-end periods. Only firms from the rounded samples with fewer than 50 total workers across years are included, which is consistent with the World Bank's definition of micro- and small-enterprises. Each period's sample size is indicated below the corresponding year.

conditional on market participation. The model is static but regime-indexed, allowing comparative static predictions based on the higher financial penalty under the 2012 employee threshold policy.

As discussed in Section 2, the 2012 employee threshold policy is only binding for firms with at least 10 formal paid (FP) workers.[17] Firms below this threshold avoid the internal work regulation (IWR) requirement while firms either at or above it must submit IWRs or risk incurring a financial penalty for non-compliance. Since informal paid and unpaid workers fall outside the government's definition of "employee," firms can respond by adjusting FP employment or substituting toward labor arrangements *outside* the regulated employment margin.

4.1 Technology, Labor Inputs, and Firm Formality Status

A representative firm i operates in a homogeneous goods sector and chooses its labor composition and formality status. These decisions are jointly determined since labor composition affects whether

[17]The model follows the 2012 employee threshold policy's definition of "employee" and thus abstracts from workers' full-time, part-time, and casual status. In the empirical analysis, however, data limitations restrict the measurement of labor categories to full-time workers (see Section 3).

the firm crosses the 10-FP employee threshold while formality status affects the costs and potential productivity benefits of compliance. Firm i chooses across three labor categories:

- $L_{FP,i}$: number of formal paid (FP) workers;
- $L_{IP,i}$: number of informal paid workers;
- $L_{U,i}$: number of unpaid workers.

Firm i's total labor is therefore:

$$L_i = L_{FP,i} + L_{IP,i} + L_{U,i}.$$

Firm i's formality status has two dimensions:

- $R_i \in \{0, 1\}$: whether firm i is registered with the government (1) or not (0);
- $K_i \in \{0, 1\}$: whether firm i is fully labor contract compliant with regards to providing formal, written labor contracts to all its workers (1) or not (0).[18]

Firm i's formality status is characterized by the pair (R_i, K_i) and it is defined as:

$$\mathcal{F}_i = R_i + K_i$$

where $\mathcal{F}_i = 0$ indicates a fully informal firm, $\mathcal{F}_i = 1$ indicates a partially formal firm, and $\mathcal{F}_i = 2$ indicates a fully formal firm. Table 2 summarizes how government registration and labor contract compliance map to each firm formality status.[19]

Table 2: Firm Choice Set for Each Formality Status

	Registered with the Government (R_i)	Formal, Written Labor Contracts For All Workers (K_i)	Formality Status ($R_i + K_i = \mathcal{F}_i$)
Fully formal	1	1	2
Partially formal			
Formal but operates informally	1	0	1
Informal but operates formally	0	1	
Fully informal	0	0	0

Given these labor composition and formality choices, firm i's output depends on effective labor input and a formality-productivity shifter. Effective labor input is a weighted sum of different labor categories' productivity:

$$L_i^{eff} = L_{FP,i} + \theta L_{IP,i} + \delta L_{U,i}, \tag{2}$$

[18]The empirical analysis measures labor contract compliance only among full-time workers because the data do not report formal, written labor contract coverage for part-time or casual workers (see Section 3).

[19]The distinction between government registration and labor contract compliance mirrors the dual margins of firm formality discussed in Ulyssea (2018).

with $0 < \delta < \theta < 1$. FP workers are normalized to be the most productive labor input, followed by informal paid workers and then unpaid workers. This hierarchy reflects that FP workers are modeled as more effective since formal employment relationships are associated with greater stability, accountability, and worker attachment; informal paid and unpaid workers are assumed to provide flexibility and lower direct compliance costs, but contribute less effective labor per worker.[20]

I capture productivity benefits from firm i's formality status through:

$$\Gamma(R_i, K_i) = e^{\gamma_R R_i + \gamma_K K_i} \tag{3}$$

where $\gamma_R \geq 0$ and $\gamma_K \geq 0$.[21] This productivity shifter is normalized to one for fully informal firms, such that $\Gamma(0,0) = 1$. Holding other inputs fixed, government registration multiplies firm i's output by e^{γ_R} while labor contract compliance multiplies firm i's output by e^{γ_K}. Therefore a fully formal firm receives the multiplier $e^{\gamma_R + \gamma_K}$. Broadly, $\Gamma(R_i, K_i)$ captures the productivity, market access, and organizational benefits of compliance across these firm formality dimensions (La Porta & Shleifer, 2008, 2014; Fietz et al., 2026), without assuming these benefits exceed compliance costs.

Firms differ in productivity, denoted by $A_i > 0$. Output price is normalized to 1 so output can be interpreted as revenue. Therefore output is produced according to:

$$y_i = A_i \Gamma(R_i, K_i)(L_i^{eff})^{\alpha} \tag{4}$$

where $\alpha \in (0,1)$ captures the diminishing returns to effective labor input. The productivity shifter $\Gamma(R_i, K_i)$ allows firm i's formality status to affect firm-level productivity separately from the worker-level productivity differences embedded in L_i^{eff}.

4.2 Costs and the Policy Threshold

Firm costs include labor costs, registration costs, labor contract compliance costs, and threshold-induced regulatory costs. Let $C_R > 0$ denote the fixed cost of government registration and $C_K > 0$ denote the fixed cost of full labor contract compliance. Formal paid labor costs are $C_{FP} = w_{FP} + b_{FP}$ per worker where w_{FP} is the wage and b_{FP} captures mandatory benefits (e.g., unemployment, social, and health insurance contributions). Informal paid labor costs are $C_{IP} = w_{IP}$ per worker. Unpaid labor does not require a wage payment, but it is not costless. Its use imposes an increasing and weakly convex non-monetary cost $\psi(L_{U,i})$ which captures the disutility, supervision burden, reduced operational efficiency, or opportunity cost associated with relying on unpaid workers—such as the limited availability of household labor. Formally, $\psi'(L_{U,i}) > 0$ and $\psi''(L_{U,i}) \geq 0$.

[20] This ordering is consistent with theoretical models (Lewis et al., 1954; Amaral & Quintin, 2006; Ulyssea, 2018) and empirical evidence from developing countries (La Porta & Shleifer, 2008; McCaig & Pavcnik, 2013; Fietz et al., 2026) on productivity differences between formal and informal firms and labor arrangements.

[21] The exponential form is a parsimonious reduced-form assumption used for tractability rather than a micro-founded specification. It ensures that $\Gamma(R_i, K_i) > 0$ and allows γ_R and γ_K to be interpreted as log output premia.

Registration status R_i is defined for all firms regardless of their level of FP employment. Under the 2012 employee threshold policy, only firms with $L_{FP,i} \geq 10$ are required to submit IWRs. For firms subject to this requirement, $T_i = 1$ denotes submission and $T_i = 0$ denotes non-submission. Submitting IWRs, in turn, requires government registration:

$$L_{FP,i} \geq 10 \wedge T_i = 1 \Rightarrow R_i = 1.$$

Hence $R_i = 0$ does not imply that firm i is below the threshold. An unregistered firm can employ at least 10 FP workers (i.e., $L_{FP,i} \geq 10$) but cannot submit IWRs (i.e., $T_i = 0$) unless it registers.

If a firm at or above the threshold submits IWRs, it incurs a fixed threshold-compliance cost $C_T > 0$. If this firm does not, it faces an expected financial penalty. Let $r \in \{0, 1\}$ denote the policy regime, where $r = 0$ is the pre-policy regime and $r = 1$ is the post-policy regime. The financial penalty satisfies $P_1 > P_0$. The expected financial penalty for IWR non-submission is $\phi(R_i)P_r$. Let $\phi(R_i)$ be the perceived probability of inspection under the 2012 employee threshold policy:

$$\phi(R_i) = \phi_L + (\phi_H - \phi_L)R_i, \qquad 0 \leq \phi_L < \phi_H \leq 1$$

where ϕ_L and ϕ_H are the perceived probabilities of inspection for unregistered (i.e., $R_i = 0$) and registered (i.e., $R_i = 1$) firms, respectively. Since $\phi_L < \phi_H$, registered firms face a higher expected financial penalty than unregistered firms (i.e., $\phi_H P_r > \phi_L P_r$). The model isolates the financial penalty channel by allowing $P_1 > P_0$ while holding perceived inspection risk and other primitives fixed, avoiding the assumption that firms anticipated the realized post-policy increase in inspections.

Full labor contract compliance requires that *all* workers have formal, written labor contracts. Thus, if $K_i = 1$, firm i cannot employ informal paid workers:

$$K_i = 1 \quad \Rightarrow \quad L_{IP,i} = 0.$$

By contrast, $L_{U,i}$ may include workers with or without formal, written labor contracts. Given that unpaid workers are excluded from the 10-FP employee threshold, they enter the model as a single threshold-excluded labor input regardless of contract status. Their contract status matters only for firm i's labor contract compliance: $K_i = 1$ requires all workers, including unpaid workers, to have formal, written labor contracts whereas $K_i = 0$ requires *at least one* worker to lack such a contract.

A key feature of the model is that threshold-specific compliance (T_i) and labor contract compliance (K_i) are separate decisions. Namely, conditional on crossing the 10-FP employee threshold, T_i captures whether firm i satisfies the threshold-specific IWR requirement. By contrast, K_i captures labor contract compliance, a broader dimension of labor regulation requiring formal, written labor contracts for all workers. Hence, a firm can cross the threshold (i.e., $L_{FP,i} \geq 10$) and submit IWRs (i.e., $T_i = 1$) while remaining partially informal along the labor contract margin (i.e., $K_i = 0$).

4.3 Regulatory Strategies

Let $s \in \{B, N, F, S\}$ capture the set of regulatory strategies available to firm i in response to the increased financial penalty associated with the 10-FP employee threshold: remaining below the threshold (B), crossing the threshold without submitting IWRs (N), crossing the threshold while complying fully with both the IWR and labor contract requirements (F), and crossing the threshold while complying selectively with the IWR requirement but not the labor contract requirement (S).

For each strategy s and policy regime r, the firm's maximized payoff is denoted by $\Pi_i^s(P_r)$. Let the following denote firm i's strategy-specific choice vector:

$$x_i^s = \left(L_{FP,i}^s, L_{IP,i}^s, L_{U,i}^s, R_i^s, K_i^s, T_i^s\right).$$

For each strategy s and policy regime r, firm i solves the following constrained profit maximization problem:

$$\begin{aligned}\Pi_i^s(P_r) \equiv \max_{x_i^s \in \mathcal{X}^s} \Big\{ & A_i \Gamma(R_i^s, K_i^s) \left(L_i^{eff,s}\right)^{\alpha} - C_{FP} L_{FP,i}^s - C_{IP} L_{IP,i}^s - \psi\left(L_{U,i}^s\right) - C_R R_i^s - C_K K_i^s \\ & - C_T \cdot \mathbb{1}\{L_{FP,i}^s \geq 10 \wedge T_i^s = 1\} - \phi(R_i^s) P_r \cdot \mathbb{1}\{L_{FP,i}^s \geq 10 \wedge T_i^s = 0\}\Big\}\end{aligned} \tag{5}$$

where $\mathcal{X}^s$ is the feasible set associated with strategy s. The formality-productivity shifter $\Gamma(R_i^s, K_i^s)$ captures productivity benefits from government registration and labor contract compliance while C_R and C_K capture their associated private costs. Compliance along both dimensions is optimal when its benefits exceed the associated compliance costs and any losses from reduced labor flexibility.

The four feasible strategy sets are defined as follows.

Strategy B: Firm i remains below the threshold, so the IWR requirement is not triggered:

$$\mathcal{X}^B = \left\{x_i^B : L_{FP,i}^B < 10,\ T_i^B = 0,\ R_i^B \in \{0,1\},\ K_i^B \in \{0,1\},\ K_i^B = 1 \Rightarrow L_{IP,i}^B = 0,\ L_{U,i}^B \geq 0\right\}.$$

For notational convenience, $T_i^B = 0$ under this strategy. Because firms below the 10-FP employee threshold are exempt from the IWR requirement, this value has no effect on payoffs.

Strategy N: Firm i crosses the threshold but does not submit IWRs, so the firm faces the expected financial penalty for IWR non-submission:

$$\mathcal{X}^N = \left\{x_i^N : L_{FP,i}^N \geq 10,\ T_i^N = 0,\ R_i^N \in \{0,1\},\ K_i^N \in \{0,1\},\ K_i^N = 1 \Rightarrow L_{IP,i}^N = 0,\ L_{U,i}^N \geq 0\right\}.$$

Strategy F: Firm i crosses the threshold, submits IWRs, registers with the government, and fully

complies with labor contract regulations:

$$\mathcal{X}^F = \left\{ x_i^F : L_{FP,i}^F \geq 10,\ T_i^F = 1,\ R_i^F = 1,\ K_i^F = 1,\ L_{IP,i}^F = 0,\ L_{U,i}^F \geq 0 \right\}.$$

Under this strategy, the firm may employ unpaid workers but any such workers must have formal, written labor contracts. Note that unpaid workers are excluded from the 10-FP employee threshold.

Strategy S: Firm i crosses the threshold, submits IWRs, and registers with the government, but *does not* fully comply with labor contract regulations:

$$\mathcal{X}^S = \left\{ x_i^S : L_{FP,i}^S \geq 10,\ T_i^S = 1,\ R_i^S = 1,\ K_i^S = 0,\ L_{IP,i}^S \geq 0,\ L_{U,i}^S \geq 0 \right\}.$$

Under this strategy, $K_i^S = 0$ means that *at least one* paid or unpaid worker lacks a formal, written labor contract. If labor contract non-compliance arises only among unpaid workers, then $L_{IP,i}^S = 0$; yet $L_{U,i}^S > 0$ does not imply such non-compliance since unpaid workers may be formally contracted.

Given these four feasible strategy sets, firm i chooses the regulatory strategy s that delivers the highest maximized payoff in policy regime r:

$$s_i^*(P_r) \in \arg \max_{s \in \{B,N,F,S\}} \Pi_i^s(P_r).$$

Because $P_1 > P_0$, the higher financial penalty changes the relative payoff of the available regulatory strategies and can therefore alter firms' labor composition and formality status choices.

4.4 Comparative Statics and Testable Predictions

The model yields three comparative static predictions that guide the empirical analysis, each derived from the increase in the financial penalty for IWR non-submission from P_0 to P_1.

Lemma 1. *The increased financial penalty for IWR non-submission weakly lowers the payoff from crossing the 10-FP employee threshold without submitting IWRs (Strategy N). Because the payoffs from remaining below the threshold (Strategy B), full compliance (Strategy F), and selective compliance (Strategy S) are not directly affected by P_r, for any alternative strategy $q \in \{B, F, S\}$:*

$$\Pi_i^q(P_1) - \Pi_i^N(P_1) \geq \Pi_i^q(P_0) - \Pi_i^N(P_0).$$

Lemma 1 follows since the expected financial penalty only enters the payoff when a firm crosses the threshold and *does not* submit IWRs. The increase from P_0 to P_1 thus makes Strategy N weakly less attractive relative to remaining below the threshold, fully complying, or selectively complying. The policy change therefore does not mechanically imply bunching just below the threshold because firms may instead choose one of these alternative strategies.

Theorem 1. *Define the net payoff gain of firm i crossing the 10-FP employee threshold under strategy $q \in \{N, F, S\}$ relative to remaining below the threshold (Strategy B) as:*

$$\Delta_{i,q}(P_r) \equiv \Pi_i^q(P_r) - \Pi_i^B(P_r).$$

Firm i remains below the threshold in the post-policy regime when:

$$\max_{q \in \{N,F,S\}} \Delta_{i,q}(P_1) < 0.$$

Equivalently, remaining below the threshold is optimal when the net gains from crossing are negative under every feasible crossing strategy. If this condition holds for a positive mass of firms near the cutoff, and FP employment can be adjusted downward at sufficiently low cost, threshold avoidance may yield bunching in the firm-size distribution just below the threshold in the post-policy regime.

Theorem 1 clarifies that threshold avoidance is a firm-level response while bunching is a potential aggregate manifestation of that response. The increase in the financial penalty weakly expands the set of firms for which remaining below the threshold dominates crossing without submitting IWRs. Yet, bunching need not occur if reducing FP employment is costly, firms are *far below* the cutoff, or other strategies are more profitable. This motivates the McCrary (2008) density tests in Section 6.1 which evaluate if firms manipulate their FP employment around the threshold post-policy.

Proof. See Appendix D.1.

Theorem 2. *For firms just below the 10-FP employee threshold, the higher financial penalty weakly increases the shadow cost of adding FP employment that triggers the IWR requirement. Conditional on maintaining a target level of effective labor input, this makes threshold-excluded labor inputs—including informal paid and unpaid labor—relatively more attractive.*

Theorem 2 follows from the effective labor function in Equation 2 and the cost components embedded in the profit function in Equation 5. Conditional on firm formality status, $\Gamma(R_i, K_i)$ scales output; however, it does not change the relative effective labor contributions of FP, informal paid, and unpaid labor. FP labor is the most productive labor category but it counts toward the policy threshold. Informal paid and unpaid labor are less productive but are *excluded* from the threshold. The relevant marginal per-effective-labor unit costs across these labor categories are:

$$c_{FP} \equiv C_{FP}, \qquad c_{IP} \equiv \frac{C_{IP}}{\theta} = \frac{w_{IP}}{\theta}, \qquad c_U(L_{U,i}) \equiv \frac{\psi'(L_{U,i})}{\delta}.$$

For firms *just below* the cutoff, let $\tau_i(P_r)$ denote the regulatory wedge associated with adding FP employment when doing so triggers the IWR requirement. Firm i substitutes toward threshold-excluded labor when either informal paid or unpaid labor provides effective labor at lower cost than

FP labor inclusive of this wedge:

$$\min\left\{\frac{w_{IP}}{\theta}, \frac{\psi'(L_{U,i})}{\delta}\right\} < C_{FP} + \tau_i(P_r).$$

Since $P_1 > P_0$ weakly increases $\tau_i(P_r)$, firms seeking to maintain effective labor can limit regulated employment by instead substituting toward unregulated labor (i.e., $L_{IP,i}$, $L_{U,i}$, or both).

Proof. See Appendix D.2.

Theorem 3. *Among firms that cross the 10-FP employee threshold, selective compliance (Strategy S) is chosen in the post-policy regime when:*

$$\Pi_i^S(P_1) > \max\{\Pi_i^N(P_1), \Pi_i^F(P_1)\}.$$

The increased financial penalty makes crossing the threshold without submitting IWRs (Strategy N) less attractive. Conditional on crossing the threshold and submitting IWRs, selective compliance is preferred when the fixed cost savings and labor flexibility from avoiding full labor contract compliance exceed the productivity and organizational benefits of full labor contract compliance.

Theorem 3 characterizes selective compliance as a strategy in which firms are government-registered and submit IWRs, thereby avoiding the expected financial penalty for IWR non-submission while remaining non-compliant along the labor contract margin. Once IWRs are submitted, the expected financial penalty no longer affects the choice between full and selective compliance. Instead, firms weigh the fixed labor contract compliance cost C_K and the value of retaining informal labor against the productivity and organizational benefits from full labor contract compliance.

Let $\widetilde{\Pi}_i^s$ denote the maximized operating payoff under strategy s, defined as output net of paid labor costs and the non-monetary cost of unpaid labor. Conditional on crossing the threshold and submitting IWRs, Strategies F and S share the common costs C_R and C_T and avoid the expected financial penalty. Thus, selective compliance is preferred to full compliance when:

$$\widetilde{\Pi}_i^S - \widetilde{\Pi}_i^F + C_K > 0.$$

Selective compliance is thus especially attractive when the cost savings and labor flexibility benefits from remaining partially informal exceed the operating advantage of full labor contract compliance.

Proof. See Appendix D.3.

5 Empirical Strategy

This section presents the empirical strategy used to estimate the causal effects of the increased financial penalty introduced by the 2012 employee threshold policy on firm outcomes. Section 5.1 describes the difference-in-discontinuities (DiDisc) identification framework, Section 5.2 defines the

outcome variables, and Section 5.3 outlines the procedure for selecting optimal bandwidths around the 10-formal paid (FP) employee threshold used for DiDisc estimation.

5.1 Difference-in-Discontinuities Identification Framework

The empirical strategy employs a DiDisc approach which combines the localized identification of regression discontinuity design (RDD) with the intertemporal variation exploited by difference-in-differences (DiD) (Grembi et al., 2016; Bennedsen et al., 2022; Tramontin Shinoki et al., 2024). Though it is increasingly used in applied research, DiDisc remains only partially formalized in the economic literature, with relatively few empirical applications to date (Takahashi, 2024). The 2012 employee threshold policy increased the financial penalty for non-compliance at a sharply defined firm-size threshold, creating a cross-sectional discontinuity and a time-based shift. These conditions are ideally suited for a DiDisc approach because it identifies the causal effect of the policy change on firm behavior at the margin before *relative to* after its implementation. Given this, a conventional RDD is insufficient because the policy generates temporal—not purely cross-sectional—identifying variation at the threshold. Additionally, Appendices E.1 and E.2 show that standard DiD estimates are sensitive to the inclusion of firms farther from the threshold despite little evidence of differential pre-policy trends near the cutoff, motivating the use of the DiDisc specification in Equation 6.

$$\begin{aligned} Y_{it} = {} & \alpha + \beta_G \, Above10_{i,2010} + \beta_T \, Post_t + \delta_0 \left(Above10_{i,2010} \times Post_t \right) \\ & + \theta_X \, Center_{i,2010} + \theta_G \left(Above10_{i,2010} \times Center_{i,2010} \right) + \theta_T \left(Post_t \times Center_{i,2010} \right) \\ & + \delta_1 \left(Above10_{i,2010} \times Post_t \times Center_{i,2010} \right) + \mu_i + \epsilon_{it} \end{aligned} \tag{6}$$

In Equation 6, Y_{it} is an outcome for firm i in time t. $Center_{i,2010}$ is the running variable and measures the number of FP full-time workers employed by firm i, centered at the 2012 employee threshold policy cutoff of 10 using pre-policy values (i.e., year-end 2010). $Above10_{i,2010}$ equals 1 if firm i has 10 or more FP full-time workers pre-policy, defining the treatment group. Firms with less than 10 FP full-time workers form the control group. $Post_t$ equals 1 for the post-policy period (i.e., year-end 2014) and 0 for the pre-policy period (i.e., year-end 2010), capturing temporal variation. Firm fixed effects (μ_i) control for unobserved, time-invariant firm traits and ϵ_{it} is the error term.[22]

Equation 6's structure parallels a standard DiD design in its first line: $Above10_{i,2010}$ captures the pre-policy group difference, $Post_t$ captures the time difference across firms, and their interaction—represented by coefficient δ_0—measures the DiDisc treatment effect. This is the primary coefficient of interest, representing the change in discontinuity at the threshold following the policy. It isolates the causal effect of the 2012 employee threshold policy on outcomes for firms just *above* the threshold (i.e., the treatment group) compared to firms just *below* it (i.e., the control group) post-policy. By comparing the size of the discontinuity at the cutoff before and after the policy, the DiDisc method nets out any pre-existing jump and recovers the policy-induced shift captured by δ_0.

[22] In Equation 6, firm fixed effects absorb all time-invariant firm-level terms, including $Above10_{i,2010}$, $Center_{i,2010}$, and $Above10_{i,2010} \times Center_{i,2010}$. These terms are nevertheless shown to make the DiDisc structure explicit.

The second line introduces the RDD components by including the running variable $Center_{i,2010}$ and its interactions with group and time. It permits different slopes on either side of the cutoff and allows those slopes to change over time. The final line integrates both the RDD and DiD components by including the triple interaction ($Above10_{i,2010} \times Post_t \times Center_{i,2010}$) with coefficient δ_1. This captures post-policy changes in the slope of the running variable at the cutoff. Though informative, δ_1 is secondary to δ_0 for interpreting the causal effect of the 2012 employee threshold policy.

5.2 Outcome Variables

The outcomes analyzed are grouped into three domains: firm size and labor composition, firm performance, and firm formality status. The first outcome domain captures whether firms adjust their labor force in response to the higher financial penalty for non-compliance associated with the 2012 employee threshold policy. While the policy threshold is defined in terms of FP workers, data limitations restrict measurement to FP *full-time* workers (see Section 3 for details). Because FP full-time workers is the running variable, it is not used as an outcome. Instead, Section 6.1 formally tests whether firms endogenously adjusted their FP full-time employment around the cutoff post-policy. For outcomes, I measure firm size using total workers, defined as the sum of full-time, part-time, and casual workers. I also examine firm labor composition to determine whether firms strategically restructured their labor force post-policy. Specifically, I evaluate four labor composition outcomes: part-time workers, casual workers, informal paid full-time workers, and unpaid full-time workers. The latter two are especially relevant because they fall outside the 2012 employee threshold policy's definition of "employee" and therefore do not count toward the compliance threshold. These labor composition outcomes directly relate to Theorem 2, which predicts substitution toward unregulated labor arrangements when threshold-induced costs increase.

The second outcome domain captures potential spillover effects on firm performance, including sales revenue per full-time worker, gross profit per full-time worker, labor costs per full-time worker, labor productivity, and capital productivity.[23] These outcomes test whether labor and compliance adjustments are accompanied by broader changes in firm profitability, productivity, or costs.

The third outcome domain captures firm formality status. I estimate Equation 6 for whether a firm is fully formal, formal but operating informally, informal but operating formally, or fully informal (see Table 2). These outcomes evaluate whether the 2012 employee threshold policy induces full formalization or instead encourages partial compliance along one dimension of informality. They relate to Theorem 3 which suggests that firms just above the threshold pre-policy may be registered with the government while relying on informal labor arrangements post-policy.

[23] All size-normalized firm economic account metrics represent year-end values. Gross profit is calculated as sales revenue minus total costs, which include raw materials, indirect expenses, and labor costs. Labor costs include wages, allowances, recruitment and training expenses, as well as health, social, and unemployment insurance contributions. Labor productivity is measured as value added per full-time worker, where value added equals sales revenue minus production costs (including raw materials and indirect expenses). Capital productivity is measured as value added per asset, including both physical and financial assets.

5.3 Optimal Bandwidth Selection

To reiterate, the running variable is a firm's number of FP full-time workers. Appendix Figure A.2 presents its distribution for the analytical sample pre- and post-policy.[24] Three features violate the continuity assumptions that underpin conventional RDD and, by extension, DiDisc estimation: (i) the distribution is highly left-skewed, (ii) the running variable is discrete and takes on only a few values, and (iii) there is a pronounced mass point at zero where approximately 78 percent of firms cluster in both periods. This mass at zero suggests that many firms strategically avoid regulatory exposure *altogether* by not employing any workers who satisfy the government's legal definition of "employee." This persistent pattern indicates widespread evasion of compliance requirements and underscores the extent to which firms operate beyond the reach of size-dependent formalization policies in settings characterized by high levels of informality. These features violate the smoothness and density continuity assumptions required for local polynomial estimation, the most commonly used implementation of RDD. They therefore motivate the use of the local randomization approach, which remains valid when the running variable is discrete and contains pronounced mass points.

The local randomization approach (Cattaneo et al., 2024) treats units within a selected window around the cutoff (i.e., 10 FP full-time workers) as if they were randomly assigned to treatment or control, circumventing the need for continuity or smoothness assumptions that are violated in this context.[25] Implementation involves conducting a series of covariate balance tests to determine the largest symmetric and asymmetric windows within which the "as-if" random assignment assumption is empirically plausible. This procedure tests for differences in the distribution of selected pre-policy covariates between treatment and control groups across possible candidate windows. The covariates selected capture distinct dimensions of firm heterogeneity expected to be associated with firms' pre-policy position around the 10-FP employee threshold and their potential outcomes. Table 3 defines these covariates and provides the rationale for their inclusion.[26] The procedure applies a modified greedy expansion rule adapted from Cattaneo et al. (2024) to identify the widest admissible window. Candidate windows are evaluated from narrowest to widest and are admissible when the minimum covariate-balance test p-value is at least 0.15. Because the running variable is discrete and some candidate windows have sparse support, initial windows are allowed to fail. However, once the first admissible window is found, the procedure expands only through consecutive admissible windows. The optimal window is thus the widest admissible window before the next candidate window fails.

Figure 4 shows the running variable's distribution for the pre-policy period within the optimal symmetric (h, h) and asymmetric (h_-, h_+) windows selected employing the aforementioned modified greedy expansion procedure. Within each selected optimal window, firms below the cutoff constitute the control group while firms at or above the cutoff constitute the treatment group. Applying the covariate balance tests to the pre-policy covariates listed in Table 3 supports a symmetric window of

[24] Appendix C describes the construction of the analytical sample from the Vietnam SME dataset.

[25] A "window" refers to the full range of values around the cutoff spanning both sides while a "bandwidth" refers to the distance from the cutoff on either side.

[26] As a secondary redundancy check, I examined pairwise Pearson and Spearman correlations among candidate pre-policy covariates in the analytical sample and avoided those with substantial conceptual and empirical overlap.

$[1, 18]$, corresponding to a bandwidth (h) of 9. The right-side bandwidth always includes the cutoff value because the 2012 employee threshold policy applies to firms with *at least* 10 FP workers. This optimal symmetric window yields a binomial test p-value of 0.687, which indicates no significant difference in treatment assignment, and a minimum covariate balance test p-value of 0.236.[27] Figure 5 plots the minimum covariate balance test p-values for all symmetric candidate windows around the cutoff, with the dotted horizontal line denoting the 0.15 threshold. The plotted values indicate that the widest symmetric window satisfying covariate balance is $[1, 18]$ while the next widest $[0, 19]$ yields a minimum p-value below 0.15 and thus fails the covariate balance test. Building on this, I apply the same modified greedy expansion logic to select the optimal *asymmetric* window. Starting from the optimal symmetric window of $[1, 18]$, I expand the right bandwidth sequentially, testing covariate balance after each candidate expansion. Because support is sparse at some values of the discrete running variable, the procedure allows initial expansions to fail while searching outward for the first admissible window; thereafter, expansion proceeds through consecutive admissible windows and stops at the first subsequent failure. These steps are then applied to the *left* bandwidth using the updated admissible window. The procedure repeats the expansion process for each bandwidth using the resulting admissible window and stops when neither can be increased without violating covariate balance. Table 3 reports the corresponding results, which support an optimal asymmetric window of $[1, 25]$ with a left-side bandwidth (h_-) of 9 and a right-side bandwidth (h_+) of 15.

Using these optimal windows, the analytical sample narrows to firms whose pre-policy FP full-time employment falls within those bandwidths. The optimal symmetric window of $[1, 18]$ defines a subsample of 154 firms—74 in the treatment group and 80 in the control—whereas the optimal asymmetric window of $[1, 25]$ results in 187 firms—107 treatment and 80 control. Both bandwidth-defined subsamples constitute the updated estimation samples used in the subsequent analyses.

[27] Results were obtained using the `rdwinselect` Stata command as recommended by Cattaneo et al. (2024).

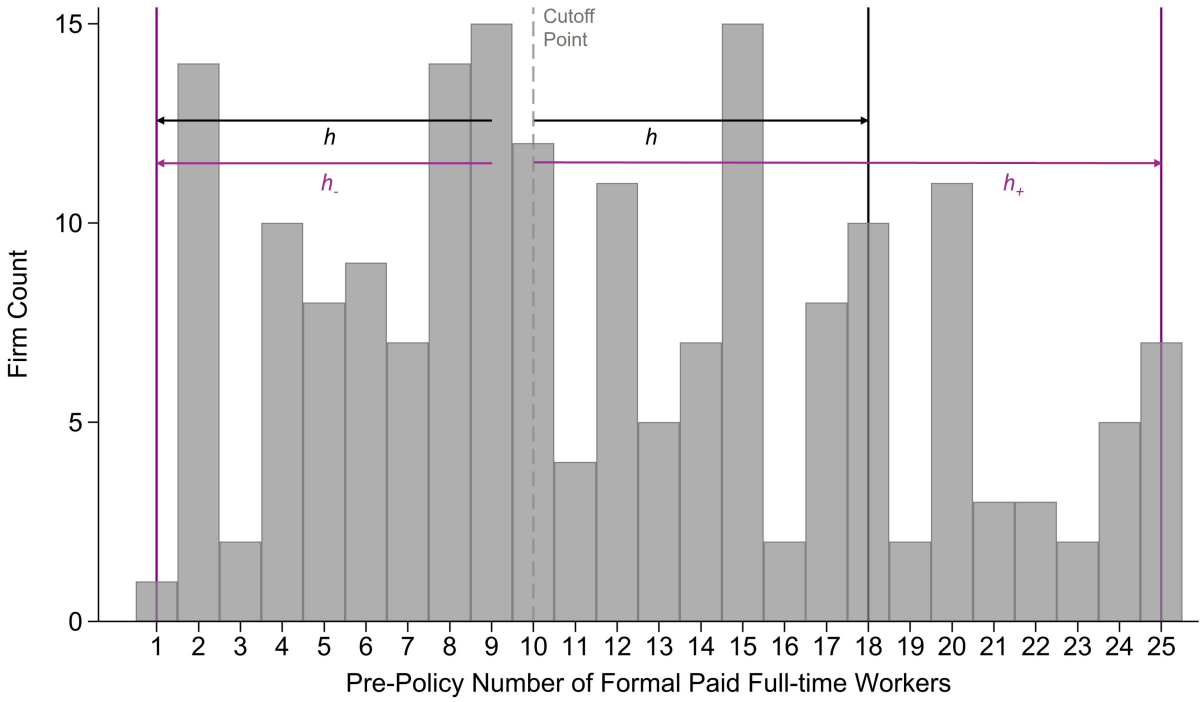


Figure 4: Distribution of Running Variable Pre-Policy within Optimal Windows

Notes: This figure shows the distribution of the running variable for the pre-policy period within the optimal windows. The optimal symmetric window is $[1, 18]$ and has a symmetric bandwidth (h) of 9. The optimal asymmetric window is $[1, 25]$ and has an asymmetric bandwidth to the left of the cutoff (h_-) of 9 and to the right of the cutoff (h_+) of 15. These optimal windows were identified using the local randomization approach for a discrete running variable and the modified greedy expansion rule described in Section 5.3.

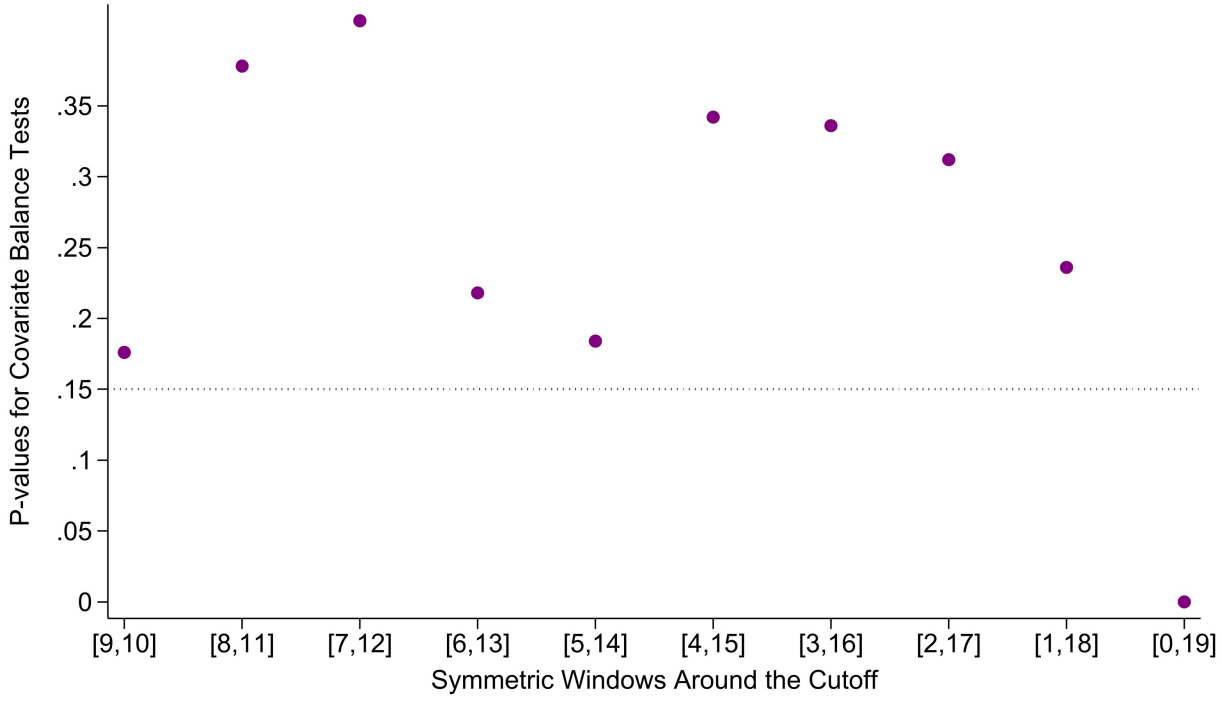


Figure 5: Minimum P-values for Covariate Balance Tests Across Symmetric Candidate Windows

Notes: This figure shows the minimum p-value for covariate balance tests for each candidate symmetric window around the cutoff, expressed in terms of the number of pre-policy formal paid full-time workers. Each dot corresponds to the lowest p-value across the covariates listed in Table 3. The dotted horizontal line marks the 0.15 threshold used to evaluate covariate balance. Windows whose dots are above this line pass the covariate balance test and are admissible; those with dots below the line fail. Using the modified greedy expansion rule described in Section 5.3, the optimal symmetric window is $[1, 18]$, corresponding to the widest admissible window before the next candidate window fails.

Table 3: Pre-Policy Covariate Balance Within Optimal Asymmetric Window of [1,25]

Covariate	Means		Fisherian P-value	Description	Rationale
	Control	Treatment			
Owner/Manager Characteristics					
Male (1 = yes, 0 = no)	0.58	0.51	0.546	Gender	Gender-linked firm differences
Upper secondary school completion (1 = yes, 0 = no)	0.80	0.89	0.160	Highest level of education completed is upper secondary school	Managerial human capital
Firm Characteristics					
Road access (1 = yes, 0 = no)	0.85	0.85	1.000	A main, paved road leads to the firm	Infrastructure access
Age	10.23	9.57	0.690	Age of firm in years	Firm maturity
Value added per full-time worker	86,681	104,830	0.360	Year-end 2010 value added divided by the number of full-time workers	Firm productivity
Assets per full-time worker	732,109	571,051	0.316	Year-end 2010 physical and financial assets divided by the number of full-time workers	Firm resource endowment
Firms	80	107			

Notes: Firms with less than 10 formal paid (FP) full-time workers pre-policy form the control group; those with 10 or more constitute the treatment group. In Vietnam, completion of upper secondary school is equivalent to a U.S. high school diploma and is *optional* after lower secondary school. Value added is calculated as sales revenue minus production costs (including raw materials and indirect expenses) while gross profit equals value added minus labor costs (e.g., wages, allowances, recruitment, training, as well as health, social, and unemployment insurance). All size-normalized firm economic account metrics are reported in thousands of Vietnamese Dong. Covariate balance tests were conducted using the `rdrandinf` Stata command as recommended by Cattaneo et al. (2024). The finite sample two-sided Fisherian p-values test if the distribution of each covariate differs statistically across the cutoff within the optimal asymmetric window of 1 to 25 FP full-time workers, based on bandwidths of 9 units to the left and 15 units to the right of the cutoff. The right-side bandwidth includes the cutoff value because the 2012 employee threshold policy applies to firms with **10 or more** FP workers.

6 Results

The model presented in Section 4 predicts how threshold-based regulations distort firms' labor and formality decisions. Building on this framework, I empirically evaluate how the 2012 employee threshold policy shapes firm behavior around the 10-formal paid (FP) employee threshold. Section 6.1 tests for endogenous adjustments in firms' employment of FP full-time workers at this threshold while Section 6.2 estimates average effects on labor composition, performance, and formality status.

6.1 McCrary Density Tests

To evaluate if firms strategically manipulated their size distribution around the 10-FP employee threshold, I conduct the McCrary (2008) density test separately for the pre- and post-policy periods using the optimal symmetric and asymmetric windows from Section 5.3. This approach tests for discontinuities in the distribution of the running variable (i.e., the number of FP full-time workers) at the cutoff in each period to determine whether manipulation emerged after the 2012 employee threshold policy's implementation. A statistically significant jump in density just below the 10-FP employee threshold post-policy would suggest that firms adjusted their FP employment downward to avoid crossing the threshold, consistent with the bunching behavior described in Theorem 1.

Figure 6 shows the distribution of firms by their number of FP full-time workers in the pre- and post-policy periods within each optimal window. Panel A uses the optimal symmetric window of $[1, 18]$ while Panel B uses the optimal asymmetric window of $[1, 25]$. Under the optimal symmetric window, the McCrary (2008) density test finds evidence of a discontinuity at the cutoff in the pre-policy period: the estimated log-difference in density across the threshold is negative ($T = -2.01$) and statistically significant ($p = 0.045$). The negative estimate indicates fewer firms were at or just above the 10-FP employee threshold than just below it. This discontinuity is consistent with firms sorting below the threshold to avoid the compliance requirement and the associated financial penalty under the policy's previous version, when the fine for non-compliance was only 1 million VND. Yet, the estimate becomes insignificant in the post-policy period ($T = -1.15$, $p = 0.251$). The wider asymmetric window yields the same findings: a negative and statistically significant discontinuity in the pre-policy period ($T = -2.05$, $p = 0.041$) followed by an insignificant estimate in the post-policy period ($T = -1.04$, $p = 0.298$). Visual inspection of Panel A and B reinforces this interpretation: the pre-policy periods exhibit a pronounced "hollowing" at 10 FP full-time workers with excess mass just below, whereas the post-policy period distributions appear smoother across the cutoff. This suggests that the sharper discontinuity evident pre-policy does not persist after the introduction of the 2012 employee threshold policy. These patterns—a significant negative density discontinuity pre-policy and no discontinuity post-policy—also hold when using the analytical sample[28].

The absence of a post-policy discontinuity suggests that the conditions for bunching in Theorem 1 were not met for a sufficient mass of firms near the cutoff. Section 4 shows that bunching requires

[28]The analytical sample (see Appendix C for details) includes 228 treatment and 1,211 control firms. Applying the McCrary (2008) density test yields $T = -2.00$ ($p = 0.045$) pre-policy and $T = -0.585$ ($p = 0.559$) post-policy.

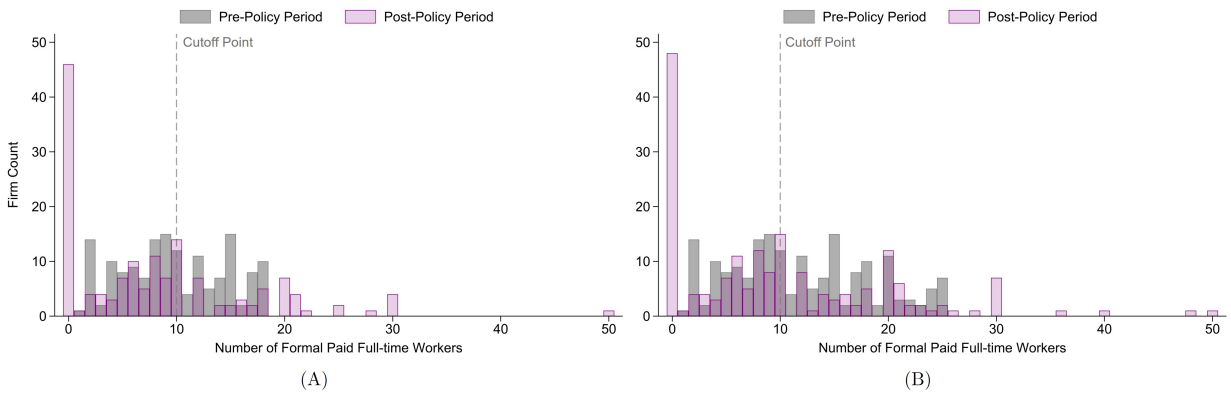


Figure 6: Distribution of Running Variable Pre- and Post-Policy within Optimal Windows

Notes: This figure shows the distribution of firms around the 10-formal paid (FP) employee threshold, with Panels A and B using the optimal symmetric and asymmetric windows, respectively. The gray bars depict the distribution of firms within these optimal windows in the pre-policy period while the purple bars depict the distribution of *the same firms* in the post-policy period. Panels A and B show a statistically significant pre-policy drop in firm density just above the threshold relative to just below—consistent with bunching—but the post-policy estimates are insignificant.

remaining below the threshold to yield a higher payoff than every feasible crossing strategy and that firms can decrease their FP employment at sufficiently low cost. Therefore, the null result does not imply that the policy was irrelevant; rather, it points to possible adjustments on margins other than manipulating FP employment to locate just below the cutoff. Lemma 1 clarifies one reason this may occur: the increased financial penalty for IWR non-submission makes crossing the threshold without submitting IWRs less attractive, especially in an environment where enforcement was credible (see Figure 3).[29] This logic motivates analyzing whether firms adjusted along observable margins other than FP employment manipulation. Section 6.2 examines two model-implied responses: changes in labor composition among firms just below the threshold pre-policy and changes in formality status among firms just above it pre-policy, corresponding to Theorems 2 and 3, respectively.[30]

6.2 Difference-in-Discontinuities

Table 4 presents estimates from Equation 6 using the symmetric and asymmetric optimal windows, reporting only outcomes with statistically significant effects.[31] The findings show that firms just *above* the 10-FP employee threshold responded to the higher costs of informality introduced by the 2012 employee threshold policy compared to firms just *below* the threshold. They are consistent

[29]The financial penalty for non-compliance under the 2012 employee threshold policy is 5 to 10 million VND—five to 10 times higher than under the policy's previous version.

[30]The absence of post-policy manipulation also supports the data reliability of the Vietnam SME database. If firms had strategically underreported FP full-time workers during survey administration, we would expect excess mass just below the 10-FP employee threshold. Instead, the smoother distribution suggests accurate reporting (see Section 3).

[31]Results for all outcomes described in Section 5.2 are reported in Appendix Tables A.3, A.4, and A.5. Appendix Table A.6 reports wild cluster bootstrap p-values for the estimates in Table 4. The similarity between clustered and bootstrap p-values suggests that the qualitative inference is not sensitive to this alternative inference procedure.

with the behavioral margins predicted by the firm profit maximization model in Section 4. Specifically, the results support Theorems 2 and 3 which predict substitution toward threshold-excluded labor and selective compliance with multi-dimensional firm formality requirements, respectively. Following multiple-hypothesis adjustments that control the false discovery rate within pre-specified testing families using the Benjamini-Hochberg and two-stage Benjamini-Krieger-Yekutieli procedures, statistical support remains for selected estimates (see Appendix Table A.7).[32] Additionally, placebo threshold tests in Appendix F show that comparable effects do not systematically emerge *away* from the 10-FP employee threshold.

Table 4: Impact of Threshold-Based Labor Policy on Selected Firm Outcomes

	Unpaid Full-time Workers		Formal Firm But Operates Informally		Gross Profit per Full-time Worker		Labor Productivity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Post_t \times$	-0.376	-0.352*	0.150*	0.182**	100,207*	139,326**	88,750	123,164*
$Above10_{i,2010}$	(0.251)	(0.209)	(0.089)	(0.088)	(56,791)	(67,209)	(59,167)	(69,049)
Control mean	0.762	0.762	0.038	0.038	55,148	55,148	86,681	86,681
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.755	0.767	0.593	0.570	0.509	0.508	0.514	0.514
Optimal window	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]
Firm-years	308	374	308	374	308	374	308	374

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. Reported coefficients correspond to $\hat{\delta}_0$ from Equation 6. "Control mean" reports the pre-policy mean of the outcome variable for the control group (i.e., firms just below the threshold) within the corresponding optimal window. "Firm-years" reports firm-period observations, with each firm contributing one year-end 2010 and one year-end 2014 observation. Estimates in Columns 5-8 are reported in thousands of Vietnamese Dong and reflect year-end values. Only firm outcomes with statistically significant effects are presented. Appendix Figure A.3 presents binned scatter plots with fitted linear regressions for each selected firm outcome, illustrating the change in discontinuity at the 10-FP employee threshold between the pre- and post-policy periods. Wild cluster bootstrap p-values and FDR-adjusted q-values for these estimates are reported in Appendix Tables A.6 and A.7, respectively.

For labor composition outcomes, Column 2 shows that the number of unpaid full-time workers decreased by approximately 0.35 workers for firms just above the threshold compared to those just below it post-policy. The effect is only significant for firms within the optimal asymmetric window. This change in the discontinuity suggests that firms just *below* the threshold relied relatively more on unpaid labor arrangements, potentially involving family members, to limit their FP employment while maintaining operational capacity. The finding is consistent with Theorem 2, which predicts that a higher cost of non-compliance increases $\tau_i(P_r)$—the regulatory wedge associated with adding FP employment at the cutoff—and thereby makes threshold-excluded labor more attractive. The

[32] Appendix Table A.7 reports false discovery rate (FDR)-adjusted q-values computed within pre-specified testing families using the Benjamini-Hochberg and the two-stage Benjamini-Krieger-Yekutieli procedures for the outcomes reported in Appendix Tables A.3, A.4, and A.5. Regarding Table 4, the corresponding (p, q) pair under the optimal symmetric window is $(0.0915, 0.2746)$ for "formal firm but operates informally." Under the optimal asymmetric window, the corresponding pairs are $(0.0939, 0.1878)$ for "unpaid full-time workers," $(0.0397, 0.1192)$ for "formal firm but operates informally," $(0.0395, 0.0791)$ for "gross profit per full-time worker," and $(0.0761, 0.1522)$ for "labor productivity." Thus, only the gross profit result remains significant after adjustment for multiple hypothesis testing.

estimate suggests the policy change distorted firms' labor composition. Descriptive evidence further reinforces this behavioral pattern: Appendix Figure A.2 shows that about 78 percent of firms in the analytical sample—which includes firms within and outside of the optimal windows—employ zero FP full-time workers across the pre- and post-policy periods. This mass point conveys widespread avoidance of the regulatory threshold, indicating that rather than risk triggering compliance obligations most firms structure their labor force to exclude *any* workers that satisfy the legal definition of "employee." These results suggest that firms respond to the policy's higher financial penalty and stricter enforcement (see Figure 3) by adjusting their labor margins to reduce regulatory exposure.

For firm performance outcomes, Columns 5-6 show that firms just above the threshold realized sizable gains in profitability compared to firms just below it in the post-policy period. Specifically, gross profit per full-time worker increased by 100 million VND (about 3,800 USD) under the optimal symmetric window and 139 million VND (about 5,300 USD) under the optimal asymmetric window. These improvements suggest that despite the higher compliance costs, firms just above the threshold were able to maintain or even enhance their profit levels. This could reflect scale effects, improved input allocation, or cost containment through partial labor contract compliance. Moreover, Column 8 shows an increase in labor productivity[33] of 123 million VND (about 4,700 USD) for firms just above the threshold relative to those just below it post-policy. This effect, significant only under the optimal asymmetric window, is consistent with how shifting from unpaid to paid labor can improve operational efficiency for these firms.[34] While not directly targeted by the 2012 employee threshold policy, these outcomes suggest positive spillovers beyond its intended compliance margins.

Finally, estimates in Columns 3-4 are consistent with the strategic non-compliance mechanism. The probability that a firm is "formal but operates informally" (i.e., registered with the government but does not provide all of its full-time workers with formal, written labor contracts) increased by 15 to 18 percentage points among firms just above the threshold relative to those just below it post-policy. Compared to a pre-policy mean of only 3.8 percent among firms just *below* the threshold, this represents a substantial behavioral shift. This finding provides empirical support for Theorem 3, which predicts that once firms cross the threshold they may adopt a partial formality status—satisfying visible requirements such as government registration while avoiding costlier or harder-to-monitor ones like universal labor contracts. The observed increase in partial firm formality reflects the strategic adjustment predicted by the model: firms balance legitimacy and cost minimization by selectively complying along the cheapest margin (Fietz et al., 2026). In practice, this means relying on informal labor even after registering. Such behavior undermines the intent of the 2012 employee threshold policy and underscores the challenges of enforcing multi-dimensional formalization.

These findings suggest a nuanced but strategic firm response to this threshold-based labor policy: firms internalize the costs of formalization and adapt along multiple margins—labor composition, profitability, efficiency, and compliance—by shifting toward the least costly forms of adjustment.

[33]Labor productivity is measured as value added per full-time worker, where value added refers to the difference between sales revenue and production costs (including raw materials and indirect expenses).

[34]Appendix Table A.8 shows that the profitability result is robust to winsorization at the 1st and 99th percentiles while the labor productivity estimate remains positive but loses statistical significance.

Appendix E.3 strengthens this interpretation by showing that these effects emerge in year-end 2014 and are generally not preceded by comparable discontinuities in year-end 2008 or 2012. Altogether, the evidence indicates that the 2012 employee threshold policy encouraged movement along the registration margin while simultaneously enabling partial compliance. This pattern raises an important policy concern: without stronger enforcement across *all* dimensions of formality, threshold-based regulations may produce nominal compliance rather than substantive formalization.

Because Table 4 averages across firms with different pre-policy formality statuses, it may conceal heterogeneous responses. Appendix G re-estimates the effects for three subgroups with support on both sides of the cutoff: fully formal firms, informal firms that operate formally, and fully informal firms. Results indicate that fully formal firms just above the threshold pre-policy adjusted through lower unpaid full-time labor and increased gross profit and labor productivity, with some evidence of a shift from full to partial formality through reduced labor contract coverage for full-time workers. Informal firms operating formally exhibited no robust response, whereas fully informal firms showed suggestive performance gains. Though some estimates are window-specific, these patterns suggest that firms' adjustment margins varied with their initial position along the formality spectrum.

7 Conclusion

This research presents causal evidence that size-dependent formalization policies can engender multi-dimensional behavioral responses among firms in developing economies. I exploit Article 119 of Vietnam's Labor Code 2012 which builds upon a pre-existing firm-size threshold by imposing a substantially higher financial penalty on firms with at least 10 formal paid (FP) employees that fail to register internal work regulations with the government. The empirical results suggest that firms respond along multiple margins instead of exhibiting canonical "bunching" behavior. To interpret these findings, I derive three predictions from a firm profit maximization model.

The predictions structure the results. McCrary (2008) density tests provide no evidence that the threshold-avoidance conditions delineated in Theorem 1 generate observable post-policy bunching just below the 10-FP employee threshold. Rather than manipulating FP employment just below the threshold, firms responded along margins consistent with Theorems 2 and 3. Firms just below the threshold increased their use of unpaid full-time labor, consistent with substitution away from regulated FP employment. In contrast, firms just above the threshold became more likely to register with the government while depending on informal labor arrangements. These "formal but operating informally" firms adopted hybrid formality strategies consistent with decreasing regulatory exposure while preserving operational flexibility. Also, firms just above the threshold realized profit and labor productivity gains, suggesting that partial compliance can coexist with efficiency improvements.

Together, these findings suggest that threshold-based labor policies can shape *how* firms comply instead of just *whether* they comply. Increased government registration among firms just above the threshold reflects adjustment along a highly observable margin of formality. Yet, limited improvements in labor contract compliance and reliance on unregulated labor show that firms selectively

satisfy the most visible regulations while circumventing costlier or less-easily-monitored regulations. Such patterns reflect strategic non-compliance: deterrence through increased financial penalties can encourage formalization along observable margins but does not ensure comprehensive compliance.

Methodologically, this paper advances the application of the difference-in-discontinuities design with local randomization to a context with a discrete and highly skewed running variable. Moreover, this research operationalizes a nuanced definition of firm formality by distinguishing between a firm's government registration status and compliance with formal labor practices. This distinction reveals gradations of firm informality that binary measures mask—showing how firms can restructure their labor force or employ hybrid formality strategies to minimize regulatory exposure.

From a policy perspective, these results highlight that effective formalization strategies hinge on *how* firms perceive and navigate regulatory obligations. When compliance thresholds trigger sharp cost jumps, firms may neither fully comply with the policy nor evade it outright, but instead reconfigure their operations to meet visible enforcement requirements while employing informal practices elsewhere. This behavior reflects a broader challenge for developing economies: when firms comply for the sake of visibility rather than substance, formalization policies risk undermining institutional credibility and perpetuating the very informality they seek to eliminate. More broadly, these findings suggest that formalization policies should be viewed not only as enforcement instruments but as development tools that influence employment quality, productivity, and long-term upgrading.

While this study offers robust evidence for Vietnamese manufacturing firms, its external validity is bounded by data and contextual limitations. First, restricting the measurement of employment to FP full-time workers likely underestimates firms' *true* formal labor force, as part-time and casual workers with formal, written labor contracts are excluded. Since this definition differs from the employee definition used in the policy—which counts all formally contracted, paid workers—measured firm size may be biased downward compared to the 10-FP employee threshold. Consequently, the results may understate the extent to which such firms actually face or respond to the policy. Second, Vietnam's enforcement environment has recently been characterized by higher financial penalties and stronger inspection capacity. Hence it may differ from settings with very weak or very strong institutions. Future research should assess how variation in enforcement intensity, monitoring credibility, and policy thresholds influences firms' behavior across sectors and countries, and whether similar hybrid formality equilibria emerge. Ultimately, understanding these mechanisms is crucial for designing policies that expand the benefits of formality without prompting firm-level distortions.

References

Aghion, P., Bergeaud, A., & Van Reenen, J. (2023). The impact of regulation on innovation. *American Economic Review*, *113*(11), 2894–2936.

Amaral, P. S., & Quintin, E. (2006). A competitive model of the informal sector. *Journal of monetary Economics*, *53*(7), 1541–1553.

Amin, M., & Okou, C. (2020). Casting a shadow: Productivity of formal firms and informality. *Review of Development Economics*, *24*(4), 1610–1630.

Amirapu, A., & Gechter, M. (2020). Labor regulations and the cost of corruption: Evidence from the indian firm size distribution. *Review of Economics and Statistics*, *102*(1), 34–48.

Angelino, A., Tassinari, M., Barbieri, E., & Di Tommaso, M. R. (2021). Institutional and economic transition in vietnam: Analysing the heterogeneity in firms' perceptions of business environment constraints. *Competition & Change*, *25*(1), 52–72.

Benedek, M. D., Deb, M. P., Gracia, M. B., Saksonovs, M. S., Shabunina, M. A., & Budina, M. N. T. (2017). *The right kind of help? tax incentives for staying small*. International Monetary Fund.

Benhassine, N., McKenzie, D., Pouliquen, V., & Santini, M. (2018). Does inducing informal firms to formalize make sense? experimental evidence from benin. *Journal of Public Economics*, *157*, 1–14.

Benjamin, N. C., & Mbaye, A. A. (2012). The informal sector, productivity, and enforcement in w est a frica: A firm-level analysis. *Review of Development Economics*, *16*(4), 664–680.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289–300.

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491–507.

Bennedsen, M., Simintzi, E., Tsoutsoura, M., & Wolfenzon, D. (2022). Do firms respond to gender pay gap transparency? *The Journal of Finance*, *77*(4), 2051–2091.

Bertrand, M., Hsieh, C.-T., & Tsivanidis, N. (2021). *Contract labor and firm growth in india* (tech. rep.). National Bureau of Economic Research.

Bruhn, M., & McKenzie, D. (2014). Entry regulation and the formalization of microenterprises in developing countries. *The World Bank Research Observer*, *29*(2), 186–201.

Candela, M. A. (2013). *Size-dependent policies and firm behavior*. University of California, Berkeley.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2024). *A practical introduction to regression discontinuity designs: Extensions*. Cambridge University Press.

Dabla-Norris, M. E., Jaramillo, L., Lima, F., & Sollaci, A. (2018). *Size dependent policies, informality and misallocation*. International Monetary Fund.

De Soto, H. (1989). *The other path: The invisible revolution in the third world*. Harper & Row New York.

Fietz, K., Lakemann, T., Beber, B., Priebe, J., & Lay, J. (2026). Formalizing employment in africa's small firms: Experimental evidence from côte d'ivoire. *The World Bank Economic Review*, lhag003.

Garibaldi, P., Pacelli, L., & Borgarello, A. (2004). Employment protection legislation and the size of firms. *Giornale degli economisti e annali di economia*, 33–68.

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

Gourio, F., & Roys, N. (2014). Size-dependent regulations, firm size distribution, and reallocation. *Quantitative Economics*, *5*(2), 377–416.

Grembi, V., Nannicini, T., & Troiano, U. (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics*, 1–30.

Guner, N., Ventura, G., & Yi, X. (2006). How costly are restrictions on size? *Japan and the World Economy*, *18*(3), 302–320.

Institute of Labour Science and Social Affairs & the University of Copenhagen. (2009, July). *Manual - guidelines: Survey of small and medium scale manufacturing enterprises (smes) in vietnam* (tech. rep.).

La Porta, R., & Shleifer, A. (2008). *The unofficial economy and economic development* (tech. rep.). National Bureau of Economic Research.

La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of economic perspectives*, *28*(3), 109–126.

Lewis, W. A., et al. (1954). Economic development with unlimited supplies of labour.

McCaig, B., & Pavcnik, N. (2013). *Moving out of agriculture: Structural change in vietnam* (tech. rep.). National Bureau of Economic Research.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, *142*(2), 698–714.

McKenzie, D. (2017). Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition. *American Economic Review*, *107*(8), 2278–2307.

Morikawa, M. (2023). Compliance costs and productivity: An approach from working hours. *Journal of Regulatory Economics*, *63*(3), 117–137.

Mulligan, C. B. (2020). The employer penalty, voluntary compliance, and the size distribution of firms: Evidence from a survey of small businesses. *Tax Policy and the Economy*, *34*(1), 139–171.

Padmakumar, K. (2022). Small by choice? reassessing the aggregate implications of size-based regulations.

Perry, G. (2007). *Informality: Exit and exclusion*. World Bank Publications.

Prado, M. (2011). Government policy in the formal and informal sectors. *European Economic Review*, *55*(8), 1120–1136.

Qian, L., & Vereshchagina, G. (2022). *Capital financing constraints, size-dependent distortions, and aggregate productivity* (tech. rep.). mimeo.

Rand, J., & Tarp, F. (2020). *Micro, small, and medium enterprises in vietnam*. Oxford University Press.

Schivardi, F., & Torrini, R. (2008). Identifying the effects of firing restrictions through size-contingent differences in regulation. *Labour Economics*, *15*(3), 482–511.

Simes, R. J. (1986). An improved bonferroni procedure for multiple tests of significance. *Biometrika*, *73*(3), 751–754.

Takahashi, M. (2024). Difference-in-discontinuities design. *Available at SSRN 4686891*.

Tramontin Shinoki, S., Picchetti, P., & Campos de Xavier Pinto, C. (2024). Difference-in-discontinuities: Estimation, inference and validity tests. *Inference and Validity Tests (May 28, 2024)*.

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

Appendix

A Supplemental Tables & Figures

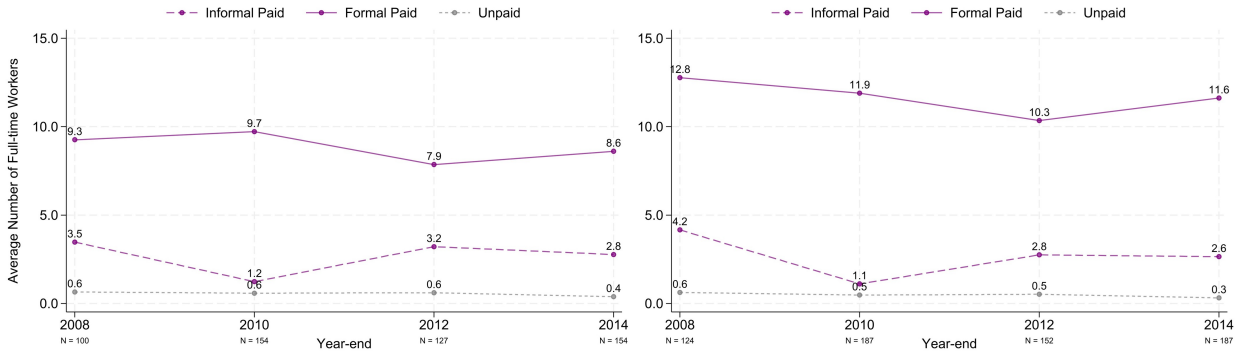


Figure A.1: Average Number of Full-time Workers by Labor Category and Year-End Period within Optimal Windows

Notes: This figure shows the average number of (i) informal paid, (ii) formal paid, and (iii) unpaid full-time workers across multiple year-end periods. Panels A and B use the optimal symmetric and asymmetric windows defined in Section 5.3, respectively. Each year-end period's sample size is indicated below the corresponding year. Since the panel is unbalanced in year-end 2008 and 2012, paired t-tests compare FP full-time employment only among firms observed in both years of each comparison. In Panel A, the p-values associated with these comparisons are $p = 0.9181$ for year-end 2008 to 2010, $p = 0.0028$ for year-end 2010 to 2012, and $p = 0.3477$ for year-end 2012 to 2014. In Panel B, the corresponding p-values are $p = 0.3072$, $p = 0.0386$, and $p = 0.1433$. These findings further support treating year-end 2012 as an interim period and using year-end 2010 as the preferred pre-policy period (see Section 3).

Table A.1: Firm Formality Status Distribution by Year-End Period

	2008	2010	2012	2014
Fully formal	15.2%	5.3%	8.5%	9.9%
Partially formal				
Formal but operates informally	49.4%	4.9%	12.3%	39.9%
Informal but operates formally	0.9%	13.0%	8.7%	5.5%
Fully informal	34.5%	76.8%	70.5%	44.7%
Firms	2,334	2,297	2,371	2,452

Notes: Data reflect year-end values (i.e., December 31st). Only firms from the raw survey samples with fewer than 50 total workers across years are included, which is consistent with the World Bank's definition of micro- and small-enterprises. Firm formality status definitions are in Table 2.

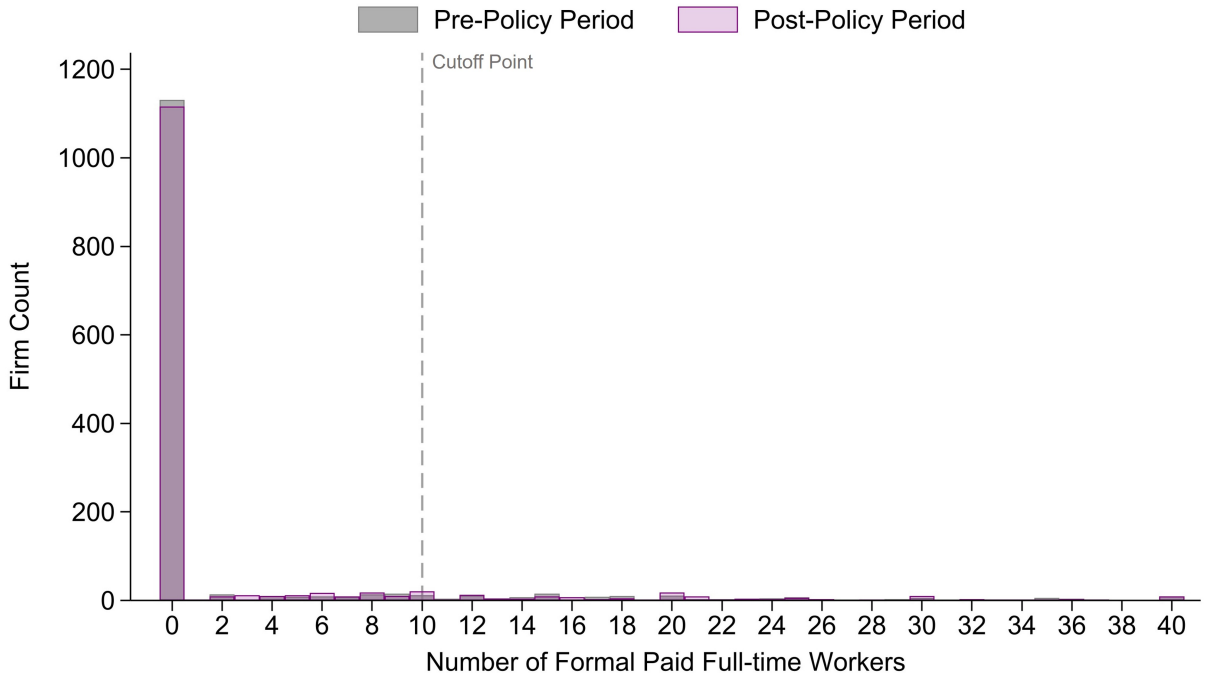


Figure A.2: Distribution of Running Variable Pre- and Post-Policy

Notes: This figure displays the distribution of the running variable (i.e., the number of formal paid full-time workers) for the pre- and post-policy periods for the analyitcal sample (see Appendix C). Three issues are apparent: (i) the distribution is highly left-skewed, (ii) the running variable is discrete and takes on few values, and (iii) there is a mass point at zero where about 78 percent of firms cluster in both periods. These violate continuity assumptions required for conventional RDD estimation and instead motivate the use of a local randomization approach (see Section 5.3).

Table A.2: Financial Penalty Relative to Firm Economic Account Metrics Pre- and Post-Policy within Optimal Windows

	Pre-Policy Period			Post-Policy Period		
	Control	Treatment		Control	Treatment	
Percent of Average Annual	[1,18] & [1,25]	[1,18]	[1,25]	[1,18] & [1,25]	[1,18]	[1,25]
Sales revenue	0.04%	0.02%	0.01%	0.12%-0.24%	0.02%-0.04%	0.02%-0.05%
Gross profit	0.25%	0.13%	0.09%	1.07%-2.15%	0.30%-0.60%	0.31%-0.61%
Labor costs	0.39%	0.15%	0.13%	0.76%-1.52%	0.62%-1.23%	0.49%-0.99%
Firms	80	74	107	80	74	107

Notes: Firms with less than 10 formal paid full-time workers pre-policy form the control group; those with 10 or more constitute the treatment group. The optimal symmetric and asymmetric windows are [1,18] and [1, 25], respectively. The financial penalty for non-compliance increased from 1 million VND pre-policy to between 5 and 10 million VND post-policy. Gross profit is the difference between sales revenue and total costs—including raw materials, indirect expenses, and labor costs (e.g., wages, allowances, recruitment, training, as well as health, social, and unemployment insurance). Overall, the revised financial penalty tied to the 2012 employee threshold policy accounts for a larger share of sales revenue, gross profit, and labor costs in the control group than in the treatment group across both optimal windows. This suggests that firms just *below* the threshold pre-policy generally faced a substantially higher financial burden for non-compliance under the 2012 employee threshold policy.

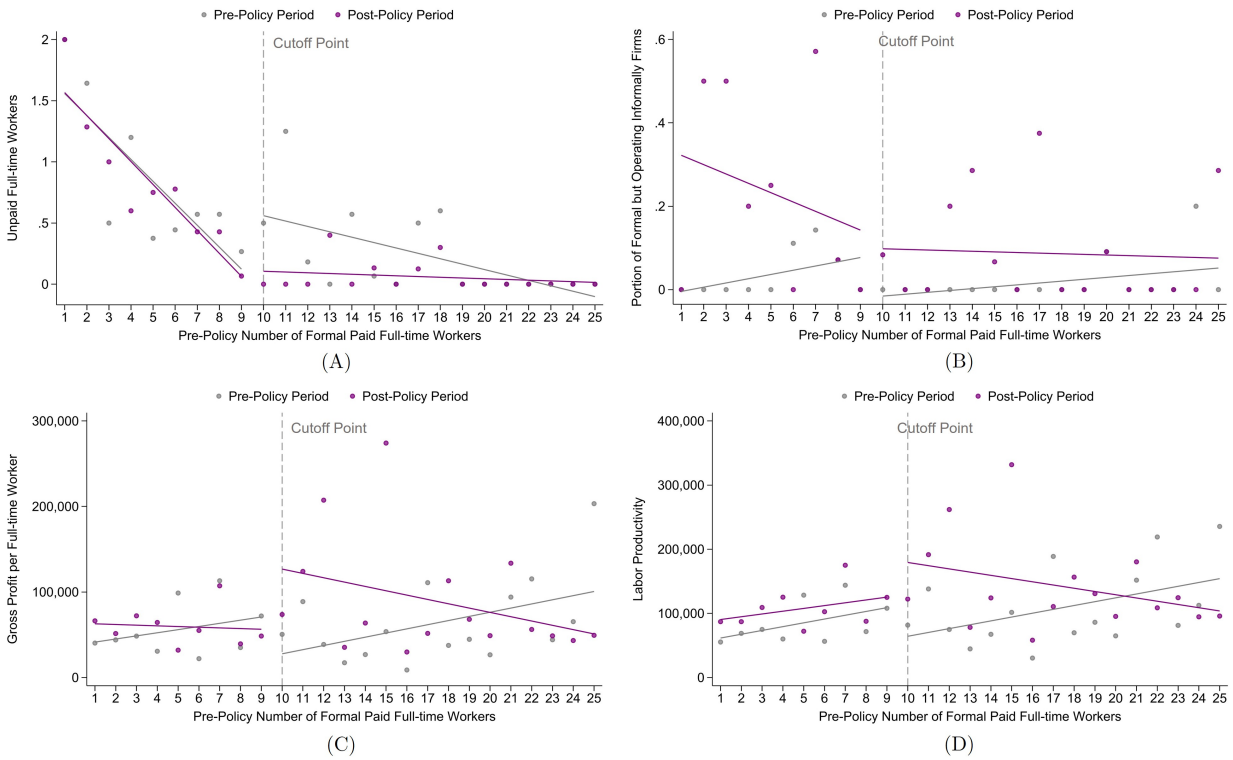


Figure A.3: Visualizing Changes in Discontinuity at the 10-Formal Paid Employee Threshold: Pre- and Post-Policy Period Comparisons for Selected Firm Outcomes

Notes: Each panel shows a bin scatter plot overlaid with fitted linear regressions, comparing the association between the selected firm outcomes listed in Table 4 and the pre-policy number of formal paid full-time workers across the pre- and post-policy periods. The vertical line denotes the cutoff introduced by the 2012 employee threshold policy. All firms above this threshold are subject to a higher financial penalty for non-compliance in the post-policy period. These plots illustrate changes in the discontinuity at the threshold over time within the optimal windows, providing visual evidence of the policy's localized temporal impact that complements the causal estimates reported in Table 4. Results in Panels C and D are reported in thousands of Vietnamese Dong and reflect year-end values.

Table A.3: Impact of Threshold-Based Labor Policy on Firm Size and Labor Composition

	Total Workers		Part-time Workers		Casual Workers		Informal Paid Full-time Workers		Unpaid Full-time Workers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$Post_t \times$	-10.581	-15.845	-0.785	1.179	-6.153	-12.276	-0.283	0.571	-0.376	-0.352*
$Above10_{i,2010}$	(11.498)	(11.564)	(0.984)	(1.098)	(10.121)	(10.002)	(1.275)	(1.325)	(0.251)	(0.209)
Control mean	9.125	9.125	0.412	0.412	0.875	0.875	1.263	1.263	0.762	0.762
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.714	0.724	0.515	0.500	0.564	0.556	0.518	0.518	0.755	0.767
Optimal window	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]
Firm-years	308	374	308	374	308	374	308	374	308	374

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. Reported coefficients correspond to $\hat{\delta}_0$ from Equation 6. "Control mean" reports the pre-policy mean of the outcome variable for the control group (i.e., firms just below the threshold) within the corresponding optimal window. "Firm-years" reports firm-period observations, with each firm contributing one year-end 2010 and one year-end 2014 observation.

Table A.4: Impact of Threshold-Based Labor Policy on Firm Performance

	Sales Revenue per Full-time Worker		Gross Profit per Full-time Worker		Labor Costs per Full-time Worker		Labor Productivity		Capital Productivity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$Post_t \times$	992,352	2,047,275	100,207*	139,326**	-11,457	-16,162	88,750	123,164*	0.147	0.451
$Above10_{i,2010}$	(789,137)	(1,770,930)	(56,791)	(67,209)	(14,648)	(14,618)	(59,167)	(69,049)	(0.464)	(0.516)
Control mean	298,647	298,647	55,148	55,148	31,533	31,533	86,681	86,681	0.560	0.560
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.506	0.504	0.509	0.508	0.582	0.585	0.514	0.514	0.527	0.534
Optimal window	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]
Firm-years	308	374	308	374	308	374	308	374	308	374

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. Reported coefficients correspond to $\hat{\delta}_0$ from Equation 6. "Control mean" reports the pre-policy mean of the outcome variable for the control group (i.e., firms just below the threshold) within the corresponding optimal window. "Firm-years" reports firm-period observations, with each firm contributing one year-end 2010 and one year-end 2014 observation. Estimates are reported in thousands of Vietnamese Dong and reflect year-end values. Gross profit equals sales revenue minus total costs—including raw materials, indirect expenses, and labor costs (e.g., wages, allowances, recruitment, training, as well as health, social, and unemployment insurance). Labor productivity is defined as value added per full-time worker while capital productivity is defined as value added per asset. Value added is calculated as sales revenue minus production costs (including raw materials and indirect expenses).

Table A.5: Impact of Threshold-Based Labor Policy on Firm Formality Status

	Fully Formal Firm		Formal Firm But Operates Informally		Informal Firm But Operates Formally		Fully Informal Firm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Post_t \times$	-0.001	0.016	0.150*	0.182**	-0.131	-0.227	-0.019	0.030
$Above10_{i,2010}$	(0.146)	(0.138)	(0.089)	(0.088)	(0.181)	(0.162)	(0.155)	(0.141)
Control mean	0.175	0.175	0.038	0.038	0.587	0.587	0.200	0.200
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.623	0.648	0.593	0.570	0.606	0.609	0.545	0.549
Optimal window	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]
Firm-years	308	374	308	374	308	374	308	374

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. Reported coefficients correspond to $\hat{\delta}_0$ from Equation 6. "Control mean" reports the pre-policy mean of the outcome variable for the control group (i.e., firms just below the threshold) within the corresponding optimal window. "Firm-years" reports firm-period observations, with each firm contributing one year-end 2010 and one year-end 2014 observation. Firm formality status definitions are in Table 2.

Table A.6: Bootstrap Inference for DiDisc Estimates of Selected Firm Outcomes

	Unpaid Full-time Workers		Formal Firm But Operates Informally		Gross Profit Per Full-time Worker		Labor Productivity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coefficient	-0.376	-0.352	0.150	0.182	100,207	139,326	88,750	123,164
Clustered p-value	0.135	0.094	0.092	0.040	0.080	0.040	0.136	0.076
Bootstrap p-value	0.141	0.094	0.095	0.043	0.049	0.022	0.128	0.070
Optimal window	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]
Firm-years	308	374	308	374	308	374	308	374

Notes: This table reports inference robustness for the selected DiDisc estimates in Table 4 where coefficients correspond to $\hat{\delta}_0$ from Equation 6. "Clustered p-value" reports the conventional firm clustered p-value while "bootstrap p-value" reports the corresponding wild cluster bootstrap p-value clustered at the firm level. "Firm-years" reports firm-period observations, with each firm contributing one year-end 2010 and one year-end 2014 observation. Estimates in Columns 5-8 are reported in thousands of Vietnamese Dong and reflect year-end values. Overall, the qualitative inference from Table 4 is robust to this alternative inference procedure.

Table A.7: P-values and FDR-Adjusted Q-values for the Impact of the Threshold-Based Labor Policy on All Outcomes

Outcome	[1,18]			[1,25]		
	P-values (1)	BH Q-values (2)	BKY Q-values (3)	P-values (4)	BH Q-values (5)	BKY Q-values (6)
Labor Force Size						
Total workers	0.3589	—	—	0.1723	—	—
Work-Time Flexibility						
Part-time workers	0.4259	0.5441	0.5445	0.2844	0.2844	0.2845
Casual workers	0.5441	0.5441	0.5445	0.2212	0.2844	0.2845
Unregulated Full-Time Labor						
Informal paid full-time workers	0.8244	0.8244	0.4125	0.6673	0.6673	0.3340
Unpaid full-time workers	0.1355	0.2709	0.2710	0.0939	0.1878	0.1880
Normalized Financial Returns						
Sales revenue per full-time worker	0.2105	0.2105	0.1895	0.2491	0.2491	0.1250
Gross profit per full-time worker	0.0796	0.1593	0.1595	0.0395	0.0791	0.0795
Normalized Labor Costs						
Labor costs per full-time worker	0.4353	—	—	0.2703	—	—
Productivity						
Labor productivity	0.1357	0.2713	0.2715	0.0761	0.1522	0.1525
Capital productivity	0.7516	0.7516	0.3760	0.3831	0.3831	0.1920
Full Formality						
Fully formal firm	0.9971	—	—	0.9101	—	—
Deviations from Full Formality						
Formal firm but operates informally	0.0915	0.2746	0.2750	0.0397	0.1192	0.1195
Informal firm but operates formally	0.4704	0.7056	0.4705	0.1624	0.2437	0.1625
Fully informal firm	0.9053	0.9053	0.6040	0.8338	0.8338	0.3225

Notes: Columns 1 and 4 report the raw p-values for $\hat{\delta}_0$ in Equation 6, estimated in separate regressions for each outcome variable using the optimal symmetric and asymmetric windows of [1,18] and [1, 25], respectively. P-values correspond to the results in Appendix Tables A.3, A.4, and A.5. Columns 2 and 5 report false discovery rate (FDR)-adjusted q-values using the Benjamini-Hochberg (BH) linear step-up procedure (Benjamini & Hochberg, 1995) and implemented using the Simes method (Simes, 1986). Columns 3 and 6 report "sharpened" q-values from the two-stage adaptive BH procedure of Benjamini et al. (2006) (i.e., BKY). The BH and BKY q-values are computed within each of the eight testing families listed. In some cases, BH and BKY q-values are nearly identical because BKY "sharpening" provides little refinement when there are few very small p-values across outcomes within a testing family. For singleton testing families, the multiple-testing adjustment is trivial since q-value = p-value, so only p-values are reported.

Table A.8: Impact of the Threshold-Based Labor Policy on Winsorized Firm Performance Outcomes

	Sales Revenue per Full-time Worker		Gross Profit per Full-time Worker		Labor Costs per Full-time Worker		Labor Productivity		Capital Productivity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$Post_t \times$	212,504	209,623	53,265*	60,819**	-4,354	-8,583	41,039	44,362	-0.184	-0.020
$Above10_{i,2010}$	(140,392)	(141,128)	(28,310)	(27,212)	(8,664)	(8,660)	(32,560)	(31,483)	(0.168)	(0.165)
Control mean	298,647	298,647	55,360	55,360	31,682	31,682	86,849	86,849	0.424	0.424
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.608	0.598	0.528	0.528	0.623	0.620	0.539	0.542	0.689	0.711
Optimal window	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]
Firm-years	308	374	308	374	308	374	308	374	308	374

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. Reported coefficients correspond to $\hat{\delta}_0$ from Equation 6. "Firm-years" reports firm-period observations, with each firm contributing one year-end 2010 and one year-end 2014 observation. Estimates are reported in thousands of Vietnamese Dong and reflect year-end values. Each firm performance outcome variable is winsorized at the 1st and 99th percentiles of its pre-policy distribution within the optimal asymmetric window of $[1, 25]$. Cutoffs are outcome-specific and calculated from the pooled distribution of firms above and below the threshold. They are then applied uniformly to both pre- and post-policy observations used in both optimal windows. Winsorization caps extreme values but does not drop firms or alter the selected optimal windows. "Control mean" reports the pre-policy mean of the winsorized outcome variable for the control group (i.e., firms just below the threshold) within the corresponding window. Overall, the gross profit effects remain positive and significant while that for labor productivity is positive but imprecise in the optimal asymmetric window.

B Formal Paid Worker Calculation Details

This appendix details the derivation for Equation 1 given the following definitions for each variable:

- $C_{it} \in [0, 1]$: percentage of full-time workers with a formal, written labor contract for firm i at time t
- FT_{it}: number of full-time workers
- PT_{it}: number of part-time workers
- Cas_{it}: number of casual workers
- $U_{it} \in [0, 1]$: percentage of total workers (i.e., summation of full-time, part-time, and casual workers) that are unpaid

First, the number of unpaid total workers is:

$$Unpaid_{it} = U_{it} \times (FT_{it} + PT_{it} + Cas_{it}).$$

Assuming that unpaid workers are distributed proportionately across all work-time categories (i.e., full-time, part-time, and casual workers), the share of full-time unpaid workers would be:

$$\frac{FT_{it}}{FT_{it} + PT_{it} + Cas_{it}}.$$

Thus the number of unpaid full-time workers is:

$$UnpaidFT_{it} = U_{it} \times (FT_{it} + PT_{it} + Cas_{it}) \times \frac{FT_{it}}{FT_{it} + PT_{it} + Cas_{it}} = U_{it} \times FT_{it}.$$

This means that the number of **paid**, full-time workers can be expressed as:

$$PaidFT_{it} = FT_{it} - (U_{it} \times FT_{it}) = FT_{it} \times (1 - U_{it}).$$

Then, the number of formally contracted, paid (FP) full-time workers is:

$$FP_{it} = C_{it} \times FT_{it} \times (1 - U_{it}), \text{ which is the same as Equation 1.}$$

C Analytical Sample Construction Process

This appendix details how the analytical sample (see Sections 3, 5, and 6) is constructed from the 2011 and 2015 survey samples of the UNU-WIDER Vietnam SME database, which correspond to the pre-policy period (i.e., year-end 2010) and post-policy period (i.e., year-end 2014), respectively. Because several labor variables in these surveys are recorded as proportions, converting them into headcounts can generate fractional worker counts but headcounts must be integers. Thus, I apply integrity checks and a bounded rounding procedure to the raw survey samples, exclude unreconciled firm-years, and define the resulting observations as year-specific rounded samples. Merging the pre- and post-policy rounded samples and retaining firms in both periods yields the analytical sample.

To provide context for these transformations, survey documentation and dataset variable labels indicate that the "percentage of the regular full-time labor force [that] has a formal (written down) labor contract" pertains to *full-time* workers only. Let FT_{it} be the number of full-time workers and $C_{it} \in [0, 1]$ denote the percentage of full-time workers with a formal, written labor contract for firm i at time t (the same definitions are used in Appendix B). The implied number of formally contracted full-time workers (i.e., $FT_{it} \times C_{it}$) should be an integer. To preserve the meaning of C_{it}, I do *not* round at this stage. Instead, I drop firm-years where this product is non-integer, which I treat as data entry errors. Table C.1 reports attrition by year-end period from this integrity check.

Referencing the assumptions from Appendix B, let $U_{it} \in [0, 1]$ denote the percentage of full-time workers that are not paid. I then decompose full-time workers into four mutually exclusive groups:

$$FP_{it} = C_{it} \times FT_{it} \times (1 - U_{it}) \tag{C.1}$$

$$IP_{it} = (1 - C_{it}) \times FT_{it} \times (1 - U_{it}) \tag{C.2}$$

$$FU_{it} = C_{it} \times FT_{it} \times U_{it} \tag{C.3}$$

$$IU_{it} = (1 - C_{it}) \times FT_{it} \times U_{it} \tag{C.4}$$

so that $FP_{it} + IP_{it} + FU_{it} + IU_{it} = FT_{it}$ by construction.[35] Equations C.1-C.4 estimate the number of (i) formal paid (FP), (ii) informal paid, (iii) formal unpaid, and (iv) informal unpaid full-time workers, respectively, for a given firm and period.[36]

Because C_{it} and U_{it} are proportions, Equations C.1-C.4 may yield fractional full-time worker counts. To conservatively recover integer counts for these four full-time labor categories, I use a bounded rounding procedure that reconciles small numerical discrepancies in three stages:

1. A full-time labor category value is snapped to the nearest integer only when it lies within a very small tolerance of an integer (i.e., ≤ 0.101) and the *other* three full-time labor category values are already integers.

2. Full-time labor category values are rounded up only when the fractional part is clearly above

[35] Note that Equation C.1 equals Equation 1.

[36] $FU_{it} + IU_{it}$ gives the number of unpaid full-time workers, an outcome variable described in Section 5.2.

one-half (i.e., ≥ 0.501), rounded down only when it is clearly below one-half, and unchanged when it is exactly one-half (i.e., $= 0.500$).

3. Observations with unreconcilable half-integer patterns (e.g., the fractional part of all full-time labor category values are equal to one-half, that for two out of the four full-time labor category values are equal to one-half, etc.) are dropped.

After implementing these stages, I again verify that the rounded full-time labor category values sum to the total number of full-time workers. These stages apply to the pre- and post-policy period samples. Table C.1 reports the step-by-step sample changes at each stage by year-end period.

Table C.1: Integrity Checks & Rounding Procedure Applied to Construct the Analytical Sample

Description	Pre-Policy Firms (2010)	Post-Policy Firms (2014)
Baseline sample	2,512	2,648
Integrity checks (pre-rounding)		
Drop if mismatched full-time worker counts across modules	8	—
Drop if missing C_{it}	16	—
Drop if non-integer $FT_{it} \times C_{it}$	122	208
Rounding procedure		
Stage 1: near-integer snapping	126	111
Stage 2: half-rule rounding	293	267
Stage 3: dropping unreconcilable half-type patterns	53	68
Integrity check (post-rounding)		
Drop if $FP_{it} + IP_{it} + FU_{it} + IU_{it} \neq FT_{it}$	1	—
Firms rounded (stages 1-2)	419	378
Firms dropped (integrity checks and stage 3)	200	276
Rounded samples	2,312	2,372
Analytical sample	1,439	

Notes: Entries report the number of firms dropped or rounded at each step when converting the raw survey samples for the pre- and post-policy periods into the analytical sample. Pre-rounding integrity checks exclude firms with (i) inconsistent full-time worker counts across survey modules, (ii) missing contract shares that could not be reliably imputed using data from the previous survey year, or (iii) non-integer implied full-time worker counts where the product of a headcount and a proportion should be integer-valued. The three-stage rounding procedure only applies to the four full-time labor categories. "Rounded samples" report the remaining number of firms in each period and "analytical sample" reports the number of firms in the balanced panel used to select the optimal symmetric and asymmetric windows as described in Section 5.3.

To determine if this process biases sample composition, I conduct two checks. First, I compare firms with *any* rounded full-time labor category values to firms with no rounded values within the optimal symmetric and asymmetric windows. Figure C.1 indicates that the prevalence of rounding within these optimal windows is modest and fairly balanced across periods and on both sides of the threshold, making it unlikely to influence sample composition. For example, Panel A shows that six of the 14 firms with two formal paid (FP) full-time workers pre-policy had *at least one*

rounded full-time labor category value. Panel B provides the post-policy equivalent for these same firms; it shows that among the four firms with two FP full-time workers post-policy, only one had *at least one* rounded full-time labor category value. Since such cases are limited, they are not expected to influence the research design. Second, I extend this assessment to the analytical sample construction process by comparing excluded and retained firms on the pre-policy covariates used to select the optimal windows (see Section 5.3). Table C.2 reports these comparisons separately for (i) the integrity checks and bounded rounding procedure and (ii) formation of the balanced analytical sample. Though some differences are statistically significant, standardized differences are generally small to moderate which suggests limited compositional divergence.

To confirm that the analytical sample construction process does not change pre- or post-policy period selection, I re-select these periods using the restricted unrounded samples.[37] The patterns in Figure C.2 mirror those shown in Figure 2: the average number of FP full-time workers remains stable through year-end 2010 before declining while the averages for informal paid and unpaid full-time workers decrease steadily and remain relatively flat over time, respectively. Adjacent year-end period comparisons for the average number of FP full-time workers do not differ significantly except for the decline from year-end 2010 to 2012.[38] These results establish that designating year-end 2010 and 2014 as the pre- and post-policy periods, respectively, is *not* an artifact of rounding.

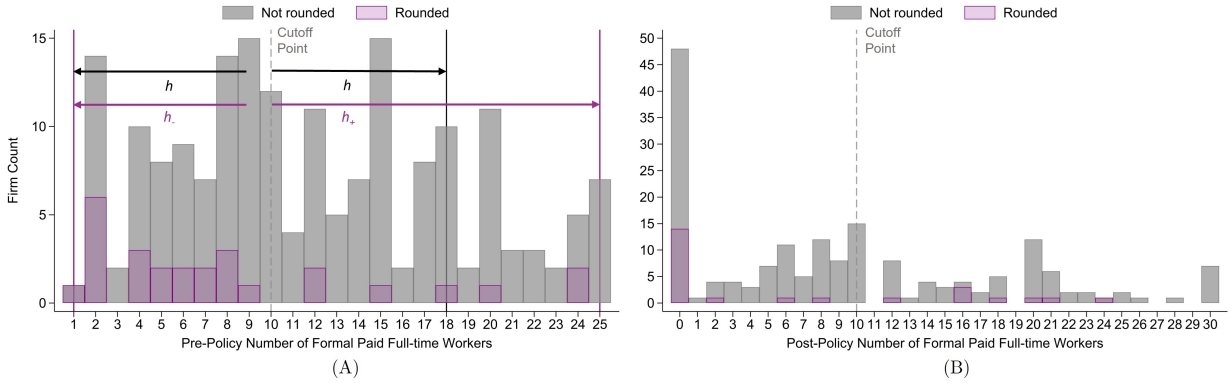


Figure C.1: Distribution of Running Variable (Rounded versus Unrounded) Pre- and Post-Policy within Optimal Windows

Notes: This figure shows the distribution of the running variable disaggregated by whether firms' number of formal paid full-time workers is affected by the analytical sample construction process. Panels A and B illustrate the pre- and post-policy periods, respectively. The black and purple vertical lines in Panel A indicate optimal symmetric and asymmetric window boundaries, respectively (see Section 5.3). Rounded values in both periods are evident but small.

The empirical findings in Section 6 are also stable to using the unrounded analytical sample.[39]

[37]The year-specific unrounded samples exclude firms that do not pass the pre-rounding integrity checks delineated in Table C.1 and those with a non-integer value in any of the four constructed full-time labor categories. The restricted unrounded samples further restrict these unrounded samples to firms with fewer than 50 total workers across years, consistent with the World Bank's definition of micro- and small- enterprises.

[38]Statistical details for these year-end comparisons are reported in the notes to Figure C.2.

[39]Restricting the pre- and post-policy unrounded samples to firms observed in both periods yields a balanced panel of 1,006 firms. This unrounded analytical sample is used to select the optimal symmetric and asymmetric windows.

Table C.2: Baseline Characteristics by Sample-Retention Status

	Means		Comparison		
	Retained	Excluded	Difference	Standardized Difference	P-value
(A) Integrity Checks & Rounding Procedure					
Owner/Manager Characteristics					
Male (1 = yes, 0 = no)	0.64	0.55	0.09	0.17	0.021
Upper secondary school completion (1 = yes, 0 = no)	0.60	0.77	-0.17	-0.37	0.000
Firm Characteristics					
Road access (1 = yes, 0 = no)	0.77	0.88	-0.11	-0.28	0.000
Age	12.56	11.05	1.51	0.17	0.014
Value added per full-time worker	68,523	98,452	-29,929	-0.12	0.232
Assets per full-time worker	528,981	521,269	7,712	0.01	0.909
Observations	2,312	200			
(B) Balanced Analytical Sample					
Owner/Manager Characteristics					
Male (1 = yes, 0 = no)	0.64	0.63	0.01	0.02	0.613
Upper secondary school completion (1 = yes, 0 = no)	0.56	0.68	-0.12	-0.26	0.000
Firm Characteristics					
Road access (1 = yes, 0 = no)	0.75	0.79	-0.04	-0.10	0.017
Age	12.89	12.01	0.88	0.09	0.032
Value added per full-time worker	65,072	74,212	-9,141	-0.10	0.021
Assets per full-time worker	489,932	593,348	-103,417	-0.09	0.041
Observations	1,439	873			

Notes: Characteristics are measured in the pre-policy period (i.e., year-end 2010) and correspond to the covariates used for optimal window selection in Section 5.3. Panel A compares firms retained after the integrity checks and rounding procedure reported in Table C.1 with firms excluded as a result of the analytical sample construction process. Panel B restricts the sample to firms retained after that process and compares firms included in the balanced analytical sample with those excluded. Differences and standardized differences are calculated as the retained-group mean minus the excluded-group mean. Standardized differences divide the mean difference by the square root of the average of the two group variances. P-values are from unequal variance two-sample t-tests.

Figure C.3 presents the running variable (i.e., number of FP full-time workers) distribution for both periods leveraging the optimal windows derived from the unrounded analytical sample. In applying the local randomization approach (see Section 5.3 for details) to this unrounded analytical sample, the resulting optimal symmetric and asymmetric windows are [2,17] and [1,30], respectively.[40] The

[40]To preserve the local nature of the analysis, the modified greedy expansion procedure caps the right-side bandwidth at 20, limiting the optimal asymmetric window to 30 FP full-time workers. Accordingly, the selected optimal

optimal symmetric window defines a subsample of 103 firms—52 in the treatment group and 51 in the control—while the optimal asymmetric window yields 151 firms—100 treatment and 51 control. McCrary (2008) density tests reveal no evidence of spurious discontinuity at the cutoff post-policy. Thus, the absence of bunching behavior is consistent with the results reported in Section 6.1.[41]

Re-estimating Equation 6 for the selected firm outcomes in Table 4 within the optimal windows derived from the unrounded analytical sample yields coefficients with similar signs and significance levels (see Table C.3). Specifically, the impact of the higher financial penalty introduced by the 2012 employee threshold policy on unpaid full-time labor is negative and significant for both optimal windows. Also, the probability that a firm is "formal but operates informally" and gross profit per full-time worker both remain positive and significant under the optimal asymmetric window. Yet, labor productivity is no longer significant. Therefore, the results in Section 6.2 are largely robust to using the unrounded analytical sample. Table C.4 reports these estimates' associated p-values and false discovery rate-adjusted q-values. Though evidence of policy-related effects remains, firm outcomes surviving the multiple hypothesis adjustment differ from those in Appendix Table A.7.

Table C.3: Impact of Threshold-Based Labor Policy on Selected Firm Outcomes Based on the Unrounded Analytical Sample

	Unpaid Full-time Workers		Formal Firm But Operates Informally		Gross Profit per Full-time Worker		Labor Productivity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Post_t \times$	-0.609**	-0.391**	0.144	0.173**	41,062	103,887*	33,068	79,522
$Above10_{i,2010}$	(0.271)	(0.198)	(0.088)	(0.073)	(26,176)	(61,206)	(34,355)	(64,510)
Control mean	0.412	0.412	0.039	0.039	60,449	60,449	94,440	94,440
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.680	0.696	0.636	0.641	0.500	0.498	0.501	0.500
Optimal window	[2,17]	[1,30]	[2,17]	[1,30]	[2,17]	[1,30]	[2,17]	[1,30]
Firm-years	206	302	206	302	206	302	206	302

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. Reported coefficients correspond to $\hat{\delta}_0$ from Equation 6. "Control mean" reports the pre-policy mean of the outcome variable for the control group (i.e., firms just below the threshold) within the corresponding optimal window. "Firm-years" reports firm-period observations, with each firm contributing one year-end 2010 and one year-end 2014 observation. Estimates in Columns 5-8 are reported in thousands of Vietnamese Dong and reflect year-end values. Only firm outcomes with statistically significant effects are presented. FDR-adjusted q-values for these estimates are listed in Appendix Table C.4. The unrounded analytical sample excludes firms with a non-integer value in any of the four constructed full-time labor categories.

asymmetric window of $[1, 30]$ for the unrounded analytical sample reaches this pre-specified upper endpoint.

[41] Statistical details for these McCrary (2008) density tests are reported in the notes to Figure C.3.

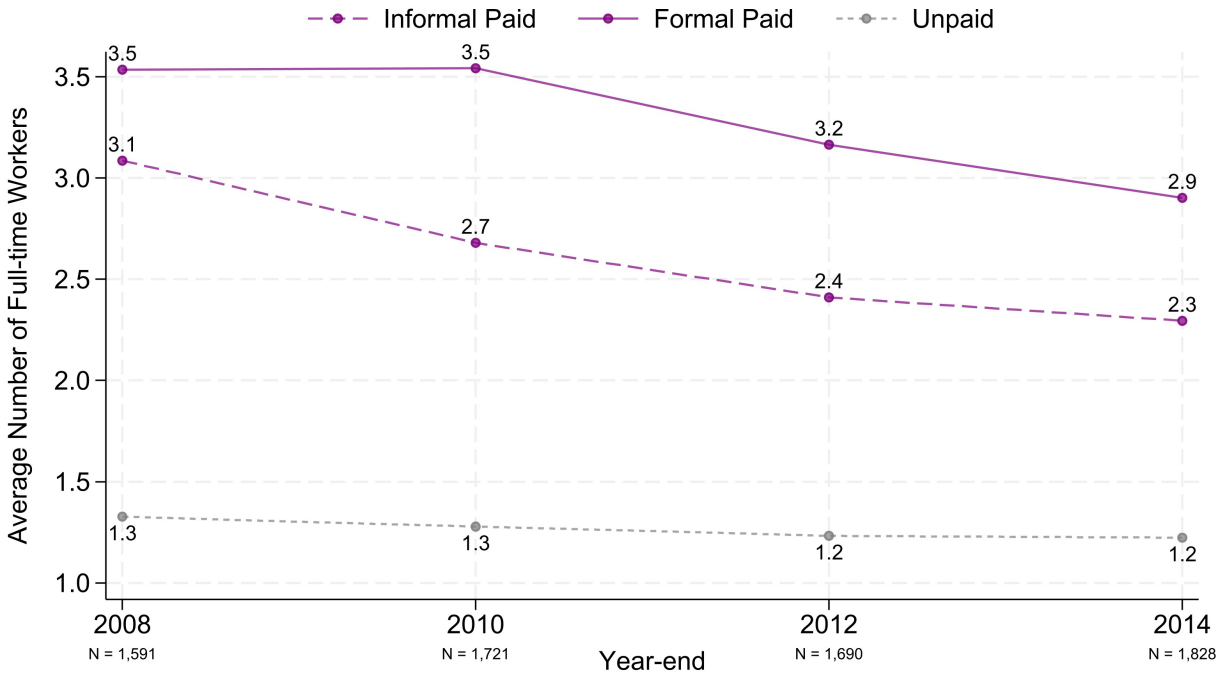


Figure C.2: Average Number of Full-time Workers by Labor Category for Restricted Unrounded Samples by Year-End Period

Notes: This figure shows the average number of (i) informal paid, (ii) formal paid, and (iii) unpaid full-time workers at each year-end period within the unrounded samples. These unrounded samples exclude firms with a non-integer value in any of the four constructed full-time labor categories. These samples are also restricted to firms with fewer than 50 total workers across years, consistent with the World Bank's definition of micro- and small-enterprises. Each year-end period's sample size is indicated below the corresponding year. Two-sided tests of mean differences—with standard errors clustered at the firm level—compare average FP full-time employment across year-end periods. Mean differences are not statistically significant between year-end 2008 and 2010 ($p = 0.9756$) or year-end 2012 and 2014 ($p = 0.1544$). However, the decline from year-end 2010 to 2012 is statistically significant ($p = 0.0469$). These patterns and findings are consistent with those shown in Figure 2 and discussed in Section 3

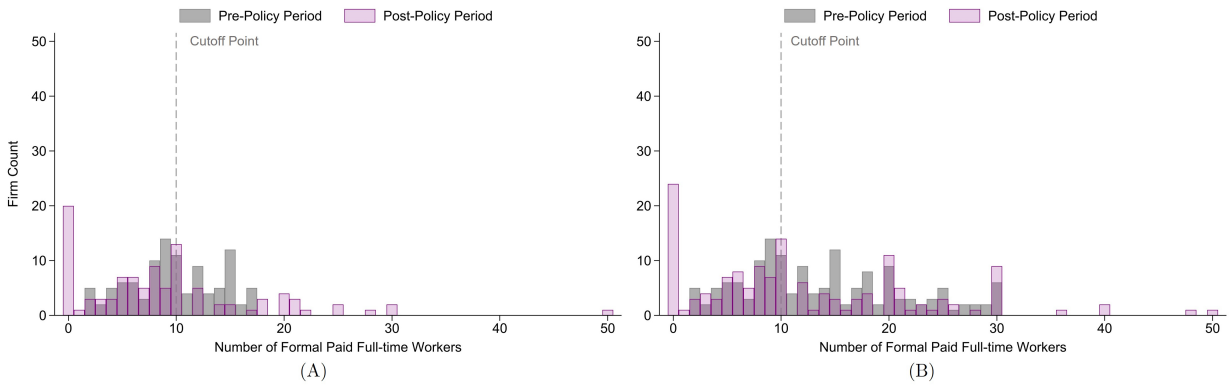


Figure C.3: Distribution of Running Variable Pre- and Post-Policy within Optimal Windows Based on the Unrounded Analytical Sample

Notes: This figure shows the distribution of firms around the 10-formal paid (FP) employee threshold, with Panels A and B using the optimal symmetric and asymmetric windows, respectively, derived from the unrounded analytical sample. This sample excludes firms with a non-integer value in any of the four constructed full-time labor categories. In Panel A, the McCrary (2008) density test detects a significant negative discontinuity at the cutoff in the pre-policy period ($T = -2.378$, $p = 0.0177$) but the post-policy estimate is not statistically significant ($T = -1.045$, $p = 0.2961$). Panel B shows a significant negative pre-policy discontinuity ($T = -2.004$, $p = 0.0451$) while the post-policy estimate is again insignificant ($T = -0.857$, $p = 0.3916$). These results align with those reported in Section 6.1. Using the *full* unrounded analytical sample—which includes 197 treatment firms and 809 control firms—shows the same pattern: McCrary (2008) density tests yield $T = -1.839$ ($p = 0.0660$) pre-policy and $T = -0.539$ ($p = 0.5899$) post-policy.

Table C.4: P-values and FDR-Adjusted Q-values for the Impact of the Threshold-Based Labor Policy on All Outcomes Based on the Unrounded Analytical Sample

Outcome	[2,17]			[1,30]		
	P-values (1)	BH Q-values (2)	BKY Q-values (3)	P-values (4)	BH Q-values (5)	BKY Q-values (6)
Labor Force Size						
Total workers	0.5370	—	—	0.1186	—	—
Work-Time Flexibility						
Part-time workers	0.4317	0.7831	0.7835	0.3584	0.3584	0.3585
Casual workers	0.7831	0.7831	0.7835	0.1466	0.2931	0.2935
Unregulated Full-Time Labor						
Informal paid full-time workers	0.3997	0.3997	0.2000	0.6794	0.6794	0.3400
Unpaid full-time workers	0.0264	0.0529	0.0530	0.0497	0.0994	0.0995
Normalized Financial Returns						
Sales revenue per full-time worker	0.9170	0.9170	0.4590	0.2904	0.2904	0.2250
Gross profit per full-time worker	0.1198	0.2396	0.2400	0.0917	0.1834	0.1835
Normalized Labor Costs						
Labor costs per full-time worker	0.6890	—	—	0.2082	—	—
Productivity						
Labor productivity	0.3381	0.6085	0.6085	0.2196	0.2696	0.2700
Capital productivity	0.6085	0.6085	0.6085	0.2696	0.2696	0.2700
Full Formality						
Fully formal firm	0.7676	—	—	0.8298	—	—
Deviations from Full Formality						
Formal firm but operates informally	0.1027	0.3082	0.3085	0.0188	0.0565	0.0570
Informal firm but operates formally	0.3250	0.4876	0.4455	0.3023	0.4534	0.3025
Fully informal firm	0.9947	0.9947	0.6635	0.8000	0.8000	0.5335

Notes: Columns 1 and 4 report the raw p-values for $\hat{\delta}_0$ in Equation 6, estimated in separate regressions for each outcome variable using the optimal symmetric and asymmetric windows of [2,17] and [1,30], respectively, derived from the unrounded analytical sample. This sample excludes firms with a non-integer value in any of the four constructed full-time labor categories. See Table A.7 for details about the false discovery rate (FDR)-adjusted q-values reported in Columns 2, 3, 5, and 6. "Firm-years" reports firm-period observations, with each firm contributing one year-end 2010 and one year-end 2014 observation.

D Proofs of Theoretical Predictions

This appendix provides the derivations and intermediate results used to support Lemma 1 and Theorems 1, 2, and 3. These steps are omitted from Section 4 but included here for completeness.

D.1 Proof of Lemma 1 and Theorem 1

Let $\pi_i^s(x_i^s; P_r)$ denote the objective function in Equation 5 evaluated at a fixed feasible choice vector $x_i^s \in \mathcal{X}^s$. For Strategy N, firm i crosses the 10-FP employee threshold but does not submit internal work regulations (IWRs) (i.e., $L_{FP,i}^N \geq 10$ and $T_i^N = 0$). Therefore, the expected financial penalty $\phi(R_i^N)P_r$ is active. Holding the choice vector x_i^N fixed:

$$\pi_i^N(x_i^N; P_1) - \pi_i^N(x_i^N; P_0) = -\phi(R_i^N)(P_1 - P_0) \leq 0,$$

with strict inequality whenever $\phi(R_i^N) > 0$.

Since this weak inequality holds for every fixed feasible choice vector $x_i^N \in \mathcal{X}^N$, the maximized payoff under Strategy N is weakly decreasing in the financial penalty:

$$\Pi_i^N(P_1) = \max_{x_i^N \in \mathcal{X}^N} \pi_i^N(x_i^N; P_1) \leq \max_{x_i^N \in \mathcal{X}^N} \pi_i^N(x_i^N; P_0) = \Pi_i^N(P_0).$$

By contrast, Strategies B, F, and S are not directly affected by P_r. Under Strategy B, firm i remains below the 10-FP employee threshold so the expected financial penalty is inactive because $L_{FP,i}^B < 10$. Moreover, under Strategies F and S, firm i crosses the threshold and submits IWRs so the expected financial penalty is inactive because $T_i^F = T_i^S = 1$. Thus:

$$\Pi_i^B(P_1) = \Pi_i^B(P_0), \qquad \Pi_i^F(P_1) = \Pi_i^F(P_0), \qquad \Pi_i^S(P_1) = \Pi_i^S(P_0).$$

It follows that for any alternative regulatory strategy $q \in \{B, F, S\}$:

$$\Pi_i^q(P_1) - \Pi_i^N(P_1) \geq \Pi_i^q(P_0) - \Pi_i^N(P_0).$$

This proves Lemma 1: the increase in the financial penalty for IWR non-submission from P_0 to P_1 weakly lowers the relative payoff from crossing the threshold *without* submitting IWRs.

To prove Theorem 1, let the net payoff gain from crossing the 10-FP employee threshold under strategy $q \in \{N, F, S\}$ relative to remaining below the threshold (Strategy B) be defined as:

$$\Delta_{i,q}(P_r) \equiv \Pi_i^q(P_r) - \Pi_i^B(P_r).$$

If, in the post-policy regime

$$\max_{q \in \{N,F,S\}} \Delta_{i,q}(P_1) < 0,$$

then every feasible crossing strategy yields a lower payoff than remaining below the threshold:

$$\Pi_i^q(P_1) < \Pi_i^B(P_1), \qquad \forall q \in \{N, F, S\}.$$

Therefore, Strategy B yields the highest payoff among the available regulatory strategies:

$$s_i^*(P_1) = B.$$

Since the feasible set for Strategy B requires $L_{FP,i}^B < 10$, firm i remains below the 10-FP employee threshold in the post-policy regime. Because FP employment is discrete and nonnegative, Strategy B allows any:

$$L_{FP,i}^{B*}(P_1) \in \{0, \ldots, 9\}.$$

To obtain an aggregate bunching implication, let $\mu(\cdot)$ denote the mass of firms satisfying a given condition. Because bunching concerns firms located near the cutoff, define an integer neighborhood below the threshold by $k \in \mathbb{N}$, where $1 \leq k < 10$:[42]

$$B_k(P_1) = \left\{ i : s_i^*(P_1) = B \text{ and } L_{FP,i}^{B*}(P_1) \in \{10 - k, \ldots, 9\} \right\}.$$

If $\mu(B_k(P_1)) > 0$, then a positive mass of firms optimally locates within the k-unit neighborhood below the threshold in the post-policy regime. For small values of k, this corresponds to bunching just below the 10-FP employee threshold.

This aggregate implication requires two conditions. First, firms for which Strategy B is optimal must have below-threshold optima located near the cutoff rather than *far below* it. Second, because adjustment costs are not modeled explicitly, firms must be able to adjust FP employment downward at sufficiently low costs. Therefore, threshold avoidance is a firm-level response while bunching is an aggregate manifestation that arises only when a positive mass of firms optimally chooses below-threshold employment levels relatively close to the cutoff.

D.2 Proof of Theorem 2

Theorem 2 characterizes how the increased financial penalty changes the relative attractiveness of threshold-excluded labor inputs for firms located just below the 10-FP employee threshold. The phrase "just below the threshold" refers to firms with $L_{FP,i} < 10$ for whom the relevant FP labor expansion would move them to $L_{FP,i} \geq 10$. Thus, Theorem 2 applies to below-threshold firms at the crossing margin, not to firms already subject to the IWR requirement.

Consider a local reallocation of labor that holds the firm's target level of effective labor input

[42] Strategy B allows firms to locate anywhere below the threshold (i.e., $L_{FP,i}^{B*}(P_1) \in \{0, \ldots, 9\}$). However I restrict $k < 10$ so that the bunching set excludes the structural mass point at zero FP employment (see Figure A.2), which reflects broader avoidance of FP labor rather than local bunching *just below* the threshold. Allowing $k = 10$ would conflate local threshold bunching with general below-threshold operation.

fixed, conditional on a given formality status. From Equation 2:

$$dL_i^{eff} = dL_{FP,i} + \theta dL_{IP,i} + \delta dL_{U,i} = 0.$$

Compared to a counterfactual allocation that employs more FP labor and crosses the threshold, suppose a below-threshold firm at the crossing margin limits regulated FP employment such that $dL_{FP,i} < 0$. Holding effective labor input fixed then requires:

$$\theta dL_{IP,i} + \delta dL_{U,i} = -dL_{FP,i} > 0.$$

Thus, labor from at least one threshold-excluded labor category—informal paid labor, unpaid labor, or both—must increase to preserve the firm's effective labor input.

The firm evaluates reallocation using marginal costs per unit of effective labor. Given that FP, informal paid, and unpaid workers contribute 1, θ, and δ units of effective labor, respectively (see Equation 2), the relevant costs are:

$$c_{FP} \equiv C_{FP}, \qquad c_{IP} \equiv \frac{C_{IP}}{\theta} = \frac{w_{IP}}{\theta}, \qquad c_U(L_{U,i}) \equiv \frac{\psi'(L_{U,i})}{\delta}.$$

These marginal costs capture worker-level labor costs but not the discrete regulatory cost associated with crossing the threshold. For a below-threshold firm at the crossing margin, the FP worker that triggers $L_{FP,i} \geq 10$ carries an additional regulatory wedge. To define this wedge formally, let $\mathcal{G}_i^s(P_r)$ denote the regulatory and compliance costs for strategy s:

$$\mathcal{G}_i^s(P_r) \equiv C_R R_i^s + C_K K_i^s + C_T \cdot \mathbb{1}\{L_{FP,i}^s \geq 10 \wedge T_i^s = 1\} + \phi(R_i^s) P_r \cdot \mathbb{1}\{L_{FP,i}^s \geq 10 \wedge T_i^s = 0\}.$$

This component includes government registration costs, labor contract compliance costs, threshold-induced regulatory costs, and the expected financial penalty for IWR non-submission. It excludes paid labor costs and the non-monetary cost of unpaid labor. The threshold-related shadow cost of moving from a below-threshold allocation to the least costly feasible crossing strategy is defined as the following finite-difference regulatory wedge:[43]

$$\tau_i(P_r) \equiv \min_{q \in \{N,F,S\}} \left\{ \mathcal{G}_i^q(P_r) - \mathcal{G}_i^B(P_r) \right\}.$$

Because inspection risk is held fixed across regimes and $P_1 > P_0$, only Strategy N is directly affected by the higher financial penalty. For any fixed Strategy N choice,

$$\mathcal{G}_i^N(P_1) = \mathcal{G}_i^N(P_0) + \phi(R_i^N)(P_1 - P_0) \geq \mathcal{G}_i^N(P_0)$$

with strict inequality when $\phi(R_i^N) > 0$. By contrast, Strategies B, F, and S avoid the expected financial penalty: B remains below the threshold while F and S involve submitting IWRs. Therefore,

[43] Because the threshold is discrete, $\tau_i(P_r)$ is not a marginal derivative or Lagrange multiplier.

for all $q \in \{B, F, S\}$,

$$\mathcal{G}_i^q(P_1) = \mathcal{G}_i^q(P_0) \equiv \mathcal{G}_i^q.$$

Define the least-cost IWR-submission strategy as

$$\overline{\mathcal{G}}_i \equiv \min\{\mathcal{G}_i^F, \mathcal{G}_i^S\}.$$

Then the threshold-related shadow cost can be written as:

$$\tau_i(P_r) = \min\{\mathcal{G}_i^N(P_r), \overline{\mathcal{G}}_i\} - \mathcal{G}_i^B.$$

Since $\mathcal{G}_i^N(P_r)$ weakly increases from P_0 to P_1 while $\overline{\mathcal{G}}_i$ and $\mathcal{G}_i^B$ are unchanged, it follows that:

$$\tau_i(P_1) \geq \tau_i(P_0).$$

The inequality is strict when the increase in Strategy N's expected financial penalty increases the minimum crossing cost. It holds with equality when the least-cost crossing option is instead pinned down by Strategy F or S, or when Strategy N faces zero perceived inspection risk (i.e. $\phi(R_i^N) = 0$).

The threshold-adjusted marginal per-effective-labor cost of the FP worker that triggers the IWR requirement is therefore defined as:

$$c_{TFP}(P_r) = C_{FP} + \tau_i(P_r).$$

Substitution toward threshold-excluded labor is cost-reducing when at least one threshold-excluded labor category provides effective labor at a lower marginal cost than threshold-adjusted FP labor:

$$\min\left\{\frac{w_{IP}}{\theta}, \frac{\psi'(L_{U,i})}{\delta}\right\} < c_{TFP}(P_r).$$

Because $\tau_i(P_1) \geq \tau_i(P_0)$, this inequality is weakly more likely to hold in the post-policy regime. Hence, for firms seeking to maintain effective labor input while limiting regulated FP employment, the higher financial penalty increases the relative attractiveness of substituting toward unregulated labor arrangements: $L_{IP,i}$, $L_{U,i}$, or both.

D.3 Proof of Theorem 3

Focusing only on firms that cross the 10-FP employee threshold, the relevant crossing strategies are: crossing without submitting IWRs (Strategy N), full compliance (Strategy F), and selective compliance (Strategy S). Firm i chooses selective compliance in the post-policy regime when:

$$\Pi_i^S(P_1) > \max\{\Pi_i^N(P_1), \Pi_i^F(P_1)\}.$$

Equivalently, let the payoff gains from selective compliance relative to the other crossing strate-

gies be defined as:

$$\Delta_i^{SN}(P_1) \equiv \Pi_i^S(P_1) - \Pi_i^N(P_1), \qquad \Delta_i^{SF}(P_1) \equiv \Pi_i^S(P_1) - \Pi_i^F(P_1).$$

Strategy S is chosen whenever:

$$\min\{\Delta_i^{SN}(P_1), \Delta_i^{SF}(P_1)\} > 0.$$

The comparison between Strategy S and N follows from Lemma 1. Strategy N activates the expected financial penalty because the firm crosses the threshold without submitting IWRs, whereas Strategy S avoids the expected financial penalty by submitting IWRs. Thus:

$$\Pi_i^S(P_1) - \Pi_i^N(P_1) \geq \Pi_i^S(P_0) - \Pi_i^N(P_0).$$

Therefore, the increase from P_0 to P_1 makes selective compliance weakly more attractive relative to crossing the threshold without submitting IWRs.

Conditional on crossing the threshold and submitting IWRs, the relevant comparison is between Strategy S and F. These regulatory strategies require $L_{FP,i} \geq 10$, $T_i = 1$, and $R_i = 1$ so both incur the registration cost C_R and the threshold-compliance cost C_T. Neither strategy pays the expected financial penalty for IWR non-submission. These common terms cancel in the payoff comparison.

The key regulatory difference is therefore labor contract compliance: Strategy F requires $K_i = 1$ and incurs the fixed cost C_K while Strategy S requires $K_i = 0$ and avoids incurring C_K. This also changes feasible firm labor composition because full compliance rules out informal paid labor while selective compliance permits informal paid labor and the informal component of unpaid labor.

Let $\widetilde{\Pi}_i^s$ represent the maximized operating payoff under strategy s, defined as output net of paid labor costs and the non-monetary cost of unpaid labor, excluding regulatory and compliance costs:

$$\widetilde{\Pi}_i^s \equiv A_i \Gamma(R_i^s*, K_i^s*) \left(L_i^{eff,s*}\right)^{\alpha} - C_{FP} L_{FP,i}^{s*} - C_{IP} L_{IP,i}^{s*} - \psi\left(L_{U,i}^{s*}\right).$$

For Strategies S and F, the full maximized payoffs in the post-policy regime can be written as:

$$\Pi_i^S(P_1) = \widetilde{\Pi}_i^S - C_R - C_T,$$

and

$$\Pi_i^F(P_1) = \widetilde{\Pi}_i^F - C_R - C_T - C_K.$$

Thus, the payoff difference between selective and full compliance is:

$$\Pi_i^S(P_1) - \Pi_i^F(P_1) = \widetilde{\Pi}_i^S - \widetilde{\Pi}_i^F + C_K.$$

Selective compliance is therefore preferred to full compliance whenever:

$$\widetilde{\Pi}_i^S - \widetilde{\Pi}_i^F + C_K > 0.$$

Equivalently:

$$C_K > \widetilde{\Pi}_i^F - \widetilde{\Pi}_i^S.$$

Because Strategy F sets $(R_i, K_i) = (1, 1)$ while Strategy S sets $(R_i, K_i) = (1, 0)$, the relevant productivity comparison is:

$$\Gamma(1,1) = e^{\gamma_R + \gamma_K} \quad \text{and} \quad \Gamma(1,0) = e^{\gamma_R}.$$

Therefore, full compliance can generate an operating advantage through the formality-productivity shifter, strictly so when $\gamma_K > 0$. Selective compliance is therefore not mechanically preferred; firm i chooses it only when the savings from avoiding C_K and the value of labor flexibility exceed the operating advantage from full compliance.

Finally, note that $L_{U,i} > 0$ should not by itself be interpreted as labor contract non-compliance. Because $L_{U,i}$ aggregates formal and informal unpaid workers, positive unpaid labor can be consistent with full compliance if those unpaid workers are formally contracted. Labor contract non-compliance is represented by $K_i = 0$ under Strategy S, meaning that *at least one* worker (e.g., FP, informal paid, or unpaid) employed by firm i lacks a formal, written labor contract.

E Diagnostics for the Validity of Difference-in-Differences

This appendix presents diagnostic analyses that evaluate the suitability of a standard difference-in-differences (DiD) approach and assess the timing of the main difference-in-discontinuities (DiDisc) results. First, I examine the sensitivity of global DiD estimates to sample composition around the 10-formal paid (FP) employee threshold. I then analyze pre-policy dynamics near the cutoff using DiD event-study evidence to evaluate local continuity. Finally, I estimate a dynamic DiDisc event-study specification to test whether the key effects from Section 6.2 emerge in the post-policy period rather than reflecting earlier discontinuities at the threshold. Altogether, these analyses show that a global DiD approach is ill-suited for this application, motivate the static DiDisc design in Section 5, and provide complementary timing evidence for the main empirical results.

E.1 Sensitivity of Global DiD Estimates to Sample Composition

As a first diagnostic, I evaluate how a standard global DiD estimator performs under alternative sample restrictions around the 10-FP employee threshold. This exercise is motivated by concerns about the interpretation of the parallel trends assumption when treatment is defined by a size-based threshold. In this setting, treatment assignment is endogenously determined by a discrete running variable (i.e., the firm's number of FP full-time workers in the pre-policy period) and firms on either side of the cutoff differ systematically in regulatory incentives and outcome-relevant characteristics. As a result, including firms far from the threshold can cause global DiD estimates to conflate policy effects with underlying size-related heterogeneity, even in the absence of differential pre-trends.

To illustrate this issue, I estimate a two-way fixed effects (TWFE) DiD specification that compares firms above and below the 10-FP employee threshold before and after the implementation of the 2012 employee threshold policy. Equation E.1 defines this regression specification.

$$Y_{it} = \beta_0(Above10_{i,2010} \times Post_t) + \mu_i + \gamma_t + \epsilon_{it} \tag{E.1}$$

In Equation E.1, Y_{it} is an outcome for firm i at time t, $Above10_{i,2010}$ equals 1 if firm i has 10 or more FP full-time workers pre-policy and 0 otherwise, and $Post_t$ equals 1 in the post-policy period (i.e., year-end 2014) and 0 in the pre-policy period (i.e., year-end 2010). The terms μ_i and γ_t represent firm and year fixed effects, respectively, and ϵ_{it} is the error term. The coefficient of interest is β_0; it captures the average differential change in outcomes for firms above the threshold relative to those below it post-policy. Standard errors are clustered at the firm level. Equation E.1 intentionally omits the running variable, reflecting the global comparison inherent in a DiD design.

Equation E.1 is estimated for each outcome variable described in Section 5.2 across a sequence of nested windows defined by firms' pre-policy FP full-time employment. The broadest window is the analytical sample (see Appendix C) which includes firms with up to 321 FP full-time workers pre-policy and provides the global DiD comparison. The second window excludes firms with zero FP full-time workers, thereby removing the dominant mass point in the running variable's distribution shown in Figure A.2 and mitigating concern that the estimates are driven by comparisons between

employing and owner-only firms. The third window further restricts the sample to firms with fewer than 50 FP full-time workers, approximating the World Bank's definition of micro- and small-enterprises based on total employment.[44] The fourth and fifth windows are the optimal asymmetric and symmetric windows of $[1, 25]$ and $[1, 18]$, respectively (see Section 5.3). The optimal asymmetric window maximizes sample size subject to pre-policy covariate balance while the optimal symmetric window is the widest symmetric window satisfying covariate balance. The sixth window imposes a fixed symmetric restriction of ± 5 FP full-time workers around the threshold, providing a transparent comparison of firms equidistant from the cutoff.[45] This sequence of windows directly evaluates how global DiD estimates vary with sample composition and with the inclusion of firms farther from the threshold. However, restricting the sample just to firms closer to the threshold does *not* resolve the underlying design limitation: the DiD design completely ignores the running variable and therefore cannot account for pre-existing discontinuities or size-related trends at the cutoff.

Figure E.1 plots the estimated $\hat{\beta}_0$ coefficients for all outcomes analyzed in the DiDisc analysis. The estimates differ substantially in magnitude, sign, and statistical significance as the sample is restricted toward the threshold. This instability suggests that global DiD estimates are driven by comparisons with firms far from the threshold—where parallel trends are less plausible—rather than by firms near the threshold. Consequently, the global DiD estimator does not recover a well-defined causal effect as the observed sensitivity indicates that the DiD estimand itself is ill-defined.

Notably, when Equation E.1 is estimated on the same samples that satisfy covariate balance at the cutoff—namely the optimal symmetric and asymmetric windows—only one of the 28 estimated treatment effects is statistically significant and it does not correspond to any outcome identified as significant under the DiDisc design. This occurs despite these being the samples in which the DiDisc design identifies economically meaningful and statistically significant policy effects (see Section 6.2). The contrast indicates that DiD's failure is not attributable to limited statistical power nor sample size, but rather to its inability to condition on the policy-induced discontinuity at the threshold.

E.2 Pre-Policy Continuity at the Threshold

Appendix E.1 demonstrates that conventional TWFE DiD estimates are sensitive to the estimation sample when treatment is defined by a size-based threshold. One potential explanation for this is a violation of the parallel trends assumption. As a second diagnostic, I assess that possibility directly by analyzing whether firms above and below the 10-FP employee threshold exhibit differential outcome trends *prior to* the 2012 employee threshold policy. Using an event-study specification, I test for pre-policy differences in outcomes between firms on either side of the cutoff. Although the DiDisc design employed in Section 5 does not require parallel trends between treatment and control groups, its validity depends on a related but weaker condition: absent the policy change, outcomes must evolve smoothly across the threshold over time. This continuity condition is inherently local and differs from the parallel trends assumption required for the DiD design. To determine whether

[44] See https://ieg.worldbankgroup.org/sites/default/files/Data/Evaluation/files/SME_Synthesis.pdf for details.

[45] The number of unique firms in the first through sixth windows is 1,439, 308, 243, 187, 154, and 107, respectively.

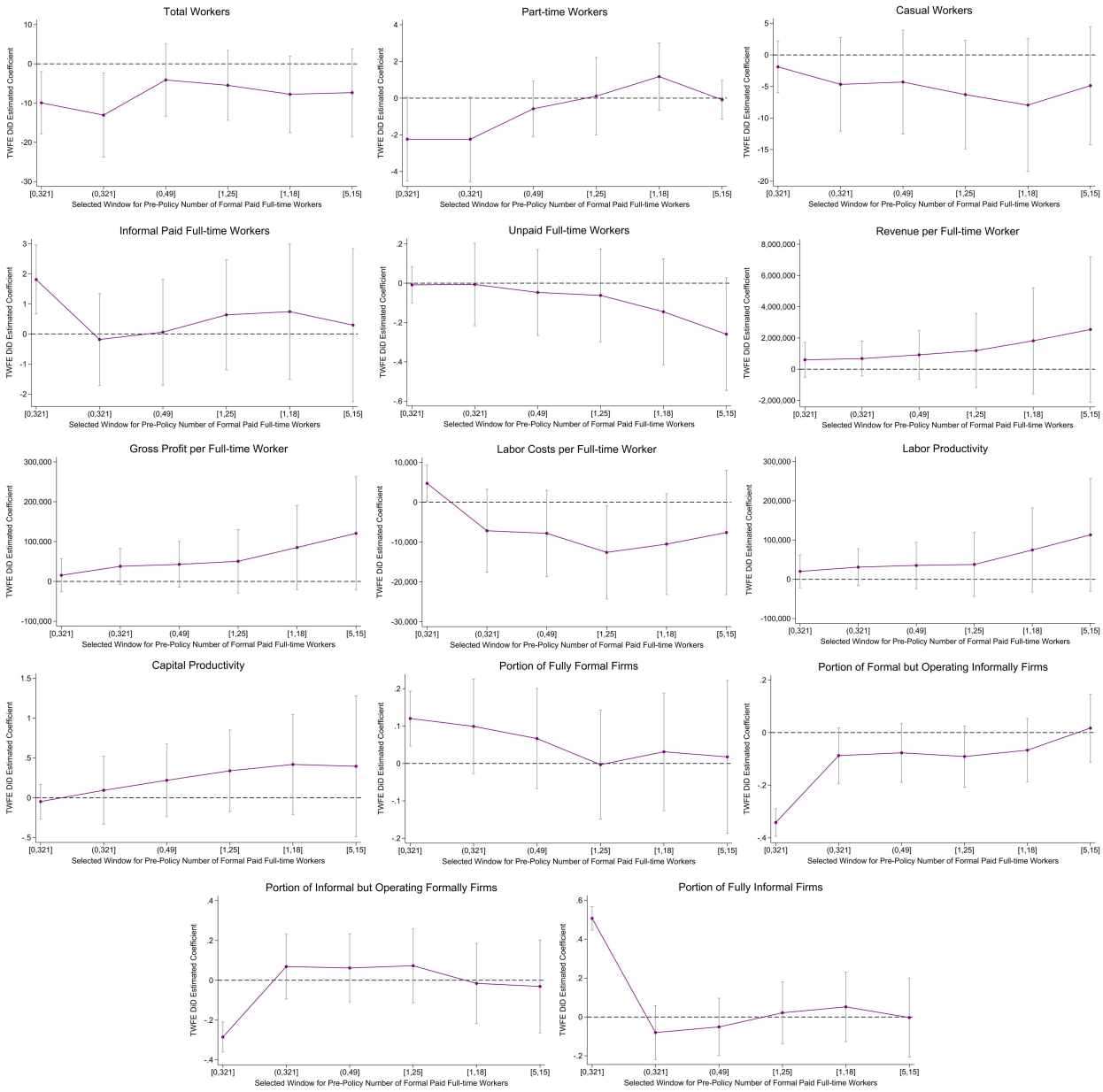


Figure E.1: Sensitivity of TWFE DiD Estimates Around the 10-Formal Paid Employee Threshold

Notes: This figure plots the estimated two-way fixed effects (TWFE) difference-in-differences (DiD) $\hat{\beta}_0$ coefficient from Equation E.1 for all the outcomes analyzed in the difference-in-discontinuities (DiDisc) analysis (see Section 5). Each point reports $\hat{\beta}_0$ on progressively narrower samples defined by a firm's number of formal paid (FP) full-time workers pre-policy. Whiskers denote 95 percent confidence intervals. Overall, estimates vary widely across selected windows, indicating that global DiD results are sensitive to sample composition around the threshold. The maximum number of FP full-time workers pre-policy in the analytical sample (see Appendix C) is 321. All size-normalized firm economic account metrics are reported in thousands of Vietnamese Dong.

it is plausible in this setting, I analyze outcome dynamics in periods *preceding* the policy's passage and implementation. Given the structure of the UNU-WIDER Vietnam SME database (see Section 3) and the policy's timeline (see Figure 1), I focus on year-end 2008 and 2010 as pre-policy periods.

To examine local pre-trends, I estimate the event-study specification in Equation E.2 separately for each outcome described in Section 5.2.

$$Y_{it} = \sum_{k \neq 2010} \beta_k \left(Above10_{i,2010} \times \mathbb{1}\{t = k\} \right) + \mu_i + \gamma_t + \epsilon_{it} \tag{E.2}$$

In Equation E.2, Y_{it} is an outcome for firm i at time t, $Above10_{i,2010}$ equals 1 if firm i has 10 or more FP full-time workers pre-policy and 0 otherwise, and $k \in \{2008, 2010, 2012, 2014\}$ represents year-end periods. The terms μ_i and γ_t capture firm and year fixed effects, respectively, and ϵ_{it} is the error term. The year-end 2010 period is omitted and serves as the reference period. Standard errors are clustered at the firm level.

Although the regression includes observations from multiple year-end periods, inference regarding pre-policy dynamics focuses only on the coefficient β_{2008}, which captures differences in outcomes between firms above and below the threshold in year-end 2008 relative to 2010. Observations from year-end 2008 and 2012 are included only for firms within the analytical sample (see Appendix C), which are observed in both the pre-policy (i.e., year-end 2010) and post-policy (i.e., year-end 2014) periods.[46] This restriction ensures that resulting comparisons reflect within-firm dynamics rather than changes in sample composition. Estimation is conducted using three samples: the analytical sample, the optimal asymmetric window of $[1, 25]$, and the optimal symmetric window of $[1, 18]$.[47]

Figure E.2 plots the estimated coefficient $\hat{\beta}_{2008}$ for each outcome variable and sample definition. Across all outcomes and samples, the estimated pre-policy differences are predominantly small and statistically insignificant, particularly within the optimal windows. Estimates are also more tightly centered around zero as the sample is restricted from the analytical sample to the optimal windows. This pattern is consistent with the interpretation that firms closer to the cutoff are more comparable along both observed and unobserved dimensions. Accordingly, there is no systematic evidence of pre-policy discontinuities or differential trends at the threshold across the optimal windows.

Importantly, these findings do *not* validate a global DiD design. Rather, they highlight a crucial limitation of standard pre-trends diagnostics in threshold-based settings. Though firms just above and below the cutoff exhibit similar pre-policy outcome dynamics, Appendix E.1 shows that global TWFE DiD estimates remain highly sensitive to window selection and sample composition. This instability arises not from violations of the parallel trends assumption, but from how the DiD design averages across heterogeneous firms whose outcomes and incentives differ systematically with size. When treatment is defined by an endogenous firm-size threshold, the DiD estimand itself is not well defined globally. This is because firms on either side of the cutoff do not form a stable counterfactual comparison group even when conventional event-study tests fail to reject parallel trends.

[46] These periods were selected based on analyses carried out in Section 3.

[47] The optimal symmetric and asymmetric windows are selected using the local randomization procedure described in Section 5.3. The number of unique firms in these three samples is 1,439, 187, and 154, respectively.

By contrast, while the absence of pre-policy differences near the cutoff does not validate a global DiD design, it is precisely what supports an identification strategy that explicitly conditions on the running variable at the cutoff. The DiDisc design exploits continuity at the threshold and identifies policy effects from changes in the discontinuity over time. As a result, the DiDisc design avoids the pitfalls inherent in averaging across heterogeneous and arbitrarily defined comparison groups.

Together, Appendices E.1 and E.2 demonstrate that the limitations of a standard DiD approach in this setting derive from the estimand it targets rather than from an assumption violation. When treatment is induced by a size-based threshold, the policy effect is inherently local because incentives change discontinuously at the cutoff. A global DiD estimator aggregates comparisons across firms that differ systematically in size and regulatory exposure; hence it cannot isolate this margin even when pre-policy dynamics are continuous near the threshold. In contrast, the DiDisc design isolates this policy-relevant margin, yielding a well-defined estimand aligned with the institutional structure of the policy. In turn, the DiD design does not recover the causal effect of interest.

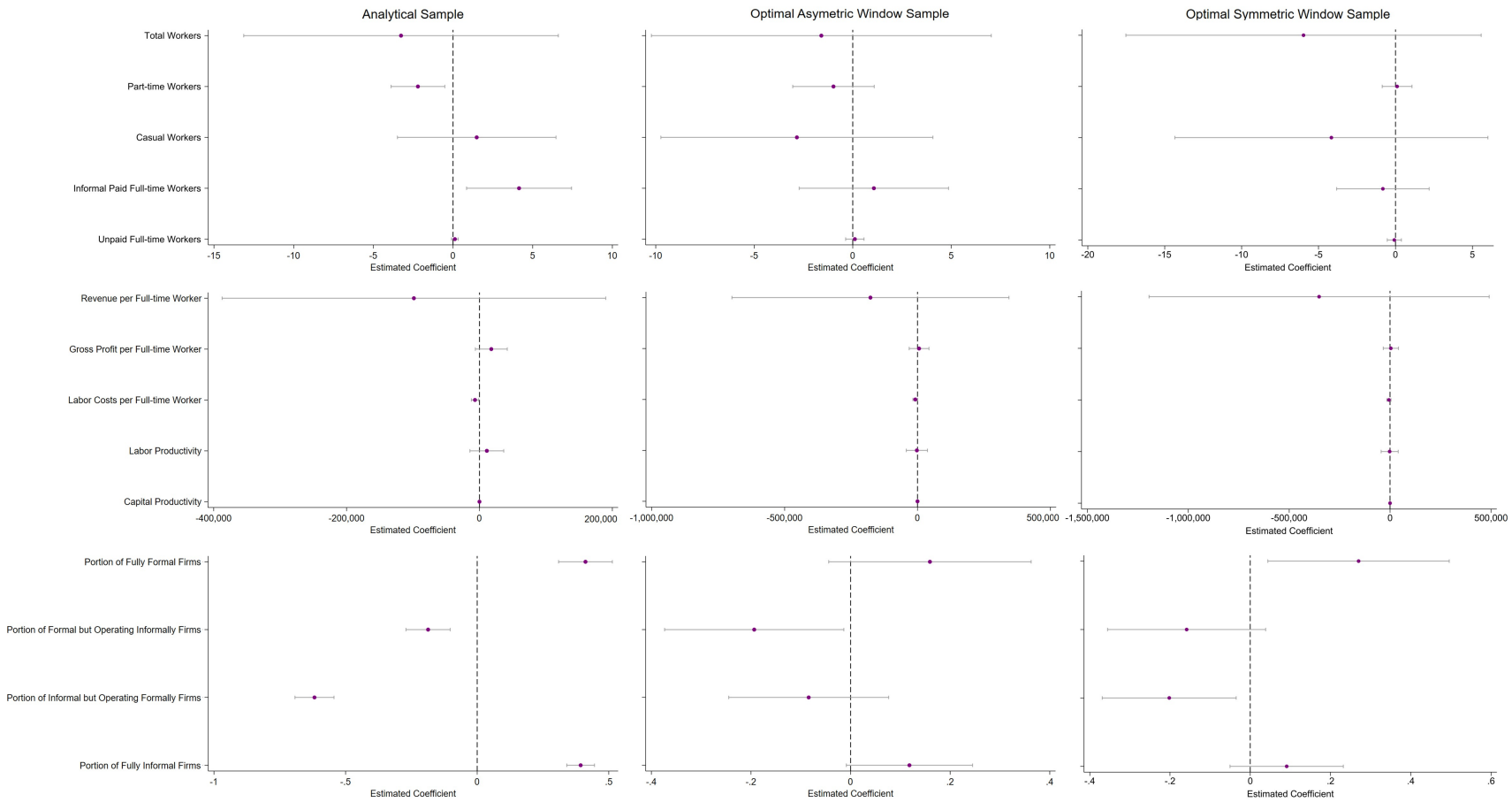


Figure E.2: Estimated Pre-Policy Discontinuities Across Samples

Notes: This figure plots estimated coefficient $\hat{\beta}_{2008}$ from Equation E.2 for all outcomes across three model specifications that correspond to the following samples: the analytical sample (see Appendix C), the optimal asymmetric window, and the optimal symmetric window (see Section 5.3). The coefficient $\hat{\beta}_{2008}$ is the estimated coefficient on $Above10_{i,2010} \times \mathbb{1}\{t = 2008\}$, where $Above10_{i,2010}$ indicates firms with at least 10 formal paid (FP) full-time workers in year-end 2010. Year-end 2010 is the omitted reference period. Whiskers denote 95 percent confidence intervals. Estimates are predominantly small and statistically insignificant, particularly in the optimal window samples. All size-normalized firm economic account metrics are reported in thousands of Vietnamese Dong.

E.3 Dynamic DiDisc Event-Study Evidence

The preceding diagnostics show that a standard DiD estimator is not the appropriate benchmark in this setting because the relevant policy variation is local to the 10-FP-employee threshold and estimates that do not condition on the running variable are sensitive to how far the sample extends from the cutoff. The remaining question is whether the DiDisc results reported in Section 6.2 reflect policy-induced changes at the threshold or instead capture discontinuities that were already present before the policy became binding. This distinction matters since the static DiDisc estimates shown in Table 4 imply economically meaningful firm adjustments at the threshold post-policy. Therefore I estimate a *dynamic* DiDisc event-study specification based on Equation 6 that preserves the optimal window comparisons and tests whether the relevant year-end 2014 discontinuities are systematically preceded by comparable effects in year-end 2008 or 2012.

Specifically, I estimate the DiDisc event-study specification in Equation E.3 for each outcome variable described in Section 5.2.

$$\begin{aligned} Y_{it} = & \sum_{k \in \{2008,2012,2014\}} \beta_k \left(\mathbb{1}\{t = k\} \times Above10_{i,2010} \right) \\ & + \sum_{k \in \{2008,2012,2014\}} \theta_k \left(\mathbb{1}\{t = k\} \times Center_{i,2010} \right) \\ & + \sum_{k \in \{2008,2012,2014\}} \rho_k \left(\mathbb{1}\{t = k\} \times Above10_{i,2010} \times Center_{i,2010} \right) + \mu_i + \lambda_t + \varepsilon_{it} \end{aligned} \tag{E.3}$$

In Equation E.3, Y_{it} is an outcome for firm i in year-end period $t \in \{2008, 2010, 2012, 2014\}$, with year-end 2010 omitted as the reference period. Thus the summations include $k \in \{2008, 2012, 2014\}$. $Above10_{i,2010}$ equals 1 if firm i has at least 10 FP full-time workers in year-end 2010 and 0 otherwise.[48] $Center_{i,2010}$ is the centered running variable, defined as firm i's year-end 2010 number of FP full-time workers minus 10. The terms μ_i and γ_t capture firm and year fixed effects, respectively, and ϵ_{it} is the error term. Standard errors are clustered at the firm level.

The coefficients of interest are β_{2008}, β_{2012}, and β_{2014}. For a given outcome variable, β_k captures the change in the discontinuity at the 10-FP employee threshold in year-end k relative to year-end 2010. Therefore, β_{2008} provides a pre-policy check, β_{2012} captures the interim period between the policy's passage and implementation (see Figure 1), and β_{2014} captures the post-policy change in the threshold discontinuity. The θ_k terms estimate year-specific changes in the slope of the outcome variable with respect to the running variable below the threshold while the ρ_k terms estimate how those slope changes differ above the threshold. Thus, this DiDisc event-study specification evaluates the timing of the local threshold effect while allowing slopes to evolve on both sides of the cutoff.

Tables E.1–E.3 report the dynamic DiDisc event-study estimates specified in Equation E.3 for all outcomes and both optimal windows. These estimates reinforce the static DiDisc results reported in Table 4: the key effects emerge in year-end 2014 and are not preceded by significant year-end 2008 or 2012 discontinuities. The only significant earlier estimate across the DiDisc event-study tables is

[48] Year-end 2010 is defined as the pre-policy period based on the analysis in Section 3.

the year-end 2012 decline in labor costs per full-time worker within the optimal asymmetric window. This timing supports interpreting the static DiDisc results as policy-induced adjustments at the threshold rather than pre-existing differences around the cutoff. Given the biennial data structure and limited pre- and post-policy periods (see Section 3), this analysis serves as a complementary robustness check; the static DiDisc estimates in Equation 6 remain the main identifying evidence.

Table E.1: DiDisc Event Study Results for Firm Size and Labor Composition

	Total Workers		Part-time Workers		Casual Workers		Informal Paid Full-time Workers		Unpaid Full-time Workers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
2008×	6.445	-3.229	0.944	1.636	3.281	-2.413	0.667	0.319	0.010	-0.057
$\text{Above10}_{i,2010}$	(5.388)	(5.069)	(0.792)	(1.036)	(4.268)	(3.892)	(2.289)	(2.184)	(0.518)	(0.396)
2012×	-1.943	-3.713	-0.926	0.265	1.137	-1.360	-0.196	0.341	-0.178	-0.273
$\text{Above10}_{i,2010}$	(2.627)	(2.423)	(0.741)	(0.975)	(2.605)	(2.165)	(2.106)	(2.014)	(0.221)	(0.206)
2014×	-10.581	-15.845	-0.785	1.179	-6.153	-12.276	-0.283	0.571	-0.376	-0.352*
$\text{Above10}_{i,2010}$	(11.553)	(11.610)	(0.989)	(1.103)	(10.170)	(10.041)	(1.281)	(1.331)	(0.252)	(0.210)
Control mean	9.125	9.125	0.412	0.412	0.875	0.875	1.263	1.263	0.762	0.762
R^2	0.581	0.593	0.347	0.289	0.438	0.428	0.424	0.357	0.635	0.576
Optimal window	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]
Firm-years	535	650	535	650	535	650	535	650	535	650

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. Reported coefficients correspond to $\hat{\beta}_{2008}$, $\hat{\beta}_{2012}$, and $\hat{\beta}_{2014}$ from Equation E.3. "Control mean" reports the year-end 2010 mean of the outcome variable for the control group (i.e., firms just below the threshold) within the corresponding optimal window. Year-end 2010 is the omitted reference period. The optimal windows are defined using the analytical sample discussed in Section 5.3: 154 and 187 firms within the optimal symmetric and asymmetric windows, respectively. The DiDisc event study preserves these windows and includes year-end 2008 and 2012 observations when available, so reported observations are firm-year observations from an unbalanced panel.

Table E.2: DiDisc Event Study Results for Firm Performance

	Sales Revenue per Full-time Worker		Gross Profit per Full-time Worker		Labor Costs per Full-time Worker		Labor Productivity		Capital Productivity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
2008×	-58,178	-278,695	12,629	18,989	1,565	-6	14,195	18,983	0.042	0.043
$\text{Above10}_{i,2010}$	(228,582)	(447,367)	(25,425)	(30,974)	(6,702)	(6,434)	(28,223)	(32,758)	(0.337)	(0.330)
2012×	-166,960	-259,420	16,685	7,354	-9,264	-13,452*	7,422	-6,098	0.459	0.514
$\text{Above10}_{i,2010}$	(249,773)	(267,998)	(30,196)	(38,228)	(7,615)	(8,032)	(33,510)	(42,456)	(0.496)	(0.495)
2014×	992,352	2,047,275	100,207*	139,326**	-11,457	-16,162	88,750	123,164*	0.147	0.451
$\text{Above10}_{i,2010}$	(792,939)	(1,777,923)	(57,064)	(67,474)	(14,719)	(14,676)	(59,452)	(69,322)	(0.466)	(0.519)
Control mean	298,647	298,647	55,148	55,148	31,533	31,533	86,681	86,681	0.560	0.560
R^2	0.263	0.260	0.313	0.324	0.466	0.484	0.331	0.342	0.365	0.379
Optimal window	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]
Firm-years	535	650	535	650	535	650	535	650	535	650

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. Reported coefficients correspond to $\hat{\beta}_{2008}$, $\hat{\beta}_{2012}$, and $\hat{\beta}_{2014}$ from Equation E.3. "Control mean" reports the year-end 2010 mean of the outcome variable for the control group (i.e., firms just below the threshold) within the corresponding optimal window. Year-end 2010 is the omitted reference period. The optimal windows are defined using the analytical sample discussed in Section 5.3: 154 and 187 firms within the optimal symmetric and asymmetric windows, respectively. The DiDisc event study preserves these windows and includes year-end 2008 and 2012 observations when available, so reported observations are firm-year observations from an unbalanced panel. Estimates are reported in thousands of Vietnamese Dong and reflect year-end values. Gross profit equals sales revenue minus total costs—including raw materials, indirect expenses, and labor costs (e.g., wages, allowances, recruitment, training, as well as health, social, and unemployment insurance). Labor productivity is defined as value added per full-time worker while capital productivity is defined as value added per asset. Value added is calculated as sales revenue minus production costs (including raw materials and indirect expenses).

Table E.3: DiDisc Event Study Results for Firm Formality Status

	Fully Formal Firm		Formal Firm But Operates Informally		Informal Firm But Operates Formally		Fully Informal Firm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2008×	-0.081	0.082	0.077	0.088	-0.057	-0.228	0.061	0.057
$\text{Above10}_{i,2010}$	(0.240)	(0.221)	(0.186)	(0.170)	(0.160)	(0.149)	(0.107)	(0.103)
2012×	-0.075	0.021	0.052	0.062	-0.004	-0.106	0.027	0.024
$\text{Above10}_{i,2010}$	(0.134)	(0.126)	(0.117)	(0.111)	(0.194)	(0.181)	(0.170)	(0.158)
2014×	-0.001	0.016	0.150*	0.182**	-0.131	-0.227	-0.019	0.030
$\text{Above10}_{i,2010}$	(0.147)	(0.138)	(0.089)	(0.088)	(0.182)	(0.162)	(0.156)	(0.141)
Control mean	0.175	0.175	0.038	0.038	0.587	0.587	0.200	0.200
R^2	0.520	0.553	0.470	0.451	0.546	0.549	0.437	0.446
Optimal window	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]
Firm-years	535	650	535	650	535	650	535	650

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. Reported coefficients correspond to $\hat{\beta}_{2008}$, $\hat{\beta}_{2012}$, and $\hat{\beta}_{2014}$ from Equation E.3. "Control mean" reports the year-end 2010 mean of the outcome variable for the control group (i.e., firms just below the threshold) within the corresponding optimal window. Year-end 2010 is the omitted reference period. The optimal windows are defined using the analytical sample discussed in Section 5.3: 154 and 187 firms within the optimal symmetric and asymmetric windows, respectively. The DiDisc event study preserves these windows and includes year-end 2008 and 2012 observations when available, so reported observations are firm-year observations from an unbalanced panel. Firm formality status definitions are in Table 2.

F Placebo Threshold Tests at Alternative Cutoffs

This appendix presents placebo threshold tests assessing whether the difference-in-discontinuities (DiDisc) results in Section 6.2 are driven by the 10-formal paid (FP) employee policy threshold or instead reflect broader size-gradient differences across firms. The identifying variation in the main analysis comes from the discontinuous increase in the financial penalty for non-compliance at 10 FP employees. If the estimated effects were simply capturing smooth or non-policy-related differences between firms with higher and lower FP employment, then similar discontinuities should appear at alternative firm-size cutoffs where no policy change occurs. By contrast, if the effects are specific to the policy threshold, placebo DiDisc estimates at non-policy cutoffs should be imprecise, unstable, or unavailable under the same local randomization procedure described in Section 5.3.

To implement this falsification test, I repeat optimal window selection and DiDisc estimation at the following placebo FP full-time worker cutoffs: $c \in \{15, 20, 25, 30\}$. For each placebo cutoff, I re-estimate Equation 6 after redefining the running and treatment variables relative to the updated placebo cutoff. Specifically, $Center_{i,2010}$ in Equation 6 is replaced by the following cutoff-specific centered running variable:

$$Center_{ic} = FPEmp_{i,2010} - c$$

where $FPEmp_{i,2010}$ denotes the number of FP full-time workers employed by firm i in the pre-policy period (i.e., year-end 2010). Similarly, $Above10_{i,2010}$ is replaced by the following:

$$AboveC_{ic} = \mathbb{1}\{FPEmp_{i,2010} \geq c\}$$

which equals 1 if firm i employs FP full-time workers at or above the placebo cutoff c and 0 otherwise. The interaction $Above10_{i,2010} \times Post_t$ in Equation 6 is thus replaced by $AboveC_{ic} \times Post_t$ and its coefficient δ_0^c captures the DiDisc estimate at placebo cutoff c. All other elements of the specification remain unchanged, including the firm fixed effects, the post-policy binary variable $Post_t$, the local linear controls in the centered running variable, and the corresponding interaction terms.

To separate the placebo tests from the true policy margin, I restrict candidate placebo windows to firms with at least 11 FP full-time workers pre-policy. Doing so prevents placebo windows from including the actual 10-FP employee threshold, where the financial penalty changes. Equivalently, because the running variable is centered at each placebo cutoff c, this lower bound becomes $11 - c$ in centered units, ensuring all placebo comparisons remain strictly *above* the policy-relevant threshold.

For each placebo cutoff, I carry out the same local randomization procedure described in Section 5.3. Candidate placebo windows are evaluated using the pre-policy covariates from Table 3 and are admissible when the minimum covariate balance test p-value across all covariates is 0.15 or higher (Cattaneo et al., 2024). The optimal symmetric and asymmetric placebo windows are then selected using the modified greedy expansion procedure discussed in Section 5.3.

To preserve the local nature of the placebo comparisons, candidate *asymmetric* expansions are capped at the maximum absolute bandwidth selected in Section 5.3, where the optimal asymmetric window of $[1, 25]$ corresponds to a left-side bandwidth of 9 and a right-side bandwidth of 15. Thus I

restrict placebo candidate expansions to a maximum absolute bandwidth of 15 around each placebo cutoff. This ensures that the placebo estimates are based on comparable local neighborhoods which limits the influence of broader size-gradient differences unrelated to the placebo threshold.

Figure F.1 presents DiDisc estimates for the outcome variables described in Section 5.2 at each placebo cutoff using cutoff-specific optimal symmetric and asymmetric windows. Estimates are not shown for the placebo 20-FP employee threshold as the local randomization procedure identifies no admissible symmetric window satisfying covariate balance. Thus this lower cutoff does not support locally valid placebo comparisons. The procedure identifies admissible symmetric windows for the placebo 15-, 25-, and 30-FP employee thresholds. The corresponding optimal symmetric windows are $[13, 16]$, $[22, 27]$ and $[22, 37]$, while the optimal asymmetric windows are $[11, 22]$, $[18, 40]$ and $[21, 45]$, respectively. Across these placebo cutoffs, the estimates vary in magnitude, precision, and sign and do not reproduce the pattern observed at the *true* 10-FP employee threshold in Table 4.

Together, these results support the interpretation that the DiDisc estimates in Section 6.2 are not mechanical artifacts of comparing firms at arbitrary points in the FP employment distribution. Instead, the effects are concentrated around the *actual* 10-FP employee threshold, where the financial penalty for non-compliance increases. At these placebo cutoffs, either the local randomization balance criterion fails or the estimates are noisy and unstable, providing no systematic evidence of comparable threshold effects away from the policy cutoff. These falsification tests complement the dynamic DiDisc event-study in Appendix E.3 which finds that the post-policy threshold effects are not generally preceded by comparable pre-policy discontinuities.

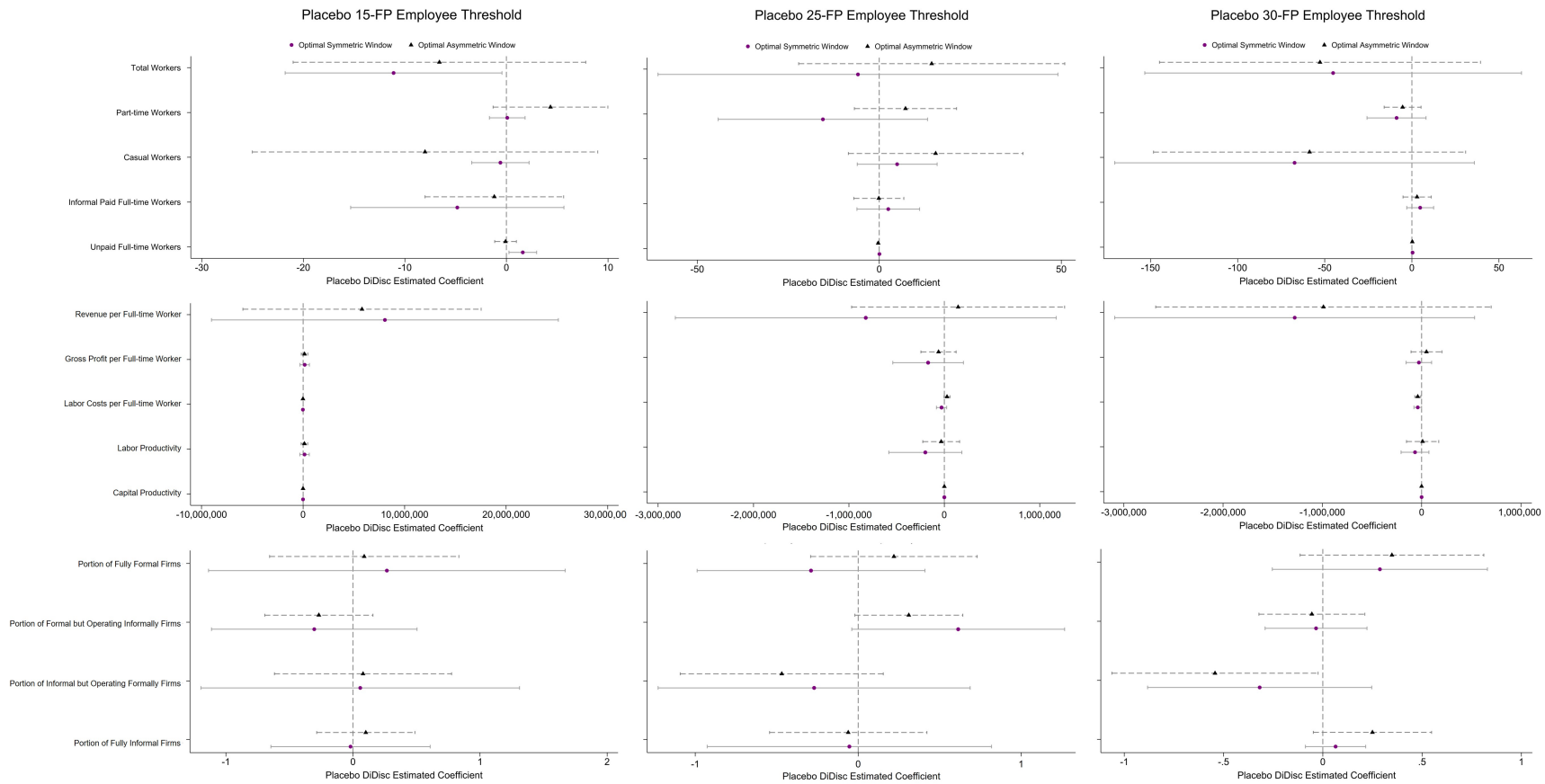


Figure F.1: Placebo DiDisc Estimates at Non-Policy Thresholds

Notes: This figure presents placebo DiDisc estimates from Equation 6 after replacing $Center_{i,2010}$ and $Above10_{i,2010}$ with the cutoff-specific variables $Center_{ic} = FPEmp_{i,2010} - c$ and $AboveC_{ic} = \mathbb{1}\{FPEmp_{i,2010} \geq c\}$, respectively. Estimates are shown for placebo cutoffs $c \in \{15, 25, 30\}$; however, estimates for $c = 20$ are omitted since no admissible symmetric window satisfies the minimum covariate balance criterion. Circles and triangles report placebo estimates using cutoff-specific optimal symmetric and asymmetric windows, respectively. Whiskers denote 95 percent confidence intervals. Overall, these estimates lack precision and consistency across cutoffs and optimal windows, reinforcing that comparable threshold effects are not evident *away* from the actual 10-FP employee threshold.

G Heterogeneity by Pre-Policy Firm Formality Status

The average causal effects in Section 6.2 potentially mask important differences in how firms with distinct *pre-policy* formality statuses responded to the higher financial penalty for non-compliance under the 2012 employee threshold policy. To analyze this heterogeneity, this appendix focuses on three firm subgroups with observed support on both sides of the cutoff within the optimal windows identified in Section 5.3: fully formal firms, informal firms that operate formally, and fully informal firms. Prior to policy implementation, fully formal firms were registered with the government and provided all full-time workers with formal, written labor contracts; informal firms operating formally provided such contracts but were *not* government-registered; and fully informal firms were neither registration nor labor contract compliant. Given that each subgroup occupies a different point on the formality spectrum, their incentives under the policy change differ. Examining these subgroups separately evaluates if the average effects identified earlier were driven by fully-compliant, partially-compliant, or non-compliant firms. This helps clarify the mechanisms underlying the policy's impact and identifies which firm *types* are most responsive to such threshold-based labor policies.

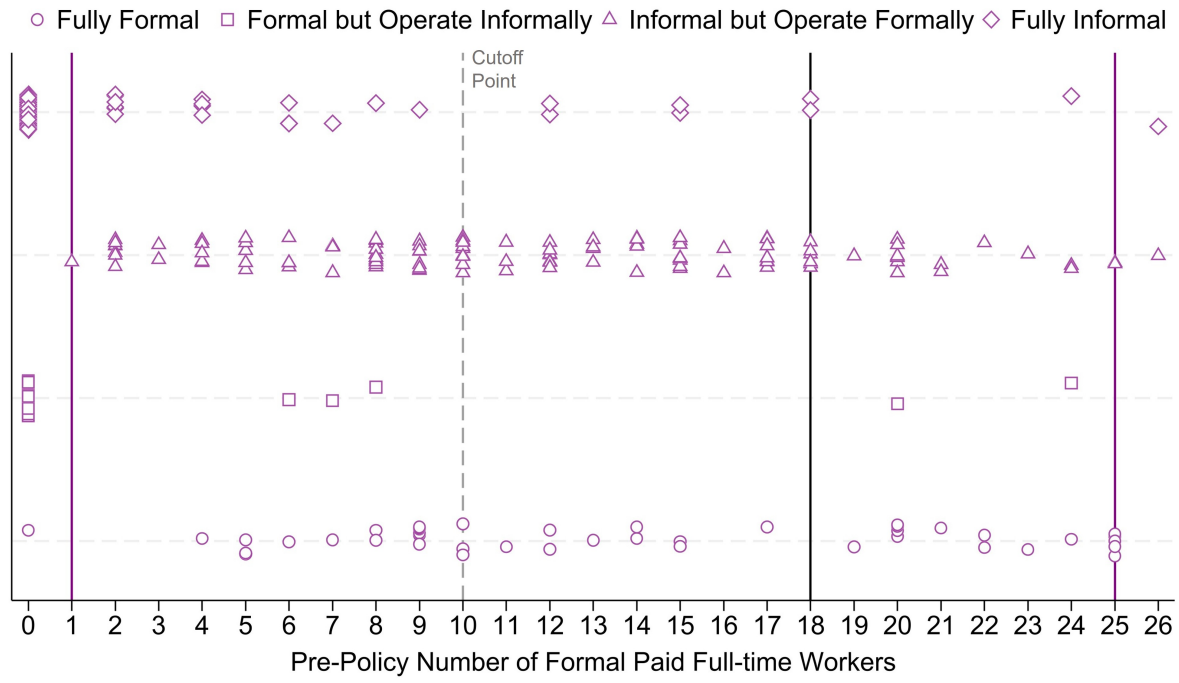


Figure G.1: Distribution of Firms within Optimal Windows by Pre-Policy Formality Status

Notes: This figure shows the distribution of firms within the optimal symmetric and asymmetric windows identified in Section 5.3 by their pre-policy formality status. Each formality status is represented by a different marker shape. The black and purple vertical lines denote the boundaries of the optimal symmetric and asymmetric windows, respectively.

Figure G.1 shows the distribution of firms around the cutoff for the analytical sample by their

pre-policy formality status within the optimal symmetric and asymmetric windows.[49] Firms that were already fully formal, informal but operating formally, or fully informal have observed support on both sides of the 10-FP employee threshold, permitting subgroup DiDisc estimation.

Table G.1 reports estimates from Equation 6 by pre-policy formality status while holding fixed the optimal symmetric and asymmetric windows. Among fully formal firms, Columns 1-2 of Panel A indicate that firms just *above* the 10-FP employee threshold had between 1.1 to 1.7 fewer unpaid full-time workers relative to firms just *below* it post-policy. Since the optimal symmetric window estimate fails the covariate balance criterion, the causal interpretation rests primarily on the balanced optimal asymmetric window estimate of 1.1 fewer unpaid full-time workers. This decline suggests that fully formal firms just *below* the threshold—seeking to avoid threshold-triggered compliance costs—substituted toward unpaid labor arrangements, consistent with Theorem 2. Column 4 shows that fully formal firms just above the threshold were 33 percentage points more likely to be "formal but operating informally" than firms below it post-policy, as predicted by Theorem 3. Hence some firms strategically relaxed compliance along one dimension by reducing their formal labor contract coverage. Altogether, these findings suggest that in response to the 2012 employee threshold policy, fully formal firms became less reliant on unpaid labor and reshaped their compliance margins by remaining government-registered while selectively adjusting their broader labor strategies.

As a robustness check, I re-select the optimal symmetric and asymmetric windows separately for each pre-policy formality status subgroup and then re-estimate Equation 6. Table G.2 reinforces the finding that the labor composition effects are concentrated among firms that were fully formal pre-policy. However, the earlier firm formality status effect does not persist. Moreover, fully formal firms just above the threshold experienced increases in gross profit per full-time worker of 91 to 117 million VND (about 3,500 to 4,500 USD) and in labor productivity of 119 to 128 million VND (about 4,500 to 4,900 USD). Fully informal firms exhibit similarly positive performance estimates, but these are significant only under their subgroup-specific optimal asymmetric window. Overall, initially fully formal firms show the strongest and most consistent labor composition and performance responses, with performance gains indicating improved efficiency rather than distortion.

These results suggest that the increase in the financial penalty introduced by the 2012 employee threshold policy did not produce a uniform shift toward full firm formality. Instead, firms' responses varied with their pre-policy formality status. The concentration of effects among fully formal firms suggests that the policy change operated predominantly through organizational adjustment among already-compliant firms rather than broad-based formalization. This heterogeneity highlights how firms' initial formality status can shape their responses to threshold-based labor policies, a nuance obscured by average treatment effects.

[49] Appendix C describes the construction of the analytical sample from the Vietnam SME dataset.

Table G.1: Impact of Threshold-Based Labor Policy on Selected Firm Outcomes by Pre-Policy Formality Status within Optimal Windows

	Unpaid Full-time Workers		Formal Firm But Operates Informally		Gross Profit Per Full-time Worker		Labor Productivity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A) Fully Formal Firms								
$Post_t \times$	-1.666**[a]	-1.143*	0.017[a]	0.333**	45,506[a]	78,014	66,530[a]	79,218
$Above10_{i,2010}$	(0.691)	(0.566)	(0.097)	(0.162)	(56,019)	(72,612)	(57,561)	(73,986)
Control mean	0.214	0.214	0.000	0.000	101,043	101,043	139,814	139,814
R^2	0.623	0.601	0.736	0.574	0.589	0.532	0.578	0.537
Firm-years	52	82	52	82	52	82	52	82
(B) Informal Firms Operating Formally								
$Post_t \times$	-0.168[a]	-0.249[a]	0.119[a]	0.111[a]	24,869[a]	25,072[a]	1,639[a]	-1,333[a]
$Above10_{i,2010}$	(0.250)	(0.232)	(0.106)	(0.103)	(34,803)	(30,599)	(43,197)	(39,503)
Control mean	0.830	0.830	0.000	0.000	49,987	49,987	79,557	79,557
R^2	0.768	0.775	0.632	0.626	0.550	0.560	0.549	0.559
Firm-years	206	236	206	236	206	236	206	236

Continued on next page

Continued from previous page

	Unpaid Full-time Workers		Formal Firm But Operates Informally		Gross Profit Per Full-time Worker		Labor Productivity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(C) Fully Informal Firms								
$Post_t \times$	0.649	0.649	-0.179	-0.179	1,522,760	1,395,027	1,511,157	1,384,173
$Above10_{i,2010}$	(0.425)	(0.424)	(0.258)	(0.257)	(952,710)	(827,639)	(960,961)	(832,280)
Control mean	1.188	1.188	0.000	0.000	35,925	35,925	67,374	67,374
R^2	0.858	0.861	0.564	0.565	0.663	0.662	0.666	0.665
Firm-years	44	46	44	46	44	46	44	46
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Optimal window	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]	[1,18]	[1,25]

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. "Control mean" reports the pre-policy mean of the outcome variable for the control group (i.e., firms just below the threshold) within the corresponding optimal window for each subgroup. "Firm-years" reports firm-period observations, with each firm contributing one year-end 2010 and one year-end 2014 observation. Fully formal firms are registered with the government and provide formal, written labor contracts to all full-time workers. Informal firms that operate formally are also labor contract compliant but are *not* registered. Fully informal firms are not registration nor labor contract compliant. Results come from subgroup analyses by firms' pre-policy formality status *within* the optimal symmetric and asymmetric windows defined in Section 5.3. The p-values for the estimates in Columns 5, 6, 7, and 8 of Panel C are 0.125, 0.106, 0.131, and 0.110, respectively. Estimates are generally robust to subgroup-specific optimal symmetric and asymmetric windows selected separately by pre-policy firm formality status (see Table G.2). Estimates marked with [a] correspond to subgroup-optimal window combinations that fail the pre-policy covariate balance criterion. In Panel A, imbalance within the optimal symmetric window is driven by higher assets per full-time worker among firms just above the threshold ($p = 0.094$). However, in Panel B, imbalance is driven by higher upper secondary school completion among firm owners or managers just above the threshold within both the optimal symmetric ($p = 0.088$) and asymmetric ($p = 0.054$) windows. This table only reports selected firm outcomes from Table 4.

Table G.2: Impact of Threshold-Based Labor Policy on Selected Firm Outcomes by Pre-Policy Formality Status: Subgroup-Specific Optimal Windows

	Unpaid Full-time Workers		Formal Firm But Operates Informally		Gross Profit Per Full-time Worker		Labor Productivity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A) Fully Formal Firms								
$Post_t \times$	-2.015**	-1.665**	—[b]	-0.035	117,359**	91,496**	127,960**	119,104**
$Above10_{i,2010}$	(0.879)	(0.705)	—	(0.087)	(55,011)	(41,790)	(55,103)	(45,366)
Control mean	0.200	0.200	—	0.000	71,352	94,307	110,891	130,493
R^2	0.659	0.626	—	0.667	0.720	0.572	0.779	0.578
Optimal window	[6,13]	[0,16]	[6,13]	[0,16]	[6,13]	[0,16]	[6,13]	[0,16]
Firm-years	34	52	34	52	34	52	34	52
(B) Informal Firms Operating Formally								
$Post_t \times$	0.088	0.144	0.067	0.046	32,297	25,460	2,322	-8,931
$Above10_{i,2010}$	(0.244)	(0.211)	(0.116)	(0.102)	(33,031)	(30,111)	(44,848)	(42,385)
Control mean	0.553	0.553	0.000	0.000	54,303	54,303	86,280	86,280
R^2	0.763	0.779	0.562	0.548	0.563	0.587	0.568	0.577
Optimal window	[3,16]	[3,30]	[3,16]	[3,30]	[3,16]	[3,30]	[3,16]	[3,30]
Firm-years	158	234	158	234	158	234	158	234

Continued on next page

Continued from previous page

	Unpaid Full-time Workers		Formal Firm But Operates Informally		Gross Profit Per Full-time Worker		Labor Productivity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(C) Fully Informal Firms								
$Post_t \times$	0.542	0.542	-0.167	-0.167	1,510,042	1,499,468*	1,495,894	1,491,522*
$Above10_{i,2010}$	(0.616)	(0.610)	(0.265)	(0.262)	(876,690)	(837,248)	(881,549)	(841,599)
Control mean	1.000	1.000	0.000	0.000	16,958	16,958	46,486	46,486
R^2	0.976	0.976	0.567	0.568	0.654	0.657	0.658	0.660
Optimal window	[6,24]	[5,30]	[6,24]	[5,30]	[6,24]	[5,30]	[6,24]	[5,30]
Firm-years	26	28	26	28	26	28	26	28
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the firm level. "Control mean" reports the pre-policy mean of the outcome variable for the control group (i.e., firms just below the threshold) within the corresponding optimal window for each subgroup. "Firm-years" reports firm-period observations, with each firm contributing one year-end 2010 and one year-end 2014 observation. Fully formal firms are registered with the government and provide formal, written labor contracts to all full-time workers. Informal firms that operate formally are also labor contract compliant but are *not* registered. Fully informal firms are not registration nor labor contract compliant. Results use subgroup-specific optimal symmetric and asymmetric windows for *each* pre-policy firm formality status. The p-values for the estimates in Columns 5 and 7 of Panel C are 0.111 and 0.115, respectively. Cells marked with [b] are omitted because the data do not support reliable clustered standard errors or statistical tests after accounting for firm fixed effects. This table only reports selected firm outcomes from Table 4. To preserve the local nature of the analysis, the modified greedy expansion procedure caps the right-side bandwidth at 20, limiting the optimal asymmetric window to 30 FP full-time workers; accordingly, the selected optimal asymmetric windows of $[3, 30]$ in Panel B and $[5, 30]$ in Panel C reach this pre-specified upper endpoint. For fully informal firms, the optimal window of $[6, 24]$ is symmetric in the number of observed running variable support points rather than in numerical distance because this subgroup has gaps in its support.