Employment Protection Legislation and Informality: Theory and Evidence from India*

Ritam Chaurey Gaurav Chiplunkar Vidhya Soundararajan

September 2024

Abstract

We present causal evidence on how employment protection legislation (EPL) that penalize firms for hiring contractual workers can have adverse effects in contexts characterized by high levels of informality. While they increase compliance and reduce informality on the "intensive margin" as formal sector firms reduce hiring contract workers "off-the-books", they increase informality on the "extensive margin", as firms move to operating in the informal sector instead. This lowers aggregate productivity, wages and welfare in equilibrium. Counterfactual simulations demonstrate the significance of carefully designing EPL in contexts where informality is prevalent.

JEL Classification: D22, E26, J46, O12, O17

Keywords: Contract labor, Employment Protection Legislation, Firms, Informality.

*Chaurey: Johns Hopkins School of Advanced International Studies; rchaurey@jhu.edu. Chiplunkar *(corresponding author)*: Darden School of Business, University of Virginia; chiplunkarg@darden.virginia.edu. Soundararajan: Indian Institute of Technology, Mumbai; vidhyasrajan@iitb.ac.in. We thank Shilpa Aggarwal, Samuel Bazzi, Nancy Chau, Penny Goldberg, Tarun Jain, Gaurav Khanna, David Lagakos, Amartya Lahiri, Devashish Mitra, Mushfiq Mobarak, Dilip Mookherjee, Kaivan Munshi, Mark Rosenzweig, Todd Schoellman and the participants in seminars and conferences at Ashoka University, Azim Premji University, Indian School of Business, Indian Institute of Management (Bangalore), NEUDC, PacDev at University of California at San Diego, Yale University, University of Virginia, and the 1^{st} Bank of Italy-World Bank conference on 21^{st} century jobs. This paper was previously circulated as "Implementation of Labor Laws and Informality". We acknowledge the excellent research assistance by Prerana Hiriyur and Jayshrita Bagabati. All remaining errors are our own.

1 Introduction

Policy discussions surrounding employment protection legislations (EPL) have gained renewed attention worldwide in the aftermath of the Covid-19 pandemic. While there exists a substantial body of literature evaluating the impact of EPL,[1] a key distinction in understanding their effects in high- and low-income countries is a sizeable presence of the informal sector in the latter (La Porta and Shleifer, 2008, 2014).[2] Furthermore, as highlighted by Ulyssea (2018), the nature of informality itself can be characterized along two distinct margins in these countries, namely: (i) the "extensive margin", where firms must decide whether or not to formalize their operations by registering, paying taxes, and complying with EPL;[3] and (ii) the "intensive margin," where formal sector firms instead opt to hire workers "off-the-books" on informal contracts to evade EPL.

Previous empirical studies examining the causal impact of EPL in developing countries have typically focused on only one of these two margins, primarily due to the challenge of measuring both in large-scale administrative data that are commonly available in these countries.[4] However, as we show in this paper, analyzing the impact of EPL on *both* these margins of informality yields additional valuable insights, particularly regarding the design of these policies and their unintended consequences, which can potentially adversely affect wages, aggregate productivity and welfare in equilibrium.

For example, in the context of our study, a large-scale EPL reform imposed penalties on formal sector firms for hiring workers on temporary, short-term contracts instead of on regular payroll. Firms often used these contractual arrangements to circumvent labor laws pertaining to minimum wages, social security and severance payments, among others. Therefore, the policy reform directly aimed to reduce the use of contract workers in the formal sector i.e., reduce informality on the intensive margin (which it was successful in achieving). However, it also made firms less inclined to formalize their businesses, leading to an overall expansion of the informal sector on the extensive margin. Consequently, this decreased wages, reduced aggregate productivity, and thus led to an overall decline in welfare. Through counterfactual simulations, we further illustrate

[1] See for example, Besley and Burgess (2004); Autor et al. (2006); Bachas et al. (2019); Bjuggren (2018); Boeri and Garibaldi (2007); Bornhäll et al. (2015); Butschek and Sauermann (2022); Cahuc and Postel-Vinay (2002); Cahuc et al. (2016, 2023); Chaurey (2015); Schivardi and Torrini (2008); Ulyssea (2018), etc.

[2] There is a strong negative correlation between real GDP per capita and the fraction of employment and output originating from the informal sector. While more than 80% of workers and 30-40% of output in low-income countries comes from the informal sector, it is only 30% of workers and 10% of output in high-income ones (Elgin et al., 2021; La Porta and Shleifer, 2014).

[3] As documented in a comprehensive review by La Porta and Shleifer (2014), limited state capacity to monitor, poor quality of public institutions (such as corruption), and cumbersome regulations are the primary reasons for firms choosing to operate in the informal sector.

[4] For example, see Almeida et al. (2022); Besley and Burgess (2004); Bruhn (2013); Bruhn and McKenzie (2014) on the extensive margin; and Chaurey (2015); Samaniego de la Parra and Fernández Bujanda (2020); Bertrand et al. (2024) on the intensive margin.

how alternative EPL policies have the potential to effectively reduce informality, aligning with the policymakers' goals, while simultaneously yielding meaningful enhancements in welfare. Our study therefore not only addresses a significant gap in the literature by providing causal evidence on the impact of EPL reforms on different dimensions of informality, but also underscores the importance of distinguishing between these dimensions in the effective design of these policies.

Our empirical investigation focuses on India, a country characterized by a large informal sector that employs around 85% of its workforce, with poor enforcement of EPL.[5] Is India's experience distinct from other countries globally? That does not seem to be the case. For example, using data on the fraction of workers in the informal sector (Elgin et al., 2021), and the quality of enforcement of EPLs (Botero et al., 2004), Figure A1 finds that informality declines (Figure A1a) and EPL are better enforced (Figure A1b) with economic development across countries. In either case, we see that India does not appear to be an exception in anyway. This suggests that the policy issues that we discuss in the paper are perhaps more broadly useful for low- and middle-income countries.

Historically, India has grappled with intricate labor laws and bureaucratic inefficiencies, making it challenging for formal sector firms to engage workers through proper payroll channels and adhere to EPL. Consequently in the 1990s, amidst trade liberalization, industrial reforms, and robust economic growth, these firms increasingly resorted to hiring workers "off-the-books" on temporary contracts, as a means to evade compliance with EPL (Bertrand et al., 2024; Chaurey, 2015; Chakraborty et al., 2020). However, this implied that these contract workers were denied job and social security benefits, and often meted out with poor working conditions. This generated an active discussion on how policy could intervene to limit the usage of contract workers and encourage firms to hire more workers on payroll. There were two main viewpoints: one that advocated for active penalties on employing contract workers (Narayanan et al., 2005). The other advocated for a simplification of complex and cumbersome employment protection laws in India to ease the process of hiring workers on payroll.[6]

It was in the above context that the Government of Andhra Pradesh issued a notification, formally termed as *"Contract Labour (Regulation and Abolition) (Andhra Pradesh Amendment) Act, 2003"*. The policy reform prohibited the use of contract workers in *core* activities of the firm, either directly or through licensed contractors.[7] In case of any

[5]Using data from the 1998 Round of the Economic Census of India, 85% of India's labor force was employed in unregistered, informal firms. This ranged from 75% in the manufacturing sector, to around 90-95% in agriculture and services.

[6]See Ahsan and Pagés (2009); Sodhi (2014); Chaurey (2015).

[7]Core activities include the primary activity for which an establishment was set up, along with other activities essential to this primary activity. All other activities (such as catering, housekeeping, transport, etc.) were classified as "non-core" activities, and the use of contract workers was permitted in them. See Appendix B for a complete list.

violations, the penalties under the Contract Labour Act 1970 would apply (for e.g. substantial fines, potential imprisonment of Directors, and/or prosecution in court). There were three exceptions however, where employing contract workers was permitted: (i) if the work was "normally" done through contractors; or (ii) the activities themselves did not require full-time workers; or (iii) if there was a sudden increase in the volume of core activity work that needed to be completed within a certain time. The reform was therefore aimed at protecting workers by limiting the usage of contract workers and encouraging firms to hire them on regular payroll, similar to the first among the two viewpoints discussed previously.

Our empirical analysis draws on multiple nationally representative datasets of firms and workers, allowing us to comprehensively measure the impact of this policy on firms' adjustments along the margins of informality, as well as on workers' wages, type of contract, sector of employment, etc. To ensure consistent availability of different variables of interest across these datasets, our primary analysis focuses on the period between 1999-2006. We employ a standard difference-in-differences approach to compare outcomes in Andhra Pradesh (treated state) before and after the policy change (in 2003), as compared to other Indian states (control group).[8]

We begin by first examining the direct impact of the policy reform on the intensive margin of informality using a representative establishment-level panel data of formal sector manufacturing firms from the Annual Survey of Industries (ASI).[9] We find a 46-50% decline in the number of contract workers (worker-days), and a 18% drop in the fraction of contract workers. We also find a corresponding increase in payroll workers by 20-25%, with little change in overall firm size.[10] We undertake several additional analyses to examine the robustness of our results. For example, we show that the results are entirely driven by a reduction of workers in the core (as opposed to non-core) activities of a firm. The results are also robust to alternate control groups, insensitive to the set of fixed effects used, and persistent when we extend the sample period until 2009-10.

To examine the extensive margin of informality, we complement our analysis with various rounds of the National Sample Survey (NSS) and the Economic Census (EC). Unlike the ASI, which samples firms only in the formal manufacturing sector, the EC is a census of all non-agricultural firms in India, and the NSS is a nationally representative survey of workers. Put together, they provide a comprehensive picture of workers and

[8]We undertake extensive additional analyses to show the robustness of our results to alternate specifications, time frames, samples, and definitions.

[9]From the 1998 Economic Census of India, a quarter of India's labor force worked in the manufacturing sector around this time, with around 25% of manufacturing workers being employed in its formal sector.

[10]Note that finding no change in the overall firm size does not necessarily imply perfect substitution between contract and payroll workers. As we document later, the policy lowered wages on average as well, thus reducing firms' cost of hiring workers more generally.

firms in both the formal and informal sectors that we use to study the extensive margin of informality. Using multiple measures of informality in the NSS data, we find that workers in Andhra Pradesh (relative to other states) were 14.1% more likely to work on an informal contract, 9.2% more likely to work in an informal enterprise, 3.7% more likely to work as a casual (informal) worker, and reported a 4.6% decline in daily wage earnings after the policy reform.[11] We show that these results are robust to a number of extensions such as excluding self-employed individuals, restricting the sample to the manufacturing sector only (to make it more comparable to the ASI), as well as extending the time frame to all years where data is available.[12]

Turning to the Economic Census (EC), which is a census of all non-agricultural firms in India, we examine the impact of the policy on the probability that firms formalize. Specifically, we can only use the 1998 and 2005 rounds of the Economic Census since these are the only rounds where the formal registration status of the firm is available.[13] We find robustly that firms, and especially the larger ones, were more likely to remain unregistered in Andhra Pradesh (relative to other states) after the policy reform. For example, we find a 33-40% increase in the probability that a firm employing more than either 10 workers or 20 workers chose to remain unregistered. Larger firms, employing more than 50 and 100 workers (like the ones covered in the ASI for example), were even more likely (around 55%) likely to remain unregistered.

Finally, we use the ASI to measure (with some important caveats discussed in Section 5.2) entry and exit from the formal manufacturing sector. To do this, we collate annual reports published by the Ministry of Statistics and Programme Implementation (Government of India) that record (among other things) the total and "operational" number of establishments within a state-3-digit industry category. Unlike the panel data used in our earlier analysis, these are measured using repeated annual cross-sections of data and thus can measure the margin of entry and exit. Using these data, we find a substantial increase (decrease) in the fraction of inactive (new entrants) firms in the formal manufacturing sector, complementing the evidence from the EC.

To summarize, our empirical analysis puts together several data sets and comprehensively measures the impact of the policy. While these data vary in their scope and coverage, they robustly confirm a reduction (increase) in the usage of contract (payroll) workers by firms in the formal sector, an increase of firms and workers in the informal

[11] Section 4.2 explains the construction of these variables in detail. A complete list of variables across all datasets is provided in Table A1.

[12] We show that log daily wages were statistically similar in Andhra Pradesh and other states before the policy reform, with a substantial decline in Andhra Pradesh (relative to other states) only after it. The probability of workers working in casual work however, was decreasing over time in Andhra Pradesh (relative to other states) before the policy. However, the gap widens (increases) after the policy reform.

[13] This limits our ability to check for pre-policy trends in these variables and therefore, our results in the EC should be reported with appropriate caution.

sector, with workers earning lower wages on average.

How did the policy reform impact aggregate productivity and welfare in general equilibrium? Furthermore, could an alternative policy design have mitigated these adverse effects on informality and potentially improved welfare? These critical policy-relevant questions extend beyond the scope of the reduced-form analysis discussed above. To address them, we build a theoretical framework that closely follows the approach developed by Ulyssea (2018). It integrates the extensive and intensive margins of informality into a general equilibrium model with heterogeneous firms.

In the theoretical framework, firms are ex-ante heterogeneous in productivity, pay a fixed cost of entry to the market, and decide whether to enter the informal or formal sectors (the extensive margin of informality) or not enter at all. Conditional on formalizing, firms can endogenously choose to comply with EPL and hire workers on payroll or evade them by hiring workers off-the-books (the intensive margin of informality). In the informal sector, firms can hire workers without complying with taxes and EPL but face a size-dependent penalty for remaining informal.[14] The relationship between firm productivity, firm size, and the choice to operate in the formal or informal sector is inherently intertwined. Typically, more productive, larger firms opt to operate in the formal sector, while smaller, less productive ones persist in the informal sector. The model therefore provides a parsimonious way of rationalizing the implications of the aforementioned policy reform. We show for example that such policies would reduce informality on the intensive margin (as formal firms hire fewer workers off-the-books), but conversely, increase it on the extensive margin due to the increased costs of formal sector operations, particularly for larger firms. This, in turn, reduces aggregate labor demand, real wages, and overall welfare in equilibrium.

We then quantitatively estimate the model using a Simulated Method of Moments estimator and consider the implications of two counterfactual EPL policies discussed previously, both aimed at increasing compliance with EPL. The first one — similar to the Andhra Pradesh policy reform — implements penalties for employing contract workers, while the other advocates for simplifying cumbersome EPLs to ease the process of hiring workers on payroll.

To study the implications of the former, we impose a 19% *penalty* on the marginal cost of hiring a contract worker in the formal sector. This allows us to (by construction) mimic the impact of the policy reform in Andhra Pradesh by reducing the fraction of contract workers in a formal sector firm by 18%. As expected, it increases informality on the extensive margin as well, since it is costlier for firms to operate in the formal sector. However, there is an *indirect* channel that operates through general equilibrium.

[14]This penalty captures the probability of being audited and caught and more general constraints such as access to formal sources of finance and markets (La Porta and Shleifer, 2014).

Since firms are less productive in the informal sector, there is an overall decline in labor demand, which reduces real wages (by around 5%).[15,16] Lower real wages in equilibrium however, make it cheaper for all firms to hire workers, therefore offsetting some of the direct effects of the policy. Put together, aggregate productivity (TFP) decreases by 2% and real income, our measure of welfare, declines by 9%.

The second counterfactual, through the lens of our model, would translate into a lower relative marginal cost of hiring payroll workers. To make the effects comparable across the two counterfactuals, we reduce these costs by 12% so that the resulting impact on the intensive margin of informality is the same across both policies — an 18% reduction in the fraction of contract workers. Contrary to the previous policy however, this policy *reduces* informality on the extensive margin as well, and *increases* TFP and real income by 0.4% and 8% respectively.

In summary, our analysis yields several policy-relevant insights. First, it highlights the importance of careful design and targeting of labor market policies, as policymakers endeavor to reduce informality and encourage compliance with EPL. Second, penalizing formal sector firms for not complying with EPL does increase their compliance, but also discourages them from formalizing entirely, thereby leading to adverse effects on aggregate productivity, real wages, and welfare. Lastly, alleviating the costs and constraints associated with hiring workers on regular payroll on the other hand, can substantially reduce both dimensions of informality and positively impact the overall economy.

This paper contributes to multiple strands of the literature. Firstly, while there is an extensive body of research on the employment effects of EPL reforms,[17] much of it concentrates on advanced economies where informality is not widespread. A relatively smaller literature that has highlighted the role of informality in understanding these reforms, has primarily focused on studying either its extensive margin (Almeida et al., 2022; Besley and Burgess, 2004; Ponczek and Ulyssea, 2022), or the intensive margin (Chaurey, 2015; Samaniego de la Parra and Fernández Bujanda, 2020; Bertrand et al., 2024) separately.[18] Our contribution lies in demonstrating that examining both margins of informality offers additional insights on understanding the impacts of EPL. In this regard, our paper aligns

[15]Note that the model is estimated using a cross-section of data in Andhra Pradesh before the policy (see Section 6.2). However, it is able to replicate key reduced-form impacts of the policy, which increases our confidence in being able to meaningfully evaluate the impact of counterfactual policies as well.

[16]While aggregate labor supply is fixed in our model, it is possible that the policy could have changed the probability of labor force participation, unemployment, or transition from agriculture into non-agriculture. In Table A12, we show that the policy reform had no impact on any of these.

[17]See Kugler and Pica (2008); Autor, Donohue III and Schwab (2006); Bachas, Jaef and Jensen (2019); Bjuggren (2018); Boeri and Garibaldi (2007); Bornhäll, Daunfeldt and Rudholm (2015); Butschek and Sauermann (2022); Cahuc and Postel-Vinay (2002); Cahuc, Charlot and Malherbet (2016); Cahuc, Carry, Malherbet and Martins (2023); Schivardi and Torrini (2008).

[18]There is a related literature that examines the impact of trade reforms in the context of informality. See Dix-Carneiro et al. (2021); Ponczek and Ulyssea (2022); Vera et al. (2022).

closely with Ulyssea (2018) and Cisneros-Acevedo (2022). Ulyssea (2018) develops and estimates a structural model of heterogeneous firms in Brazil to show that labor market policies and informality need not move in the same direction, and may not lead to higher TFP or welfare. We complement this analysis by providing causal empirical evidence using a large-scale EPL policy reform in India. Cisneros-Acevedo (2022) shows how the Peruvian trade reforms reduced informality on the extensive margin and increased it on the intensive margin, complementing our analysis in a different policy context.

A second strand of the literature has examined the impact of EPL on the composition of a firm's workforce, with earlier studies primarily focusing on EPL for regular workers.[19] However, given the rapid rise of contract workers in recent years due to automation and globalization (Dix-Carneiro et al., 2021; Ponczek and Ulyssea, 2022), more recent literature has explored how EPL can affect their use (Aguirregabiria and Alonso-Borrego, 2014; Ardito et al., 2021; Baek and Park, 2018; Cahuc et al., 2023; Daruich et al., 2023; Hijzen et al., 2017). A particularly relevant paper in the Indian context is Bertrand et al. (2024), who show how formal sector Indian firms increased their usage of contract workers to circumvent restrictive Indian regulatory environments. While they emphasize the intensive margin of informality, our contribution complements their analysis in emphasizing the significance of carefully designing and targeting EPL policies that aim to regulate flexible and part-time work arrangements, particularly in contexts where both the intensive and extensive margins of informality are prevalent.

The rest of the paper is organized as follows: Section 2 develops the theoretical model that provides a basis for interpreting the empirical results. Section 3 discusses the policy reform in India. Section 4 describes the data and we present the empirical methodology and results in Section 5. Sections 6 and 7 discuss the quantification of the model, and counterfactual policy simulations. In Section 8, we conclude.

2 Theory

An economy consists of J industries and two sectors in each industry, namely the formal and informal sectors. There is a continuum of potential firms in each industry, with mass N_j, $j \in J$. Each firm is indexed by its productivity $z \sim F_j(z)$. Labor is the only input in production, and firms can hire homogeneous workers at a competitive wage w. Firms can endogenously decide whether to enter at all, and conditional on entry, whether to operate in the informal or formal sector, $s \in \{I, F\}$. We refer to this as the "extensive" margin of informality. Firm entry is static so that after entry, firms stay active forever. Firms in the formal sector can hire workers either on regular payroll or on temporary contracts i.e.

[19] See Botero et al. (2004); Cingano et al. (2010); Besley and Burgess (2004); Autor et al. (2007).

"off-the-books". We refer to this margin as the "intensive" margin of informality. The goods market is perfectly competitive and the formal and informal sectors are perfect substitutes in consumption.[20]

The key difference between the two sectors is that in the informal sector, a firm chooses to remain unregistered with the government. Therefore, it does not pay any taxes nor does it need to conform with EPL regulations. However, it faces a size-dependent penalty of operating in the informal sector. This penalty not only captures the cost of being large, evading taxes and getting caught, but also other constraints faced by operating in the informal sector such as access to formal channels of finance for example, where registration with a government agency is required.[21] On the other hand, in the formal sector, firms pay taxes, but can endogenously decide on its compliance with EPL by choosing the composition of its workers on payroll or on temporary contracts.

With this setup, we now turn to discussing the details of the model. In Section 2.1, we begin with the production decision of incumbent firms in the formal and informal sectors. In Section 2.2, we then discuss the entry decisions of a firm and define equilibrium in Section 2.3. Lastly, we discuss the implications of a policy that penalizes the hiring of contract labor by formal firms in Section 2.4. Moving forward, we drop firm and industry subscripts for notational clarity.

2.1 Incumbents

Output of a firm with productivity z is given by a decreasing returns to scale production function, $y(z) = zl^{\rho}$, where $0 < \rho < 1$, and l is the labor used in production. We now discuss a firm's production decision in the formal and informal sectors below.

Formal Sector: Two features define production in the formal sector. First, firms must pay a per-unit tax t on sales. Second, they can hire workers on both regular payroll (l_r) as well as "off-the-books" on informal contracts (l_c). Both types of workers are imperfect substitutes in production[22] with an elasticity of substitution ν, such that $l_F = \left[l_r^{\frac{\nu-1}{\nu}} + l_c^{\frac{\nu-1}{\nu}} \right]^{\frac{\nu}{\nu-1}}$. However, apart from the wage w, firms pay additional marginal

[20] To measure the overlap of firms manufacturing similar products or operating in the same (finely measured) industry, we use establishment-level data from the 2005 round of the Annual Survey of Industries and National Sample Survey. These data capture the distribution of firms in both the formal and informal sectors. We find that 97-98.5% of industries as measured by their 3-digit, 4-digit and 5-digit NIC classifications have both formal and informal firms.

[21] See La Porta and Shleifer (2014), Beck and Hoseini (2014) and Nikaido, Pais and Sarma (2015).

[22] In Appendix C1.1, we provide evidence to show that while payroll and contract workers might be perfect substitutes within a task, there is substantial variation in their allocation across tasks. Since we do not observe task-level allocation of workers empirically, modeling them to be imperfect substitutes in production is a reasonable first-order approximation.

costs b_r and $\tilde{b}_c = b_c z^{-\phi}$ to hire a payroll and contract worker respectively. These costs (b_c and b_r) are a reduced-form way of incorporating compliance with labor laws for regular payroll workers such as incurring hiring/firing costs, as well as search costs, intermediary contract worker payments, etc., for hiring contract workers. Note that in the estimation, we will not impose any restrictions on the value of these parameters i.e., they could be greater or less than 1. Lastly, as emphasized by Bertrand et al. (2024), as well as as shown in Figure A4 for our sample, the fraction of contract workers increases with firm size. We therefore allow for the "effective" marginal cost of hiring contract workers (denoted by $\tilde{b}_c$ above) to decline with firm productivity (and hence firm size). As we will show subsequently in Equation 1 , ϕ is the elasticity of how the ratio of contract to regular workers changes with firm productivity.[23] The variable profit of a firm manufacturing in the formal sector is therefore:

$$\pi_F(z) = (1-t)pzl_F^{\rho} - wb_r l_r - w\tilde{b}_c l_c$$

Define: $w_F = \left[b_r^{1-\nu} + \tilde{b}_c^{1-\nu}\right]^{\frac{1}{1-\nu}} w$. The ratio of contract to regular workers and the variable profit of the firm in the formal sector can therefore be given by:

$$\frac{l_c}{l_r} = \left(\frac{b_c}{b_r}\right)^{-\nu} \times z^{\phi\nu} \tag{1}$$

$$l_F(z) = \left[\rho(1-t) \times \frac{z}{w_F/p}\right]^{\frac{1}{1-\rho}} \tag{2}$$

$$\pi_F(z) = (1-\rho)(1-t)pzl_F^{\rho}(z) \tag{3}$$

See proof in Appendix C1.2. Note that although b_c/b_r captures the additional marginal costs in our model, in theory, they could also reflect (industry-specific) differences in the relative productivity of contract and regular workers. For example, let $B = b_c/b_r = A_c/A_r \times \hat{b}_c/\hat{b}_r$, where $A = A_c/A_r$ is the relative productivity of contract workers to payroll workers, and $\widehat{B} = \hat{b}_c/\hat{b}_r$ is the true difference in the relative marginal costs of hiring contract and regular payroll workers, after taking into account the productivity differences. Our current data structure does not allow us to disentangle them, and our calibration exercise in Section 6 will identify B instead of $\widehat{B}$. However, our analysis and counterfactual policies will focus on changing the marginal costs of hiring contract (or payroll) workers, keeping the relative productivity differences (A) unchanged. Therefore,

[23] More generally, we potentially allow for the marginal costs for both the payroll and contract workers to depend on firm productivity (z) i.e., $\{\tilde{b}_r, \tilde{b}_c\} = \{b_r z^{\phi_r}, b_c z^{\phi_c}\}$. As we will show later, we will effectively be only able to identify the relative difference between ϕ_c and ϕ_r i.e., $\phi = \phi_r - \phi_c$. We can therefore normalize $\phi_r = 0$.

if we define $\triangle X = x_{new}/x_{baseline}$, then $\triangle B = \triangle \widehat{B}$ (since $\triangle A = 1$).

Informal Sector: Firms in the informal sector are unregistered with the government and tax authorities. Therefore, they can only hire informal contract workers, but do not have to pay taxes, or comply with EPL regulations (b_c and b_r). However, they face size-based challenges of operating in the informal sector such as the probability of being audited (and potentially penalized) by the government, limited access to formal channels of credit, etc. We approximate this distortion in the form of a convex cost function $c(l_I) = wl_I^\theta$, where l_I is the size of the firm, and $\theta \geq 1$.[24] The variable profit function of a firm in the informal sector is therefore given by:

$$\pi_I(z) = pzl_I^\rho - wl_I^\theta$$

Define $\widetilde{\rho} = \rho/\theta < 1$. The optimal profit of the firm is then given by:

$$l_I(z) = \left[\widetilde{\rho} \times \frac{z}{w/p}\right]^{\frac{1}{\theta-\rho}} \tag{4}$$

$$\pi_I(z) = (1-\widetilde{\rho})pzl_I^\rho(z) \tag{5}$$

See proof in Appendix C1.3.

2.2 Firm Entry

Let there be a mass of potential entrants N indexed by a pre-entry signal of productivity denoted by $x \sim F(x)$. We assume that $F(x)$ is continuous with support $(0, \infty)$, has finite moments and is independent across firms. For a sector $s \in \{I, F\}$, let E_s be the fixed cost of entry (in units of output) that a firm has to pay to enter the sector. We assume that entry costs in the formal sector are higher than the informal sector so that $E_F = E_I + E_R$, where E_R capture fixed costs associated with registering the firm with the government (for example, bureaucratic procedures and red-tape).

After entry occurs, an entrepreneur draws his/her actual productivity from a conditional cumulative distribution $G(z|x)$, which is assumed to be continuous in both z and x and is strictly decreasing in x. This implies that a high pre-entry signal x is indicative of a high post-entry draw of productivity z as well.[25] An entrepreneur with pre-entry

[24] Appendix C1.4 relates the convex-costs to a size-based penalty function for informal firms.

[25] This entry structure is similar to Ulyssea (2018) in that it allows for an overlap of the firm-size distribution between the informal and formal sectors, a salient feature of firms in the Indian economy. An entry decision in the spirit of Melitz (2003) on the other hand, would predict perfect sorting of firms between the formal and informal sectors based on firm size, which is inconsistent with the data.

productivity x, operating in sector s would have an expected variable profit $V_s(x)$ given by:

$$V_s(x) = \int \pi_s(z) dG(z|x) \tag{6}$$

Therefore, the entrepreneur will enter a sector s as long as $V_s(x) - pE_s \geq \max\{V_{s'}(x) - pE_{s'}, 0\}$. Given that $E_F > E_I$ and we observe positive entry into both sectors, this implies that there is a positive threshold level of productivity x_s^* in each sector s such that:

$$\begin{aligned} V_I(x_I^*) &= pE_I \\ V_F(x_F^*) - V_I(x_F^*) &= pE_R \end{aligned} \tag{7}$$

Entrepreneurs with $x < x_I^*$ will not enter at all, those with $x \in [x_I^*, x_F^*]$ will enter the informal sector, and $x > x_F^*$ will enter the formal sector.

2.3 Equilibrium and Welfare

We now specify the demand side to close the model in equilibrium. We assume a representative consumer who inelastically supplies labor $\overline{L}$ (no disutility from labor) and does not save. Consumer income therefore simply aggregates wage income, profits and taxes across the economy and is therefore given by $I = w\overline{L} + \Pi + T$, where Π denotes the total profits (net of fixed costs) of incumbent entrepreneurs and T is the total tax-revenue collected by the government that is redistributed back to the consumers. We assume a simple demand system where this representative consumer consumes a composite bundle C, whose price is normalized to 1, and a constant share of income κ_j is spent on consumption from industry j. Total consumption (and hence real income) is therefore a natural measure of welfare. The equilibrium in this economy is therefore characterized by a set of prices for each industry $\{p_j\}_{\forall j}$ and wages w such that:

(i) the goods market clears in each industry i.e., the total output produced in an industry j (across both the informal and formal sectors) is either demanded by consumers or paid in fixed costs by the incumbent firms.

(ii) the labor market clears so that the total labor demand across industries and sectors is equal to the size of the labor force in the economy.

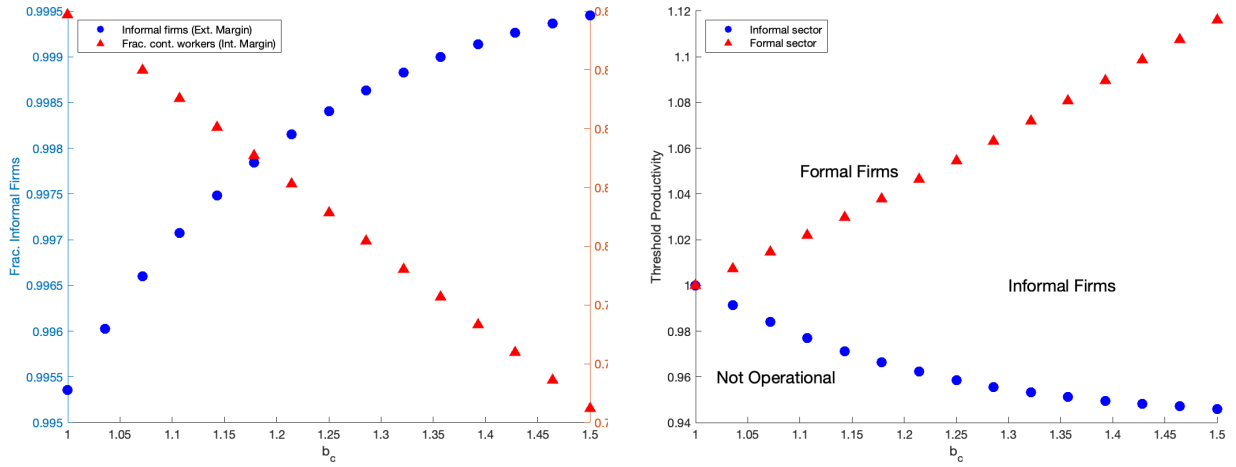
(iii) the zero-profit condition in Equation (7) holds with equality in both sectors.

2.4 Impact of Enforcing EPL and Informality

Stricter enforcement of EPL would therefore encourage firms to hire more workers on payroll and less on temporary contracts. In this section, we consider the impact of pe-

nalizing the hiring of contract workers in the formal sector on both the intensive and extensive margins of informality. From the perspective of the model, this corresponds to an increase in b_c. To understand the mechanism, we fix some parameter values and simulate the model in equilibrium and examine its implications.

Figure 1: Theoretical Implications of Penalizing Hiring Contract Workers



(a) Margins of Informality

(b) Productivity Thresholds

Note: The above figure examines the impact of penalizing the hiring of contract workers on the intensive and extensive margins of informality (Graph A) and on the productivity of the marginal firm, which is indifferent between entering the informal and formal sectors (Graph B). We start with a baseline scenario of $b_c = 1$, and examine the impact of gradually increasing b_c. We use the following parameter values for the simulation: $b_p = 1$; $E_I = 0.5$; $E_R = 50$; $\tau = 0.05$; $\sigma_x = \sigma_\varepsilon = 0.2$; $\nu = 3$; $\phi = 1$; $\rho = 0.74$; $\theta = 1.15$; $N = L = 5m$.

We start with a baseline scenario where $b_c = 1$ and increase b_c with each counterfactual simulation, leaving b_r unchanged. Figure 1a shows the impact of this increase on both margins of informality. As can be seen from the figure, the intensive margin of informality i.e., the fraction of workers hired on contracts in the formal sector (denoted by red triangles) declines as b_c increases. However, there is also an increase on the extensive margin of informality (denoted by blue dots) as more firms now choose to operate in the informal sector instead of formalizing. This is what we explore in Figure 1b, where we plot the productivity of the marginal firm that is indifferent between entering the informal sector or not (denoted by blue dots) and formalizing or not (denoted by red triangles). We see that as b_c increases, the productivity of the marginal firm in the formal sector (red triangles) also increases. Lastly, the size of the informal sector (area between the blue dots and red triangles) also increases as b_c increases.

The underlying mechanism (that we will come back to again in Section 7) is more nuanced. The policy reform increases the size of the informal sector in two ways. First, it has a direct effect: the marginal formal sector firm at baseline now chooses to remain

informal as b_c increases (red triangles in Figure 1b). However, since firms are less productive in the informal sector, the resulting changes in general equilibrium–through lower labor demand and lower real wages–has a second indirect effect: low-productivity firms now choose to enter the informal sector i.e., the threshold productivity of the marginal firm choosing to enter now decreases (blue dots in Figure 1b). Therefore, although the policy reduces informality on the intensive margin, it increases informality on the extensive margin through both the direct and indirect channels.

3 Context of the EPL Reform in India

An Overview of Labor Laws in India: Before discussing the policy reform instituted by the Government of Andhra Pradesh, it is important to understand key labor laws in India, along with their implications for hiring workers on payroll or temporary contracts. The Industrial Disputes Act (1947) makes it expensive for a firm to dismiss or layoff its hired workers on payroll. For example, firms retrenching a worker must offer severance pay, seek government permission, and issue advance notification in case of closures.[26] On the other hand, contract (or fixed-term) workers are not protected under the IDA. These workers are not directly hired on the payroll of a firm, but are usually hired on fixed-term contracts with licensed contractors. The Contract Labour Act, 1970 (CLA) allows the use of contract workers in any firm with a minimum of twenty or more workers, but mandates that such establishments register with the government to use contract workers, and their contractors obtain a license to operate. This naturally creates a tradeoff where since the IDA is not applicable to contract workers, firms normally hire them to circumvent the payment of formal benefits as well as associated high dismissal costs. In fact, several studies provide robust evidence for this channel by examining the rise in the usage of contract labor in being able to adjust to unexpected shocks (Bertrand et al., 2024), particularly by firms operating in states with stringent employment protection laws (Chaurey, 2015; Ahsan and Pagés, 2009), or in response to India's trade liberalization in the 1990s (Saha et al., 2013; Chakraborty et al., 2020).

The Need to Regularize Contract Workers: Being able to circumvent IDA implied that contract workers usually lacked job and social security benefits, and were meted out

[26]Section V-A of the IDA lays down regulations for establishments with 50 or more workers. For example, a retrenched worker is entitled to compensation equalling 15 days' average pay for each year of service, and for layoffs, every worker is paid fifty percent of the basic wages and a dearness allowance for each day that they are laid off (maximum of 45 days). Section V-B mandates that no worker may be laid-off or retrenched in large firms (of size 100 and above) without prior permission of the government. Establishments that want to close down are also required to issue a sixty days (Section V-A) or ninety days (Section V-B) notification to the government prior to the shutting down. Both these sections of the IDA make it costly for firms to fire workers.

with poor working conditions. There was an active debate on whether and how policy could intervene to limit this, with two main viewpoints: the first was to explicitly limit the usage of contract workers by firms. For example, as noted by the Bharatiya Mazdoor Sangh (BMS), one of the apex labor unions in India: *"[employers] prefer contract labour system because of two main reasons, viz. workers need be given only lower wages and they can be thrown out at any time. This is sheer exploitation. The tragic plight of contract workers is still worsened by the advent of labour law reforms during the 1990s."* (Narayanan et al., 2005). In fact, BMS strongly advocated against contractualization of workers and passing reforms to regularize them. The alternate viewpoint, advocated by entrepreneurs, and supported in the academic literature,[27] was to simplify the complex and cumbersome employment protection laws in India. As noted by a TeamLease (2006) report, *"there are nearly 2500 Central laws in force...[and] there is not even a rough estimate available about the number of laws operating in states. In one state alone the number is stated to be of the order of 1100..."*. Moreover, many of these laws were antiquated, arbitrary and their enforcement relied on multiple labor inspectors who inspected different statutes and held a lot of discretion on determining violations. It was therefore argued that simplifying the regulatory burden on firms would also encourage them to regularize their workforce.

EPL Reform in Andhra Pradesh: It was in the above context that the Government of Andhra Pradesh issued a notification, formally termed as *"Contract Labour (Regulation and Abolition) (Andhra Pradesh Amendment) Act, 2003"*. The policy reform, in line with the BMS viewpoint discussed above, explicitly prohibited the use of contract workers in *core* activities of the firm,[28] either directly or through licensed contractors. It applied to all firms in the state of Andhra Pradesh irrespective of their size and sector and in case of any violations, the penalties under the original Contract Labour Act 1970 would apply (for e.g. substantial fines, potential imprisonment of Directors, and/or prosecution in court). There were three exceptions however, where employing contract workers was permitted: (i) if the work was "normally" done through contractors; or (ii) the activities themselves did not require full-time workers; or (iii) if there was a sudden increase in the volume of core activity work that needed to be completed within a certain time.

The reform was therefore aimed at protecting workers by limiting the usage of contract workers and encouraging firms to hire them on regular payroll (thus conforming with the IDA regulations). Using data from the 1998 round of the Economic Census of India, Figure A2 shows that the fraction of unregistered, informal firms in an industry

[27] See Besley and Burgess (2004); Ahsan and Pagés (2009); Sodhi (2014); Chaurey (2015)

[28] Core activities included the primary activity for which an establishment was set up, along with other activities essential to this primary activity. All other activities (such as catering, housekeeping, transport, etc.) were classified as "non-core" activities, and the use of contract workers was permitted in them. See Appendix B for a complete list.

(agriculture, manufacturing, or services) in Andhra Pradesh, as well as the fraction of workers in them were comparable to other states in India. This suggests that Andhra Pradesh was not an exception in the Indian context. However, it also raises interest regarding the timing of when Andhra Pradesh passed the legislation. While it is not possible to definitively say much about it, some potential factors that could have played a role include the fact that despite a literate and skilled labor force, the state performed poorly on labor protection and implementation of labor laws.[29]. Moreover, the BMS was quite active in Andhra Pradesh and wielded political influence through organizing large rallies, strikes, and protests.

4 Data

4.1 Annual Survey of Industries

Our primary source of data is the panel version of the Annual Survey of Industries (ASI), administered by the Ministry of Statistics and Programme Implementation (MoSPI), Government of India. The ASI is an annual representative survey of factories registered under The Factories Act (1949), a central piece of legislation regulating manufacturing establishments in India.[30] The ASI comprises two parts: first, the Census Sector, which is a census of establishments employing 100 or more workers; second, the Sample Sector, consisting of establishments that are randomly sampled using a systematic circular sampling technique within each State $\times$ Industry $\times$ Sector $\times$ 4-digit NIC industry stratum.[31]

We use a panel of establishments from the ASI between 2000-2001 to 2005-2006 for this study. While this choice is driven by being able to consistently measure outcome variables across multiple data sources (as we will subsequently discuss), we show the persistence of our results by extending the ASI data until 2009-10. The reference period for ASI is the accounting year of the factory ending on the last day of the fiscal year between April to March. For instance, the data for the year 2005-2006 corresponds to all activities between 1^{st} April 2005 and 31^{st} March 2006. A unique feature of the data that is relevant for this study is that it contains detailed information on the number of contract and regular workers, their worker-days employed, and further distinguishes employment in core and non-core activities as well. However, the ASI does not provide

[29]For example, as noted in the TeamLease (2006) report, Andhra Pradesh ranked in the top 2 among the 6 comparable developed states in western and southern India on the quality of its labor force, but ranked in the bottom 2 among these states on the quality of its labor laws.

[30]All factories employing 10 workers or more (without using electricity) or employing 20 workers or more (with or without using electricity), are required to register under the Factories Act.

[31]Apart from the large establishments, the Census Sector also comprises of: (1) all industrial units belonging to the six less industrially developed states and Union territories viz. Manipur, Meghalaya, Nagaland, Sikkim, Tripura and Andaman & Nicobar Islands; (2) all factories filing Joint Returns.

Table 1: Summary Statistics from the ASI Data

	Andhra Pradesh			All Other States		
	N	Mean	SD	N	Mean	SD
	(1)	(2)	(3)	(4)	(5)	(6)
Contract Workers	5759	39.03	96.31	77665	22.91	66.49
Regular Workers	5759	117.66	219.41	77665	113.90	207.60
Total Workers	5759	174.61	310.54	77665	140.91	247.71
Contract Worker-days ('000)	5759	11.30	28.62	77665	6.60	19.76
Regular Worker-days ('000)	5759	33.39	64.02	77665	32.57	60.52
Frac. Contract Workers	5759	0.22	0.34	77665	0.15	0.30
Frac. Contract Worker-days	5759	0.23	0.34	77665	0.15	0.30

Notes: The above table uses data from the Annual Survey of Industries between 1999-2002. Each row reports an outcome variable. Columns (1)-(3) report the total number of observations, mean, and standard deviation for the variable in Andhra Pradesh, while Columns (4)-(6) report the same for all other Indian states.

occupation- or task-level allocation of workers (either on payroll or contract), which limits our examination (both theoretically and empirically) of the within-firm reallocation of workers in response to the policy.

We use the ASI to study the impact of the policy amendment on firms' usage of labor — contract, regular, and total workers, as well as worker-days. Data on worker-days are further disaggregated between core and non-core activities of a firm, which are useful in our context (see the discussion in Section 3) to conduct placebo tests. A key limitation of the ASI is that it covers (mostly formal) establishments in the manufacturing sector only. The manufacturing sector accounted for around 20-25% of firms and workers in India around this time, with around 25% of manufacturing workers working in the formal sector. To overcome this limitation, we complement our analysis with two additional datasets, namely the National Sample Surveys and the Economic Census that we discuss subsequently.

4.2 National Sample Surveys

We use the Employment-Unemployment schedules from the National Sample Surveys (NSS), which are repeated cross-sections of a nationally representative sample of households. Among other things, the NSS collects detailed data on the employment of individuals, including: (a) the characteristics of the workplace (for e.g., firm size); (b) nature of employment, namely, regular or casual (informal) work;[32] (c) nature of payment

[32]The NSS defines a casual worker as *"a person engaged in farm or non-farm enterprises (both household and non-household) and getting in return a wage according to the terms of the daily or periodic work contract."*

Table 2: Overlap Across Measures of Informality

	Fraction of Workers in:		
	Informal Contract	Informal Firm	Casual Work
	(1)	(2)	(3)
Informal Contract	1.00	0.79	0.89
Informal Firm	0.70	1.00	0.88
Casual Work	0.87	0.94	1.00

Notes: The above table uses data rounds of the National Sample Survey between 1999-2005. Each outcome variable takes the value 1 if an individual reports working on an informal contract (Column 1), in an informal firm (Column 2), or in casual work (Column 3). Each panel reports the fraction of individuals in each of the above categories if a worker is on an informal contract (Row 1), informal firm (Row 2), and casual work (Row 3). See Appendix Table A1 for detailed definitions for each variable.

(whether piece-rate or flat-rate), along with the frequency of payment (daily, weekly, monthly etc.); (d) number of days worked and earnings in the past week. Unlike the ASI, the NSS provides a representative picture of workers both in the formal and in the informal sectors as well as across industries. Given the nature of the reform, we restrict the sample to those in the working age between ages 14 to 65, and those employed in the non-agricultural sector (to make it consistent with the Economic Census data discussed subsequently).[33] We use information in the NSS to define various binary measures of informality described below:

1. **Whether an individual works on an informal contract:** Workers report the frequency and type (cash, kind, or both) of their payments. We classify workers who receive payments on either a daily or weekly basis, along with those who provide unpaid labor as those working on informal contracts.

2. **Whether an individual works in an informal firm:** Workers report the size of the firm that they work in, as well as whether the firm uses electricity or not. These are intentionally collected in the NSS since they are a common definition of "formality" in the Indian context. Specifically, firms with more than 20 workers, or more than 10 workers and using electricity (Amirapu and Gechter, 2020) are considered formal. We therefore define a worker to be informal if s/he does not work in such a firm. In addition, we also include individuals who work in an owner-operated enterprise, or those working in household enterprises to be informal workers as well.

3. **Whether an individual is a casual worker:** The NSS asks individuals to report

[33] In Table A12, we show that the reform had no impact on the probability that a worker participated in the labor force, was unemployed, or the probability of working in agriculture.

their "usual status of employment". These range from: (i) salaried and wage employment; (ii) self-employed as an entrepreneur; (iii) self-employed in an owner-operated enterprise; (iv) unpaid work in a household enterprise; (v) casual work in a public welfare program; (vi) other types of non-standard work. We consider an individual to be a casual worker for categories (iii)-(vi).

Lastly, we calculate wages as the ratio of total earnings divided by the number of days worked. Table A1 in the Appendix provides detailed definitions for all variables as well as their coverage across different rounds of the NSS. In our primary analysis, we restrict our attention to three rounds of the NSS for the years 1999-2000, 2003-04, and 2004-2005. Our sample consists of all individuals in the labor force, including those who are self-employed or work in unpaid jobs.

As we discuss later in Section 5.1, our results are robust to: (i) other ways to construct measures of informality; (ii) excluding self-employed individuals; (iii) restricting attention to the manufacturing sector (to make it consistent with the ASI) (iv) using additional years in which data on these variables is available.

4.3 Economic Census Data

The Economic Census is an enumeration of all non-farm establishments in India. Our primary analysis is based on the 1998 and 2005 rounds of the EC, which contains information on the firm's location and industry, along with whether the firm is registered and conforms with various government regulations (like the Factories Act, Shop & Establishment Act, the Cooperative Society/Labour Act, etc.). The EC also reports data on the total number of workers, but not data separately on contract and regular workers. We use this firm-level information to examine the impact of the policy reform on the "extensive margin" of informality i.e., the probability that a firm remains unregistered. There are two additional rounds of the Economic Census in 1990 and 2013 that we do not use for this study because they do not collect information on the registration status of firms, which is an important feature in our analysis. Our results in the EC should therefore be interpreted with appropriate caution, since we cannot speak to the pre-trends in the Economic Census data.

4.4 Summary Statistics

We present summary statistics for key outcome variables of interest in the ASI data before the policy reform (i.e., between 2000-2002) in Table 1. Columns (1)-(3) report the sample size, mean, and standard deviation for Andhra Pradesh (Treatment state), while Columns (4)-(6) report them for all other states in India (Control states). Of special note

is the fact that around around 15-20% of workers in a formal firm work on temporary contracts. Turning to the NSS, Table 2 provides an overlap between different definitions of informality. Each number in this table is the fraction of individuals defined in the category in each row that overlap with the category in each column. By construction therefore, the diagonal elements in this table are 1. From the first row, we see that 80% (89%) of workers on an informal contract also work in an informal firm (in casual work). Similarly, 70% (88%) of workers who work in an informal firm work on an informal contract (in casual work). Lastly, 87% and 94% of casual workers work on an informal contract or in an informal firm. Lastly, from the Economic Census, over 90% of firms in India are informal (or unregistered), employing around 70% of the laborforce.

5 Empirical Methodology and Results

We analyze the impact of the policy reform in a difference-in-differences (DID) setup, by comparing outcomes before and after the policy change (2003) in the treated state (Andhra Pradesh) as compared to alternate control states. In our preferred specification, we compare firm-level outcomes in Andhra Pradesh (Treatment group) to all other Indian states (Control group). Our identification strategy relies on the assumption that the treatment and control groups have similar trends before the policy reform, which we provide evidence for in Figure 2 and discuss in more detail below. Finally, to the best of our knowledge, there was no other labor market policy that was implemented in Andhra Pradesh in 2003 that affected its firms differentially as compared to other states, which increases confidence in our identification strategy.

5.1 Impact on Workers in Formal Firms (Intensive Margin of Informality)

We first examine the impact of the policy on the intensive margin of informality i.e., the use of contract workers in formal sector firms. Using the ASI data, we first test whether the treatment and control groups had parallel trends before the policy, and examine the impact of the policy over time. We then pool the pre- and post-policy years together to examine the impact of the policy on a wider set of firm and worker outcomes using the ASI and NSS datasets. We begin estimating the following event-study regression for a firm i in industry j, state s and year t:

$$Y_{it} = \alpha_i + \theta_s t + \alpha_{jt} + \sum_{t=2000}^{2006} \beta_t Treat_s + \varepsilon_{it} \tag{8}$$

where Y_{it} is our outcome of interest. We consider six outcome variables, namely: (a) log contract workers; (b) log contract worker-days; (c) fraction of contract workers (d) fraction of contract worker-days; (e) log regular workers; (f) log regular worker-days.[34] We consider, 2002-2003, one year before the policy was implemented, to be our reference year. $Treat_s$ is an indicator variable that takes the value 1 if a firm is in Andhra Pradesh and 0 otherwise. We also add firm fixed effects (α_i) that control for any time-invariant observed and unobserved heterogeneity at the firm level.[35,36] Although the DID method controls for time invariant unobservable firm characteristics, one may be concerned that the timing of passing the Act in Andhra Pradesh could be correlated with time-varying differences across industries and states. In fact, Indian states with similar patterns of labor regulation and reforms do have similar long-term trends as well (Besley and Burgess, 2004; Karak and Basu, 2020; Bhattacharjea, 2006). We therefore include a 3-digit industry-year fixed effects (α_{jt}) as well as state-specific time trends ($\theta_s t$) in all our regression specifications. However, Borusyak, Jaravel and Spiess (2024) show that an event-study specification with firm and year fixed effects along with state-specific time trends is under-identified. They recommend omitting one additional year in the estimation (1999-2000 in our case), and reporting the p-value of a F-test that jointly tests that the estimated $\widehat{\beta}_t$ before the policy (pre-period) were zero.[37] Lastly, we cluster standard errors at the state level to allow for correlations across firms within a state.

The results are reported in Figure 2, and the corresponding regression coefficients are reported in Table A2 in the Appendix. The results indicate that there were no significant differences in the usage of contract or regular workers between firms in Andhra Pradesh and other states before the policy was implemented. Both, the estimated magnitudes are small and they are statistically insignificant at conventional levels. Following Borusyak et al. (2024), for each outcome variable in Figure 2, the corresponding p-value of a F-test that jointly tests for the pre-policy $\widehat{\beta}_t$ to be equal to 0 are: 0.74 and 0.96 for log contract

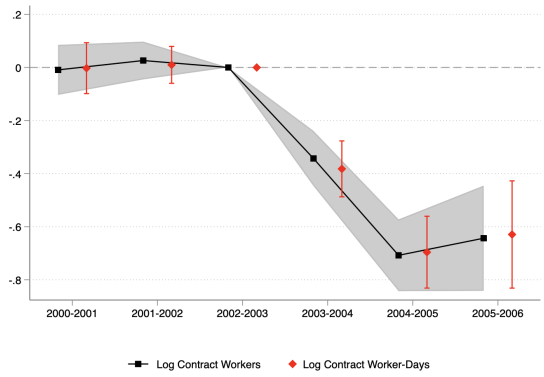
[34]The amendment only affected hiring of contract labor in core activities of the firm (as discussed in Section 3). The ASI data does not disaggregate the number of workers by core and non-core activities, but data on worker-days is available for both categories separately. In Section **??**, we take advantage of this to separately analyze the impact of the policy on contract and regular worker-days in both core and non-core activities. We show that all the effect is driven by adjustments in core activities.

[35]Note that the inclusion of firm fixed effects implies that the regression results can be interpreted only on the intensive margin of the policy change, i.e. the effect of the policy change on incumbent firms.

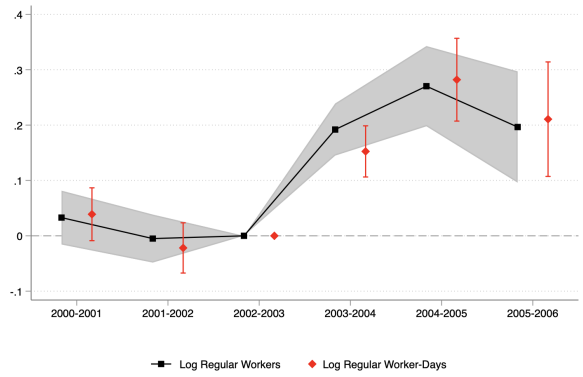
[36]Our set up follows the canonical format of the differences-in-differences design using two time periods (before and after the policy reform), and two groups. Therefore, we can directly estimate the average treatment effect for the treated (ATT) by comparing the average change in outcomes experienced by the treated group to the average change in outcomes experienced by the comparison group. Employing recent methods in the differences-in-differences estimation literature (Goodman-Bacon, 2018; De Chaisemartin and d'Haultfoeuille, 2020; Callaway and Sant'Anna, 2020) is not warranted here because they are concerned with issues arising due to multiple groups and multiple periods.

[37]The results are not sensitive to which pre-policy year is omitted from the regression and the F-test (by construction) is invariant to which pre-policy year is omitted.

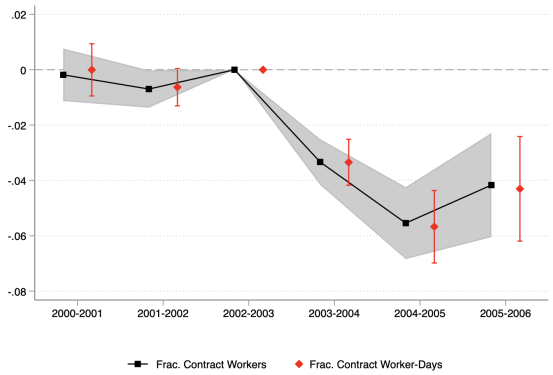
Figure 2: Impact of the Policy on the Use of Contract and Regular Workers



(a) Log Contract Workers



(b) Log Regular Workers



(c) Frac. Contract Workers

Notes: The above graphs plot the regression coefficients from a difference-in-differences specification using data from the Annual Survey of Industries. The regressions corresponding to these figures are presented in Table A2. The black squares report the results for workers, while the red diamonds report the results for worker-days. All regressions contain firm and 2-digit industry-year fixed effects, along with a state-time trend. 95% confidence intervals, clustered at the state level are indicated around the point estimate. The coefficient for 2002-2003, the year before the reform, has been normalized to zero. The outcome variables are winsorized at the 1% levels at the top and the bottom of the distribution.

workers and worker-days (Figure 2(a)); 0.38 and 0.11 for log regular workers and worker-days (Figure 2(b)); 0.12 and 0.11 for the fraction contract workers and worker-days (Figure 2(c)). Turning to the impact of the policy reform, we find large and persistent declines in the usage of contract workers (worker-days) and in contrast, a corresponding increase in the hiring of regular workers (worker-days), indicating a substantial reduction in the "intensive-margin" of informality.

For the subsequent analysis, we pool the time-periods before and after the policy and estimate the following regression for a firm i, in industry j, state s, and year t:

$$Y_{i(js)t} = \beta Treat_s \times Post_t + \alpha_i + \theta_s t + \alpha_{jt} + \gamma X_{it} + \varepsilon_{it} \tag{9}$$

where Y_{it} are various firm-level variables of interest. $Post_t$ is an indicator variable that takes the value 1 for years after 2003 and 0 otherwise. β is our coefficient of interest and captures the differential impact of the policy reform on a firm in the treatment group relative to the control group, before and after the policy reform. In addition to the fixed effects discussed in Equation 8, we control for age and square of age of the firm.

Table 3: Impact of the Policy on Contract, Payroll and Total Workers in Formal Sector Firms

	Workers		Worker-Days	
	(1)	(2)	(3)	(4)
Dep Var:	Log Contract	Frac. Contract	Log Contract	Frac. Contract
Panel A: Impact on Usage of Contract Workers				
Post X Treat	-0.38***	-0.03***	-0.41***	-0.03***
	(0.06)	(0.00)	(0.06)	(0.00)
R2	0.80	0.81	0.79	0.81
Panel B: Impact on Payroll Workers and Firm Size				
Dep Var:	Log Regular	Log Total	Log Regular	Log Total
Post X Treat	0.21***	0.05***	0.18***	0.03
	(0.03)	(0.02)	(0.03)	(0.02)
R2	0.85	0.94	0.86	0.93
N	161552	161552	161552	161552

Notes: The above table uses establishment-level data from the Annual Survey of Industries between 1999-2006. It reports the impact of the policy reform on Workers in Columns (1) and (2) and Worker-Days in Columns (3) and (4). Panel A reports the impact of the policy on the usage of contract workers, while Panel B reports the impact on the usage of regular payroll workers and total workers. Frac. Contract is the number of contract workers (or worker-days) as a fraction of total workers (worker-days) in a firm. Post is defined as 1 for years after 2003, and 0 before that. Treat is defined as 1 for Andhra Pradesh and 0 for other states. The outcome variables are winsorized from the top and bottom at 1%. All regressions contain firm and 2-digit industry-year fixed effects, a state-time trend, and control for age and age-square of the firm. Robust standard errors clustered at the state level in parentheses. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

Table 3 presents the results. Panel A reports the impact of the policy on contract workers (Columns 1-2) and worker-days (Columns 3-4). From Columns (1) and (3), we see that firms in Andhra Pradesh reduced the number of contract workers and worker-days by around 38-41 log-points (46-50%). From Columns (2) and (4), contract workers (and worker-days) as a fraction of total workers (worker-days) declined by around 3.1 p.p. (or 18% of the pre-period mean in control states).

Panel B of Table 3 then examines the impact of the policy on hiring regular workers and firm size. As reported in Columns (1) and (3) of the table, firms increased the hiring of regular workers and worker-days by 18-21 log-points (around 20-25%). Fur-

thermore, firm-size as measured by the total workers and worker-days (Columns (2) and (4)) may have increased slightly by around 3-5%. Taken together, these results show that as intended, the policy reform did result in a substantial reduction of informality on the intensive margin.

Robustness of the results: We conduct a series of additional analyses to test the robustness of our results. To begin, since the policy reform only affected workers in the primary or "core" activities of a formal firm, as opposed to allied "non-core" activities (such as cleaners, janitors office boys, and *chaiwalas* for example), we examine the effects of the policy on non-core activities of the firm as a placebo check. While the ASI does not provide the number of workers in these activities, it provides the number of worker-days in them. We use these to replicate our analysis separately for core and non-core activities. As reported in the Appendix Table A3, we find that the effects of the policy are entirely driven by a decline (increase) in the use of contract (regular) workers in core activities (Panel A), and not in non-core activities (Panel B), where the estimated magnitudes are very small and statistically insignificant at conventional levels, except a statistically significant and negative effect on regular workers in non-core activities.

Second, our preferred specification uses all other states in India except Andhra Pradesh as a control group. However, we also check the robustness of our results using only the states bordering Andhra Pradesh (such as Odisha, Chhattisgarh, Maharashtra, Karnataka, and Tamil Nadu) as a control group instead. As reported in Panel A of Appendix Table A4, this does not impact our results meaningfully. In contrast to the previous exercise, a concern might be that there are spatial spillovers across state borders. Therefore, in Panel B of Appendix Table A5, we report the results with now a control group that *excludes* states bordering Andhra Pradesh. The results are again similar to those in our preferred specification (Table 3).

Third, our main results included firm, and industry-year fixed effects, along with state-specific time trends. In table A6, we show that our results are robust to only keeping a year fixed effect and a state-time trend (Panel A), or additionally adding industry-year (Panel B) or state-industry (Panel C) fixed effects, or adding both industry-year and state-industry fixed effects (Panel D). The results are qualitatively similar, and in many cases quantitatively stronger than our preferred results (reported in Panel E for ease of comparison).

Lastly, as discussed in Section 4 and shown in Table A1, we restrict our sample between 1999-2006 for the main analysis to have a consistent set of outcome variables that are measured across all the datasets that we use. However, we now extend the sample to 2008-09 and estimate the event-study specification in Equation 8. As reported in Figure A3, we find no evidence of differential pre-trends (similar to Figure 2), and a large

and persistent impact of the policy that only strengthens over time.

5.2 Impact on Likelihood of Operating in the Informal Sector (Extensive Margin of Informality)

The ASI data covers firms in the formal manufacturing sector, and hence is only well suited to examine firm responses on the intensive margin of informality, that is, the substitution between contract and regular/payroll workers within these firms. We now complement this analysis using data from the National Sample Surveys (NSS) and the Economic Census (EC). This allows us to not only examine the impact across all non-agricultural industries, but also on the extensive margin of informality, i.e., whether firms and workers operate in the informal sector.

Analysis using the National Sample Surveys

As discussed in Section 4.2, we use multiple rounds of the NSS data to construct three measures of informality, namely whether a worker: (i) works on an informal contract; (ii) in an informal firm; (iii) in casual work (see Section 4.2 and Table A1 for detailed definitions). This allows us to measure worker transitions in informal or formal firms as well as on informal or payroll contracts. Similar to Equation 9, we estimate the following regression for a worker i in industry j, state s and year t:

$$Y_{it} = \beta Treat_s \times Post_t + \gamma X_{it} + \theta_s t + \alpha_{js} + \alpha_{jt} + \varepsilon_{it} \tag{10}$$

where Y_{it} are the measures of informality described above, along with (log) daily wages. We include 2-digit industry-state fixed effects (α_{js}), two-digit industry-year fixed effects (α_{jt}), and a state-time trend. Moreover, since socio-demographic characteristics of the work can influence their employment choices and wages, we also control for a range of worker characteristics (X_{it}) such as age, age-squared, gender, literacy, caste group, and marital status. We use the NSS sample weights for the estimation and cluster standard errors at the state level.

The results are reported in Table 4. From Column (1), the probability that a worker reported working on an informal contract increased by 5.5 p.p. (14.1%), s/he working in an informal firm increased by 6.2 p.p. (9.2%), and working in casual work increased by 2.7 p.p. (3.7%). Lastly, average wages fell by around 4.6% after the policy reform (Column 4). Taken together, these results imply that the policy reform, which targeted (and reduced) the use of contract workers in formal firms, also led to a decline in wages and an increase in the size of the informal sector. These are perfectly consistent with the

Table 4: Impact of the Policy on the Probability of Working in the Informal Sector and Daily Wages

	Probability of Working In			Log
	(1) Informal Contract	(2) Informal Firm	(3) Casual Work	(4) Daily Wage
Post X Treat	0.055***	0.062***	0.027***	-0.046**
	(0.009)	(0.012)	(0.009)	(0.017)
Control, Pre-Mean	0.39	0.67	0.73	3.94
R2	0.43	0.34	0.41	0.50
N	90260	144035	216558	91510

Notes: The above table uses data from all rounds of the National Sample Survey (NSS) between 1999-2005. Each outcome variable takes the value 1 if an individual reports working on an informal contract (Column 1), informal firm (Column 2), and casual work (Column 3). Daily log wages are reported in Column (4). Table A1 provides detailed definitions for all outcome variables. Post is a binary variable that takes the value 1 for years after 2003, while Treat is a binary variable that takes the value 1 for Andhra Pradesh and 0 otherwise. All regressions contain state-2 digit industry and 2-digit industry-year fixed effects along with a state-time trend. They also control for individual characteristics such as age, age-squared, gender, literacy, caste group, and marital status. Control, Pre Mean reports the mean of the outcome variable among the control states in the years before the policy reform. We use sample weights for the estimation and report robust standard errors clustered at the state level in parentheses below. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$. The sample size varies depending on the coverage of the underlying variable across rounds as well as whether a worker reports it.

key theoretical predictions.[38]

Robustness of Results to Alternate Definitions and Samples: Similar to the ASI, we now conduct multiple robustness exercises to test the validity of our results. First, in Table A7, we show that the results are very similar if we do not use sampling weights. Second, Table A8 shows that our results are robust — both in magnitude and statistical significance — to excluding self-employed individuals from the sample.

Third, our preferred specification defines those workers earning wages on a daily or weekly basis to be working on an informal contract. We try an alternate definition by classifying workers earning wages on a weekly basis to be on an informal contract. As is clear from the results in Table A9, the results are similar in magnitude and robust to this alternate definition.

Fourth, in Table A10, we extend our sample to include all five available rounds of the NSS between 1993-94 and 2009-10. Data on only two outcome variables, namely whether a worker was employed in casual work (Column 1), and their log daily wages

[38]In Appendix Table A12, we show that the reform had no impact on the probability that a worker was unemployed, or the probability of working in agriculture. The estimated coefficients are small in magnitude and statistically insignificant at conventional levels. This implies that the supply of labor to the non-agriculture sector was not affected due to the reform, a key assumption in our theoretical model.

(Column 2) is available across all these rounds. This allows to construct an event study (similar in spirit to Equation 8 in the ASI) by taking 1999-00 as the reference year. From Column (1), we see that if anything, the probability of workers working in casual work was decreasing over time in Andhra Pradesh relative to other states before the policy. On the other hand, there is an increase in this gap after the policy reform. Log daily earnings (Column 2) were statistically similar in Andhra Pradesh and other states before the policy reform, with a substantial decline in Andhra Pradesh (relative to other states) after the policy reform.

Lastly, Table A11 further expands the analysis along two dimensions: first, Panels A and C in the table restrict the sample to the manufacturing sector. This allows us to make the sample comparable to the ASI data. Panels B and D on the other hand include all non-agricultural sectors. Second, Panels A and B use the sample from our "main years" i.e., between 1999-2005. However, as noted in Table A1, some outcome variables are available for other years as well. Panels C and D therefore consider all years for which data is available for that outcome variable. We see that the results are qualitatively similar across all these samples.[39] Put together, we conclude from the above analyses that our results robustly indicate an increase in the probability of workers working in the informal sector, and earning lower wages on average.

Analysis using the Economic Census

We use two rounds of the Economic Census of India (1998 and 2005) to measure the impact of the policy on whether firms (and particularly the larger ones) are more likely to remain unregistered after the policy reform. Unlike the ASI, the EC (by definition) is a census of all non-agricultural firms in the country, thus allowing us to observe the entire distribution of firms — including the sector (formal/informal) that they operate in. More importantly for our analysis, the EC reports a firm's total employment, and whether the firm is registered across a range of government institutions (such as State Directorate of Industries, Textile/Jute Commissioner, Coir Board, etc.). The latter allows us to define a binary variable which takes the value 1 if a firm is unregistered (and thus informal) and 0 otherwise.[40] For a firm i operating in a district d, state s and industry j in year t, we estimate the following regression specification:

$$Y_{it} = \beta Treat_d \times Post_t + \alpha_t + \alpha_d + \alpha_{js} + \gamma X_{it} + \varepsilon_{it} \tag{11}$$

[39] All 16 coefficients have the expected sign, and 14 out of 16 coefficients are statistically significant at conventional levels as well.

[40] As discussed earlier (in Section 5.2), while rounds of the EC are available in 1993-94 and 2013-14, they do not report firms' registration status and hence we are unable to determine whether they are formal or informal. Hence we cannot use them for our analysis. With a lack of evidence on pre-trends, these results should therefore be interpreted with the appropriate caution.

Table 5: Impact of the Probability of Firms Remaining Unregistered

	Large	Small	All
	(1)	(2)	(3)
Post × Treat	0.276***	0.169***	0.170***
	(0.028)	(0.035)	(0.035)
Control, Pre Mean	0.84	0.97	0.97
R2	0.37	0.27	0.27
N (millions)	1.36	66.94	68.29

Notes: The above table uses data from the 1998 and 2005 rounds of the Economic Census. The outcome variable is a binary variable that takes the value 1 if a firm is unregistered and 0 otherwise. We classify firms as Large (Column 1) if they employ more than 10 workers, and Small (Column 2) if they employ less than 10 workers. Column 3 includes all firms. All regressions contain district, year, and state-2-digit industry fixed effects and control for firm characteristics like whether the firm is owned by a female, whether the owner is from a lower-caste, and whether the firm uses electrical power. Control, Pre Mean reports the mean of the outcome variable among the control states in the years before the policy reform. Robust standard errors are clustered at the state level and reported in parentheses below. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

where Y_{it} is a binary variable on if the firm is unregistered. Similar to the specifications throughout the paper, we include district, year, and state-industry fixed effects, and cluster standard errors at the state level. We also control for firm-level characteristics (X_{it}) such as whether the owner is from a low caste, female, or whether the firm uses electrical power.

Our theoretical model hypothesizes that the policy reform should make it costlier for firms–particularly the larger ones–to operate in the formal sector and thus the marginal firm (deciding between formalizing or not) would want to remain informal after the reform. On the other hand, the policy should have had a lesser impact on smaller firms that would have wanted to remain informal anyway. To test this implication, we classify firms as "small" (below 10 workers), and "large" (above 10 workers). This is a natural cutoff given the definition of formality in the Indian context (Amirapu and Gechter, 2020). That said, we show robustness of these results across multiple cutoffs in our discussion below. We separately estimate Equation 11 for each category (small, large, and all) and report the results in Table 5.

From Column (1), the probability that a large firm in Andhra Pradesh (relative to other states) remained unregistered after the policy reform, increased by around 27.5 p.p. (32.7%). On the other hand, the probability that a small firm in Andhra Pradesh (relative to other states) chose to remain unregistered increased by 17 p.p. (17.4%). This increase is only about half of the increase observed in the large firms (Column 1). Finally, since

most firms in India are small (see Table 1), we see a similar increase in the probability of remaining unregistered for all firms (Column 3). These results are consistent with the predictions of the model given that the policy reform increased the cost of operating in the formal sector.

Robustness of the results: Table A13 shows that our results are robust across the firm-size distribution i.e., robust to defining "large" firms using alternate cutoffs: from 10 workers (Column 1), 20 workers (Column 2), 30 workers (Column 3), 50 workers (Column 4), and 100 workers (Column 5). The cutoffs at 50 and 100 workers is noteworthy because additional EPL requirements in India are applicable to firms that employ workers above this threshold (Amirapu and Gechter, 2020). Moreover, the Census Sector of the ASI (as discussed in Section 4.1) covers firms employing more than 100 workers, making this cutoff relevant for our analysis. The results, if anything, grow larger across the firm-size distribution, ranging from 33-40% in Columns (1) and (2) where the cutoffs are 10 and 20 workers respectively, to 54-56% in Columns (5) and (6), where the cutoffs are 50 and 100 workers respectively. Lastly, Table A14 shows that the results are similar, both in magnitude and statistical significance, if we restrict our sample to the manufacturing sector only to make the sample comparable to the coverage in the ASI.

Measuring Entry-Exit using State-Industry Level Data from the ASI: A key limitation of using the panel data in the ASI is that it limits our understanding on the extensive margin entry-exit into the data. However, the Ministry of Statistics and Programme Implementation (Government of India) publishes annual reports that record (among other things) the total and operational number of establishments within a state-3-digit industry category. Unlike the panel data, these data are generated using a cross-section of firms in the ASI each year, thus allowing us to measure (to some extent) the extensive margin entry and exit into the formal manufacturing sector.

However, there is a key caveat — the data records the total number of "operational" establishments. The status of (non)operational is measured using a code list in the data set, where an establishment can be non-operational because of a variety of reasons such as actual closures, non-traceable owner, records with courts, or income tax department, factory under prosecution, de-registration, etc. This implies that non-operational status of a firm is strongly correlated with, but not necessarily equivalent to exit. In fact, this is exactly the argument made by Chatterjee et al. (2023), who show that firm exit in India is low, slow, and contentious.

With this caveat in mind, we collate data from these annual reports between 1999-2006 and for each state-industry (js) in year t. The total number of operational firms in js in year t (denoted by N_{jst}) will therefore be equal to $N_{jst-1}+$ New$_{jst}-$ Inactive$_{jst}$. We

Table 6: Entry-Exit into the Formal Sector

	Fraction of firms	
	Inactive	New Entrants
	(1)	(2)
Post x Treat	0.0253***	-0.0670**
	(0.00589)	(0.0307)
Control, Pre Mean	0.05	0.11
R-squared	0.175	0.084
N	7,991	7,991

Notes: Data are from the repeated cross-section of the Annual Survey of Industries between 1999-2006 that have been aggregated to the state-3-digit industry level. Column (1) reports the fraction of firms from year $t-1$ that become inactive in year t, while Column (2) reports the number of firms that enter the ASI sample in year t as a fraction of firms in $t-1$. Post is defined as 1 for years after 2003, and 0 before that. Treat is defined as 1 for Andhra Pradesh, and 0 for other states. Robust standard errors are clustered at the state level and reported in parentheses. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

therefore calculate the fraction of inactive firms ($\text{Inactive}_{jst}/N_{jst-1}$), the fraction of new entrants ($\text{New}_{jst}/N_{jst-1}$), and estimate the following specification:

$$Y_{jst} = \beta Treat_s \times Post_t + \alpha_s + \alpha_{jt} + \theta_s t + \varepsilon_{jst} \tag{12}$$

where Y_{jst} are the fraction of inactive and new firms. The fixed effects are similar to the ones discussed throughout the empirical analysis and we cluster standard errors at the state level. The results reported in Table 6 indicate an increase in the fraction of inactive formal firms by 2.5 p.p. (Column 1) and a decrease in the fraction of new entrants by 6.7 p.p. (Column 2) in Andhra Pradesh (relative to other states) after the policy reform. This is remarkably consistent in magnitude with the results from the Economic Census discussed above, increasing our confidence in the impact of the policy measured through these different data.

5.3 Summary of the Empirical Analysis

Our analysis offers three distinct insights on the impact of the policy reform on both margins of informality, which we now summarize below. First, on the intensive margin of informality, we find a large and persistent decline in the usage of contract workers and a consequent increase in the usage of payroll workers by formal firms. Second, this translates into a higher probability that workers also report working in informal enterprises as well as on informal contracts, and with lower wages. Lastly, the policy

reform results in an increase in the fraction of unregistered firms, particularly the larger ones who could have potentially formalized, but now choose to remain informal. This is also complemented by a reduced entry and greater inactivity of firms in the formal (manufacturing) sector. Put together, the policy reform does impact both margins of informality and is consistent with the insights provided by the theoretical model.

6 Model Estimation

After documenting the impact of the policy on both margins of informality, we now turn to estimating the model. This exercise is useful to be able to quantify the welfare effects of the policy, as well as examine the impact of alternate policies that aim at reducing the use of contract workers in the formal sector, such as reducing the red-tape around hiring payroll workers, for example. In Section 6.1, we discuss the assumptions that help in taking the model to the data. In Section 6.2 we provide details on estimation of the parameters, and discuss the results in Section 6.3. Lastly, in Section 6.3, we show the fit of the model with key data moments, and provide a structured discussion around the identification of model parameters.

6.1 Parameterization

We assume that Andhra Pradesh is a closed economy with one representative industry and that the pre-entry productivity for an entrepreneur, denoted by x, follows a log-normal distribution with mean 0 and variance σ_x^2 i.e., $x \sim \log N(0, \sigma_x^2)$. The post-entry productivity is then given by $z = x\varepsilon$ where $\varepsilon \sim \log N(0, \sigma_\varepsilon^2)$. The number of potential entrants, M, is calculated by subtracting the total number of entrepreneurs and workers in the formal and informal firms in Andhra Pradesh from the total adult workforce of Andhra Pradesh as reported in the 2001 Population Census of India.

The model has the following set of parameters: production function parameters, $\mathcal{P} = \{\rho, \nu, \phi, \sigma_x^2, \sigma_\varepsilon^2, \theta\}$, additional marginal costs incurred by formal firms in hiring contract and payroll workers $\mathcal{B} = \{b_c, b_r\}$, fixed costs $\mathcal{F} = \{E_i, E_r\}$, average sales tax ($t$). We set ρ to 0.738, the average labor share in output from the ASI data. We set the elasticity of substitution between contract and regular workers (ν) to be 3.[41] Sales tax $t = 3.3\%$ is the average tax paid by a firm in the ASI and ranges from 2%-6% across industries.

[41] We follow a method similar to Acemoglu, Autor and Lyle (2004) to validate this elasticity. Using the ASI data and Equation 1, we regress the log of relative wages for contract to permanent workers on the log ratio of contract workers to permanent workers. This gives us an estimate of the elasticity of substitution of 2.94.

6.2 Estimation Procedure

The 8 parameters – $\Phi = \{\{\phi, \sigma_x^2, \sigma_\varepsilon^2, \theta\}, \{b_c, b_r\}, \{E_i, E_r\}\}$, are then identified using the following 8 moments in the data:

(a) the number of firms in each sector and industry (2 moments)

(b) the average firm size in each sector and industry (2 moments)

(c) the ratio of contract to payroll workers in each industry (1 moment)

(d) the average variance in firm size in the informal and formal sector (2 moments)

(e) the coefficient on $Ln(FirmSize)^2$ of a quadratic regression of the fraction of contract workers and log firm-size (1 moment)[42]

We provide an intuition for identifying these parameters here and refer the reader to Section 6.3 for a more rigorous test of identification. The intuition for identification is as follows: given the distributional assumption on the productivity distribution, the number of firms in each industry-sector allows us to calibrate the productivity thresholds for entry into each industry-sector i.e. x_I^* and x_F^*. Given these thresholds, the variance of log firm-size in the formal and informal sectors (conditional on entry), allows us to identify $\{\sigma_x^2, \sigma_\varepsilon^2\}$. From Equation 4, the average firm size in the informal sector is a function of the penalty of operating in the informal sector (θ). Similarly, from Equation 2, the average firm size in the formal sector depends on taxes and the marginal cost of hiring contract and payroll workers (b_c and b_r respectively). Moreover, the ratio of contract workers to payroll workers (Equation 1) helps us identify b_c and b_r separately.

Using data on ASI establishments in Andhra Pradesh in 2002 (one year before the policy reform), Figure A4 shows a non-parametric fit of how the fraction of contract workers change with (log)firm-size. Consistent with Bertrand et al. (2024), we find that larger firms hire more contract workers. This elasticity — of firm productivity (and hence firm size) to ratio of contract and payroll workers — is governed by ϕ (Equations 1 and 2). Therefore, we regress the fraction of contract workers on a quadratic function of log-firm size and match the coefficient on $Ln(FirmSize)^2$ (equal to -0.010 from Column (1) of Table A15). Lastly, the fixed costs of entry are then identified from the zero-profit condition in the informal sector and the indifference condition for entering the formal sector (Equation 7).

[42] For a firm f in Andhra Pradesh in 2002, we regress: Frac. Contract Workers$_f = \alpha_j + \beta_1 Ln(FirmSize) + \beta_2 Ln(FirmSize)^2 + \varepsilon_f$. In Table A15, we report the results without any fixed effects in Column (1) and with 2-digit industry fixed effects in Column (2). We use the estimated coefficient on the quadratic term equal to -0.010 to calibrate ϕ.

We summarize the estimation algorithm here and provide detailed steps in Appendix C2. For a given set of parameters Φ, we can use iterative mapping to find the price vector $\mathbf{p} = \{w, p\}$ that clears the goods and labor markets. We can then simulate the set of moments (a)-(d) above (denoted by $M^{Sim}(\Phi)$) and use Simulated Method of Moments (McFadden, 1989; Duffie and Singleton, 1993) to choose the parameter values Φ that minimize the distance between the simulated moments and their data counterparts (denoted by M^{Data}).

6.3 Parameter Estimates and Model Fit

Parameter Estimates: We estimate the variance of the productivity distribution, $\sigma_x = 0.21$ and $\sigma_\varepsilon = 0.22$. The penalty parameter for operating in the informal sector (θ) is estimated to 1.15. The elasticity of how the fraction of contract workers changes with firm productivity (ϕ) is estimated 0.91. $\{b_c, b_r\}$ are estimated to be $\{0.91, 0.57\}$, which implies that conditional on productivity (z), it is 1.6 times more expensive to hire a worker on contract (as compared to on payroll) in the formal sector.

Model Fit: Section 6.3 shows the fit of the model to key targeted and untargeted moments in the data. First, Column 1 of Table A16 reports the simulated moments from the model, while Column 2 reports the moments from the data. As can be seen from the table, the model matches the key targeted moments such as the number and size of formal and informal firms as well as variance in the firm-size distribution in each sector very well. We also find that the model has a good fit of the non-targeted moments like the fraction of labor and firms in the informal sector.[43] Lastly, even though we parameterize a quadratic relationship between the fraction of contract workers and firm size and match one regression coefficient, our model is in fact able to match the non-parametric relationship very well (dash-dotted blue line in Figure A4), increasing our confidence in the model being able to replicate key data patterns.

Identification: Section 6.2 provides heuristic arguments of how various data moments help identify the key parameters of the model. However, we now adopt a more systematic approach for establishing identification in the spirit of Kaboski and Townsend (2011), Bick et al. (2021) and Chiplunkar and Goldberg (2023). Specifically, for the key model parameters, namely: distortions in hiring contract and payroll workers (b_c, b_r), produc-

[43] One place to note is that the cost of hiring contract workers is decreasing with firm size in the model. This implies that larger firms (in the model) will hire more contract workers. However, as we will show in Figure A4, the fraction of contract workers plateaus with firm-size. This implies that the model will generate on average firms with larger size (employing more labor) in the formal sector than what we empirically observe.

tivity distribution (σ_x and σ_ε), penalty of operating in the informal sector (θ), and the elasticity of hiring contract workers with firm productivity (ϕ), we compute the derivative of a moment with respect to each parameter.[44] To do so, we re-solve the model each time by increasing one parameter by 1% above its estimated value (keeping all others the same) and compute the resulting percent changes in the simulated moments. We report the results in Table A17. Each number in a row r and column c reports the percentage change in the moment in row r (averaged across industries) when the parameter in column c is increased by 1% (keeping all other parameters the same). As the table shows, the results are consistent with the discussion in Section 6.2. From Columns 1 and 2, we see that the variance in firm-size in the informal and formal sector is most sensitive to the changes in σ_x and σ_ε respectively. Conditional on this, changes in the average firm size in the informal sector is most sensitive to a change in θ (Column 3), and changes in the coefficient on the square of (log) firm-size is most sensitive to ϕ (Column 4) Lastly, a change in b_c and b_r affects the average size of a formal firm, and the fraction of contract workers (Columns 5 and 6).

7 Impact of Labor Law Reforms

We now turn to discussing the impact of counterfactual EPL policies on informality and welfare. We begin with the policy reform in Andhra Pradesh, where formal sector firms are penalized for hiring contract workers (Section 7.1. We implement the policy so that the decline in the fraction of contract workers in the simulated model corresponds to the decline that we empirically observe in Table 3 (around 18%). In Section 7.2, we then implement another counterfactual reform where we obtain the same reduction in the fraction of contract workers, but instead by reducing the costs incurred by formal firms for hiring workers on payroll (such as red-tape around hiring/firing payroll workers, for example).

7.1 Penalizing Hiring Contract Workers

Consider the policy reform implemented in Andhra Pradesh, that of penalizing a firm in the formal sector for hiring workers "off-the-books". Through the lens of the model, this would correspond to increasing b_c. For our counterfactual simulation, we increase b_c by 19% to be able to empirically match the empirically observed decline in the fraction of contract workers (Table 3) and then examine the impact of the policy on various other

[44] Note that the fixed costs of entry and formalization are not identified directly from a data moment, but computed from the the zero-profit conditions. Accordingly, we do not consider them here.

Table 7: Impact of Employment Protection Policies

	Baseline	Penalizing Contract Workers	Subsidizing Payroll Workers
		$\triangle b_c = 19\%$	$\triangle b_r = -12\%$
	(1)	(2)	(3)
Policy change		1.19	0.88
Panel A: Intensive Margin of Informality			
Frac. Contract Workers	1.00	0.82	0.82
Panel B: Extensive Margin of Informality			
Frac. Informal Firms	1.00	1.005	0.996
Frac. Informal Labor	1.00	1.29	0.95
Panel C: Productivity of the Marginal Firm			
Informal Sector	1.00	0.97	1.01
Formal Sector	1.59	1.64	1.55
Panel D: Aggregate Effects			
TFP	1.00	0.98	1.004
Real Wages	1.00	0.95	1.01
Real Income (Welfare)	1.00	0.91	1.08

Notes: Notes: The above table reports the impact of two labor laws on the margins of informality. The baseline values in Column (1) have been normalized to 1, except for the productivity of the marginal firm in the formal sector, which in the baseline has been normalized to 1 for the productivity of the marginal firm in the informal sector. Columns (2) and (3) report the impact of a 19% (12%) increase (decrease) in b_c (b_r) relative to its baseline value. The change in b_r has been calibrated to achieve the same reduction in the fraction of contract workers as in Columns (2).

outcomes of interest. We report the results in Table 7. Column (1) reports the status-quo baseline (in Column 1), which we normalize to 1 for ease of comparison across counterfactual simulations. Starting with the intensive margin of informality (Panel A in Column 2), a 19% increase in b_c (by construction) results in a 18% decline in the fraction of contract workers. Turning to the extensive margin (Panel B), we see that the fraction of firms and workers in the informal sector increase by 0.5% and 29% respectively as compared to the baseline. This is consistent with a 3 % (3.1%) decrease (increase) in the productivity of the marginal firm who is deciding to enter the informal (formal) sectors (Panel C). Lastly, aggregate TFP decreases by 2%, with a 5% and 9% decline in real wages and income (Panel D).

As discussed earlier, the underlying mechanisms are intuitive in the framework of the

model. A policy that penalizes hiring contract workers in the formal sector increases the marginal costs for a firm to operate in the formal sector. The direct effect of this policy (keeping wages and prices fixed) implies that a firm now has to be even more productive to operate in the formal sector, which results in a higher threshold productivity of formalizing (in Panel C). This implies that firms sort to operate in the informal sector (Panel B), which lowers TFP as well as aggregate labor demand, and therefore real wages (Panel D). Although this indirect channel of lower real wages (in general equilibrium) mitigates the impact of the direct effect in the formal sector, it does not offset it, resulting in lower informality on the intensive margin, but higher informality on the extensive one.

7.2 Reducing Costs of Hiring Payroll Workers

An alternate policy advocated by entrepreneurs, and supported in the academic literature as well (Ahsan and Pagés, 2009; Besley and Burgess, 2004; Chaurey, 2015; Sodhi, 2014), is the simplification of complex and cumbersome employment protection laws (such as government approvals for hiring and firing workers, long notice periods, etc.). Through the lens of our model, such policies would reduce the costs for hiring payroll workers i.e., lower b_r. We decrease b_r in such a way so that its impact on the intensive margin of informality (fraction of contract workers) is the same as in the previous case (Section 7.1). This allows us to make the impact of the two policies comparable on the intensive margin–a key objective of EPL reforms. The results, reported in Column (3) of Table 7, imply a 12% decrease in b_r.

The impact of this policy, by construction, is the same on the intensive margin of informality (Panel A) as compared to the previous one (Section 7.1). However, the effects on the extensive margin of informality (Panel B) are in sharp contrast. While the previous policy (Section 7.1) increased the extensive margin of informality with more firms and workers in the informal sector, this policy results in a reduction in the size of the informal sector. The fraction of firms and workers in the informal sector decrease by 0.4% and 5% respectively relative to the baseline. Since the marginal firm now wants to formalize, the productivity of the marginal firm in the formal sector now decreases by 2.5%, while that in the informal sector increases by 1% (Panel C). Since firms are more productive in the formal sector, overall TFP marginally increases by 0.4% (Panel D). Lastly, real wages and income increase by 1% and 8% respectively.

The above policy simulations highlight the importance of effectively targeting EPL reforms, and its impact on different margins of informality. First, penalizing non-compliance with EPL (as was the policy implemented in Andhra Pradesh) does reduce informality on the intensive margin, but increases informality on the extensive margin. Moreover, it lowers aggregate TFP, real wages and income as well. On the other hand, policy reforms

that reduce the burden of hiring payroll workers reduce informality on both margins, and increase TFP, real wages, and welfare.

7.3 Endogenous Labor Supply Responses

Our model assumes a fixed labor supply i.e., a perfectly inelastic labor supply curve. However, it is possible that the policy reform induces responses in labor supply as well in three ways: (i) increasing participation of individuals into the labor force (LFP); (ii) allowing unemployed individuals to find jobs; (iii) reallocating individuals from agriculture into non-agricultural sectors. We use the NSS data to measure each of these three dimensions and examine the impact of the policy on them. As reported in Table A12, we find no effect of the policy on either on the probability of participating in the labor force (Column 1), probability of unemployment (Column 2), or being employed in agriculture (Column 3). Both the magnitude is small and the coefficients are not statistically significant at conventional levels, indicating that this is not a meaningful dimension of adjustment in our context.

8 Conclusion

Low compliance with EPL remains an important policy challenge in developing countries. In this paper, we show, both theoretically and empirically, that the presence of an informal sector plays a key role in understanding this. We study the impact of a unique natural experiment in India the penalized the use of contract workers in the formal sector. A key insight from our analysis is that while the reform increased EPL compliance in the formal sector i.e., firms hired fewer contract workers, it also made them more likely to operate in the informal sector all together, lowering wages, aggregate productivity, and welfare. On the contrary, we argue that designing and targeting EPL, especially around lowering bureaucratic costs and constraints in hiring workers on payroll can meaningfully reduce both margins of informality and increase welfare. More broadly, as the world continues to grapple with the aftermath of the Covid-19 pandemic, and automation and globalization continue to redefine the nature of work and the relationship between workers and firms, defining and enforcing employment protection legislation–particularly in contexts with high informality–remains an important challenge. These regulations should be carefully targeted to increase the productivity of the workforce and the overall welfare of the economy.

References

Acemoglu, Daron, David H Autor, and David Lyle, "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Mid-century," *Journal of Political Economy*, 2004, *112* (3), 497–551.

Aguirregabiria, Victor and CESAR Alonso-Borrego, "Labor Contracts and Flexibility: Evidence from a Labor Market Reform in Spain," *Economic Inquiry*, 2014, *52* (2), 930–957.

Ahsan, Ahmad and Carmen Pagés, "Are All Labor Regulations Equal? Evidence from Indian Manufacturing," *Journal of Comparative Economics*, 2009, *37* (1), 62–75.

Almeida, Rita K, Lourenço S Paz, and Jennifer P Poole, "Precarization or Protection? The Role of Trade and Labour policies on Informality," *The Journal of Development Studies*, 2022, *58* (7), 1416–1435.

Amirapu, Amrit and Michael Gechter, "Labor Regulations and the Cost of Corruption: Evidence from the Indian Firm Size Distribution," *Review of Economics and Statistics*, 2020, *102* (1), 34–48.

Ardito, Chiara, Fabio Berton, Lia Pacelli, and Filippo Passerini, "Employment Protection, Workforce Mix and Firm Performance," 2021.

Autor, David H, John J Donohue III, and Stewart J Schwab, "The Costs of Wrongful Discharge Laws," *The Review of Economics and Statistics*, 2006, *88* (2), 211–231.

___, William R Kerr, and Adriana D Kugler, "Does Employment Protection Reduce Productivity? Evidence from US States," *The Economic Journal*, 2007, *117* (521), F189–F217.

Bachas, Pierre, Roberto N Fattal Jaef, and Anders Jensen, "Size-dependent Tax Enforcement and Compliance: Global Evidence and Aggregate Implications," *Journal of Development Economics*, 2019, *140*, 203–222.

Baek, Jisun and WooRam Park, "Firms Adjustments to Employment Protection Legislation: Evidence from South Korea," *ILR Review*, 2018, *71* (3), 733–759.

Beck, Thorsten and Mohammad Hoseini, "Informality and Access to Finance: Evidence from India," *Center Discussion Paper Series*, 2014.

Bertrand, Marianne, Chang-Tai Hsieh, and Nick Tsivanidis, "Contract Labor and Establishment Growth in India," Technical Report, Conditionally Accepted, Econometrica 2024.

Besley, Timothy and Robin Burgess, "Can Labor Regulation Hinder Economic Performance? Evidence from India," *The Quarterly Journal of Economics*, 2004, *119* (1), 91–134.

Bhattacharjea, Aditya, "Labour Market Regulation and Industrial Performance in India: A Critical Review of the Empirical Evidence," *Indian Journal of Labour Economics*, 2006, *49* (2), 211–232.

Bick, Alexander, Nicola Fuchs-Schündeln, David Lagakos, and Hitoshi Tsujiyama, "Structural Change in Labor Supply and Cross-Country Differences in Hours Worked," Technical Report, National Bureau of Economic Research 2021.

Bjuggren, Carl Magnus, "Employment Protection and Labor Productivity," *Journal of Public Economics*, 2018, *157*, 138–157.

Boeri, Tito and Pietro Garibaldi, "Two Tier Reforms of Employment Protection: A Honeymoon Effect?," *The Economic Journal*, 2007, *117* (521), F357–F385.

Bornhäll, Anders, Sven-Olov Daunfeldt, and Niklas Rudholm, "Employment Protection Legislation and Firm Growth: Evidence from a Natural Experiment," *Industrial and Corporate Change*, 2015, p. dtw017.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting event-study designs: robust and efficient estimation," *Review of Economic Studies*, 2024, p. rdae007.

Botero, Juan C, Simeon Djankov, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "The Regulation of Labor," *The Quarterly Journal of Economics*, 2004, *119* (4), 1339–1382.

Bruhn, Miriam, "A Tale of Two Species: Revisiting the Effect of Registration Reform on Informal Business Owners in Mexico," *Journal of Development Economics*, 2013, *103*, 275–283.

___ **and David McKenzie**, "Entry Regulation and the Formalization of Microenterprises in Developing Countries," *The World Bank Research Observer*, 2014, *29* (2), 186–201.

Butschek, Sebastian and Jan Sauermann, "The Effect of Employment Protection on Firms Worker Selection," *Journal of Human Resources*, 2022, pp. 0919–10433R1.

Cahuc, Pierre and Fabien Postel-Vinay, "Temporary Jobs, Employment Protection and Labor Market Performance," *Labour Economics*, 2002, *9* (1), 63–91.

___ **, Olivier Charlot, and Franck Malherbet**, "Explaining the Spread of Temporary Jobs and It's Impact on Labor Turnover," *International Economic Review*, 2016, *57* (2), 533–572.

___ **, Pauline Carry, Franck Malherbet, and Pedro S Martins**, "Spillover Effects of Employment Protection," *Nova SBE Working Paper Series*, 2023, (655).

Callaway, Brantly and Pedro H.C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2020.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (9), 2964–96.

Chakraborty, Pavel, Rahul Singh, and Vidhya Soundararajan, "Import Competition, Formalization, and the Role of Contract Labor," *IIM Bangalore Research Paper*, 2020, (633).

Chatterjee, Shoumitro, Kalyani Padmakumar, Kala Krishna, and Yingyan Zhao, "No Country for Dying Firms: Evidence from India," 2023.

Chaurey, Ritam, "Labor Regulations and Contract Labor Use: Evidence from Indian Firms," *Journal of Development Economics*, 2015, *114*, 224–232.

Chiplunkar, Gaurav and Pinelopi K Goldberg, "Aggregate Implications of Barriers to Female Entrepreneurship," 2023.

Cingano, Federico, Marco Leonardi, Julian Messina, and Giovanni Pica, "The Effects of Employment Protection Legislation and Financial Market Imperfections on Investment: Evidence from a Firm-level Panel of EU Countries," *Economic Policy*, 2010, *25* (61), 117–163.

Cisneros-Acevedo, Camila, "Unfolding Trade Effect in Two Margins of Informality. The Peruvian Case," *The World Bank Economic Review*, 2022, *36* (1), 141–170.

Daruich, Diego, Sabrina Di Addario, and Raffaele Saggio, "The Effects of Partial Employment Protection Reforms: Evidence from Italy," *The Review of Economic Studies*, 02 2023. rdad012.

de la Parra, Brenda Samaniego and León Fernández Bujanda, "Increasing the Cost of Informal Workers: Evidence from Mexico," Technical Report, Working Papers 2020.

Dix-Carneiro, Rafael, Pinelopi K Goldberg, Costas Meghir, and Gabriel Ulyssea, "Trade and Informality in the Presence of Labor Market Frictions and Regulations," Technical Report, National Bureau of Economic Research 2021.

Duffie, Darrell and Kenneth J Singleton, "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica: Journal of the Econometric Society*, 1993.

Elgin, Ceyhun, M Ayhan Kose, Franziska Ohnsorge, and Shu Yu, "Understanding Informality," *CAMA Working Paper*, 2021.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," Technical Report, National Bureau of Economic Research 2018.

Gourinchas, Pierre-Olivier and Jonathan A Parker, "Consumption Over the Life Cycle," *Econometrica*, 2002, *70* (1), 47–89.

Hijzen, Alexander, Leopoldo Mondauto, and Stefano Scarpetta, "The Impact of Employment Protection on Temporary Employment: Evidence from a Regression Discontinuity Design," *Labour Economics*, 2017, *46*, 64–76.

Kaboski, Joseph P and Robert M Townsend, "A Structural Evaluation of a Large-scale Quasi-experimental Microfinance Initiative," *Econometrica*, 2011, *79* (5), 1357–1406.

Karak, Anirban and Deepankar Basu, "Profitability or Industrial Relations: What Explains Manufacturing Performance Across Indian States?," *Development and Change*, 2020, *51* (3), 817–842.

Kugler, Adriana and Giovanni Pica, "Effects of Employment Protection on Worker and Job Flows: Evidence from the 1990 Italian Reform," *Labour Economics*, 2008, *15* (1), 78–95.

McFadden, Daniel, "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica*, 1989, pp. 995–1026.

Melitz, Marc J, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Narayanan, CK et al., "A Firm'No'to Labour Reforms," 2005.

Nikaido, Yuko, Jesim Pais, and Mandira Sarma, "What Hinders and What Enhances Small Enterprises Access to Formal Credit in India?," *Review of Development Finance*, 2015, *5* (1), 43–52.

Ponczek, Vladimir and Gabriel Ulyssea, "Enforcement of Labour Regulation and the Labour Market Effects of Trade: Evidence from Brazil," *The Economic Journal*, 2022, *132* (641), 361–390.

Porta, Rafael La and Andrei Shleifer, "The Unofficial Economy and Economic Development," *Brookings Papers on Economic Activity*, 2008, *2008* (2), 275–363.

___ and ___, "Informality and Development," *Journal of Economic Perspectives*, 2014, *28* (3), 109–26.

Saha, Bibhas, Kunal Sen, and Dibyendu Maiti, "Trade Openness, Labour Institutions and Flexibilisation: Theory and Evidence from India," *Labour Economics*, 2013, *24*, 180–195.

Schivardi, Fabiano and Roberto Torrini, "Identifying the Effects of Firing Restrictions Through Size-contingent Differences in Regulation," *Labour Economics*, 2008, *15* (3), 482–511.

Sodhi, JS, "Labour Law Reform in India," *The Indian Journal of Industrial Relations*, 2014, pp. 100–117.

TeamLease, "India Labour Report: A Ranking of Indian States by their Labour Ecosystem," *A Report*, 2006.

Ulyssea, Gabriel, "Firms, Informality, and Development: Theory and Evidence from Brazil," *American Economic Review*, 2018, *108* (8), 2015–47.

Vera, Celia P, Bruno Jiménez et al., "Do Immigrants Take or Create Natives' Jobs? Evidence of Venezuelan Immigration in Peru," Technical Report 2022.

APPENDIX

A Figures and Tables

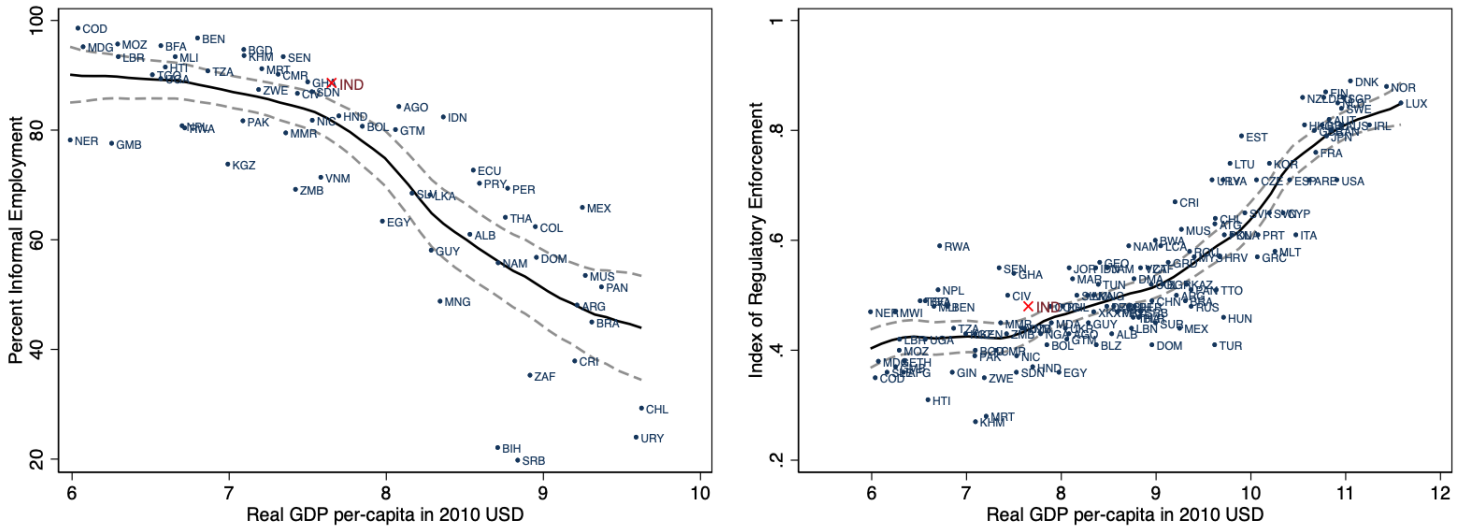
A.1 Data Definitions and Coverage Over Time

Table A1: Definition of Variables, Coverage of Data Over Time

	Variable	Coverage	Definition
	(1)	(2)	(3)
ASI	Contract workers	1999-2009	Number of contract workers employed by the firm
	Frac. contract workers	1999-2009	Contract Workers/Total Workers employed in the firm
	Regular workers	1999-2009	Number of regular workers employed in the firm
	Contract worker-days	1999-2009	Person-days worked by contract workers
	Regular worker-days	1999-2009	Person-days worked by regular workers
	Total workers	1999-2009	Total number of workers employed by the firm
NSS	Informal worker	1999, 2004, 2009	Paid on a daily or weekly basis; provide unpaid labor in a household enterprise
	Informal firm	1999, 2004, 2009	Working in a firm less than 10 workers, less than 20 workers without electricity; working in an owner-operated enterprise; or provide unpaid labor in a household enterprise
	Casual worker	1993, 1999, 2003, 2004, 2009	Employed in casual work, working in an owner-operated enterprise; or provide unpaid labor in a household enterprise
	Daily wages	1993, 1999, 2003, 2004, 2009	Daily wages earned in INR, calculated as the ratio of total weekly earnings divided by the number of days worked
EC	Firm size	1991, 1999, 2005, 2013	Number of workers employed in a firm
	Registration status	1999, 2005	Whether a firm is registered with any government agency or not

A.2 Descriptive Figures

Figure A1: Fraction of Informal Workers and EPL Enforcement

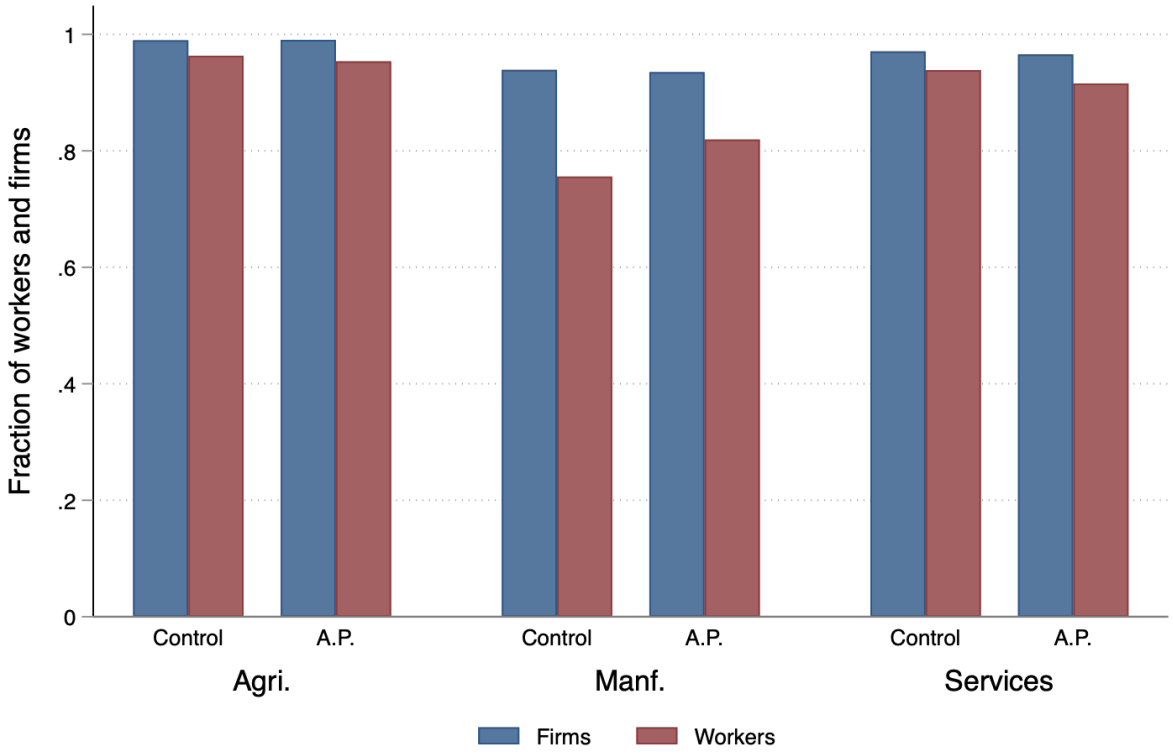


(a) Fraction of Informal Workers

(b) EPL Enforcement

Note: Data on informality is taken from Elgin, Kose, Ohnsorge and Yu (2021). Data on enforcement of labor regulation is taken from Botero, Djankov, Porta, Lopez-de Silanes and Shleifer (2004). We use 2018 values across all countries. Data on Real GDP per-capita is taken from the World Bank.

Figure A2: Fraction of Workers and Firms in the Informal Sector



Note: The above figure uses data from the 1998 round of the Economic Census. The blue bars plot the fraction of unregistered firms in each industry, while the red bars are the fraction of workers who work in them. A.P. = Andhra Pradesh; Control = Other states in India excluding Andhra Pradesh.

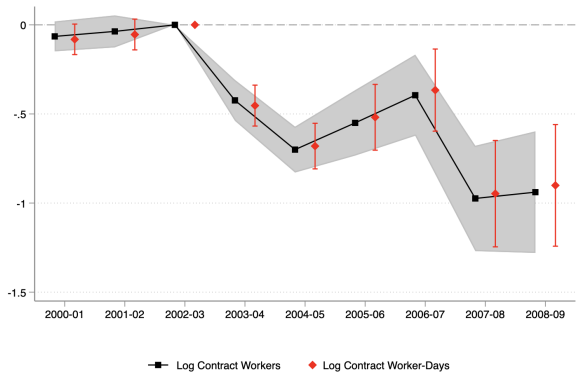
A.3 Additional Results from the Annual Survey of Industries

Table A2: Event Study Regression

	Log Contract		Log Regular		Frac. Contract	
	(1) Workers	(2) Worker-Days	(3) Workers	(4) Worker-Days	(5) Workers	(6) Worker-Days
2000-2001	-0.009	-0.003	0.033	0.039	-0.002	-0.000
	(0.055)	(0.057)	(0.029)	(0.028)	(0.006)	(0.006)
2001-2002	0.026	0.010	-0.005	-0.022	-0.007*	-0.006
	(0.042)	(0.041)	(0.026)	(0.027)	(0.004)	(0.004)
2003-2004	-0.343***	-0.382***	0.192***	0.152***	-0.033***	-0.033***
	(0.061)	(0.062)	(0.028)	(0.027)	(0.005)	(0.005)
2004-2005	-0.708***	-0.696***	0.270***	0.282***	-0.055***	-0.057***
	(0.080)	(0.080)	(0.043)	(0.044)	(0.008)	(0.008)
2005-2006	-0.643***	-0.629***	0.197***	0.211***	-0.042***	-0.043***
	(0.117)	(0.119)	(0.059)	(0.061)	(0.011)	(0.011)
p-val:	0.74	0.96	0.38	0.11	0.12	0.11
R2	0.80	0.79	0.85	0.86	0.81	0.81
N	161552	161552	161552	161552	161552	161552

Notes: Data are from the Annual Survey of Industries between 2000-2007. Columns (1)-(6) report underlying coefficients in Figure 2 for sub-figures (a)-(c) respectively. The coefficient for 2002-2003, the year before the reform, has been normalized to zero. The outcome variables are winsorized at the 1% levels at the top and the bottom of the distribution. Following Borusyak et al. (2024), the p-value of a F-test that jointly tests for the pre-policy $\widehat{\beta}_t$ to be equal to 0 are reported. Robust standard errors clustered at the state level in parentheses; *** is $p < 0.01$, ** is $p < 0.05$ and * is $p < 0.1$.

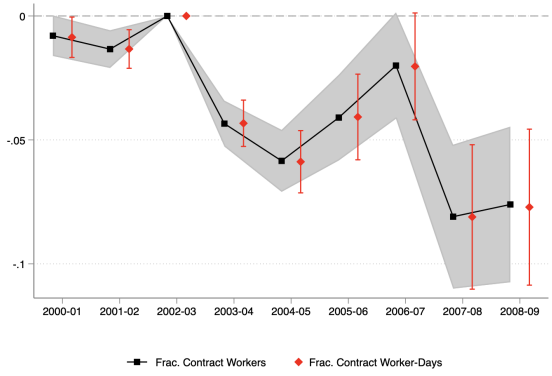
Figure A3: Impact of the Policy on the Use of Contract and Regular Workers



(a) Log Contract Workers



(b) Log Regular Workers



(c) Frac. Contract Workers

Notes: The above graphs plot the regression coefficients from a difference-in-differences specification using data from the Annual Survey of Industries between 1999-2009. The black squares report the results for workers, while the red diamonds report the results for worker-days. 90% confidence intervals, clustered at the state level are indicated around the point estimate. The coefficient for 2002-2003, the year before the reform, has been normalized to zero. The outcome variables are winsorized at the 1% levels at the top and the bottom of the distribution.

Table A3: Impact on Core and Non-Core Activities

	Log Contract Workers	Log Regular Workers	Frac. Contract Workers
	(1)	(2)	(3)
Panel A: Impact on Core Activities			
Post X Treat	-0.41***	0.14***	-0.03***
	(0.05)	(0.03)	(0.00)
R2	0.79	0.85	0.81
Panel B: Impact on Non-Core Activities			
Post X Treat	-0.01	-0.09***	0.00
	(0.01)	(0.03)	(0.00)
R2	0.61	0.80	0.62
N	165676	165676	165676

Notes: Data are from the Annual Survey of Industries between 2000-2007. Core activity (Panel A) is any activity for which the establishment is set up, and other activities essential for these core activities. Non-core activities (Panel B) are the remaining peripheral activities (listed in the Appendix Section B). Column (1)-(3) reports the results for contract workers, regular workers, and the fraction of contract workers respectively. Frac. Contract is the number of contract workers (or worker-days) as a fraction of total workers (worker-days) in a firm. Post is defined as 1 for years after 2003, and 0 before that. Treat is defined as 1 for Andhra Pradesh, and 0 for other states. Source of the data is the Annual Survey of Industries between 2000-2007. The outcome variables are winsorized from the top and bottom at 1%. Robust standard errors are clustered at the state level and reported in parentheses. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

Table A4: Results After Restricting the Control States to Neighboring States Only

	Workers		Worker-Days	
	(1) Log Contract	(2) Frac. Contract	(3) Log Contract	(4) Frac. Contract
Panel A: Impact on Usage of Contract Workers				
Post X Treat	-0.46***	-0.04***	-0.48***	-0.04***
	(0.09)	(0.01)	(0.09)	(0.01)
R2	0.78	0.79	0.78	0.79
	Log Contract	Frac. Contract	Log Contract	Frac. Contract
Panel B: Impact on Payroll Workers and Firm Size				
Post X Treat	0.22***	0.05*	0.20***	0.03
	(0.03)	(0.02)	(0.03)	(0.02)
R2	0.85	0.94	0.85	0.93
N	67046	67046	67046	67046

Notes: The above table uses data from the Annual Survey of Industries between 2000-2007. The sample is constrained to neighboring states of Andhra Pradesh, which are Chhattisgarh, Karnataka, Maharashtra, and Odisha. It reports the impact of the policy reform on workers in Columns (1) and (2) and worker-Days in Columns (3) and (4). Panel A reports the impact of the policy on the usage of contract workers, while Panel B reports the impact on the usage of regular payroll workers and total workers. Frac. Contract is the number of contract workers (or worker-days) as a fraction of total workers (worker-days) in a firm. Post is defined as 1 for years after 2003, and 0 before that. Treat is defined as 1 for Andhra Pradesh, and 0 for other states. Source of the data is the Annual Survey of Industries between 2000-2007. The outcome variables are winsorized from the top and bottom at 1%. Robust standard errors are clustered at the state level and reported in parentheses. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

Table A5: Results After Excluding Neighboring States

	Workers		Worker-Days	
	(1) Log Contract	(2) Frac. Contract	(3) Log Contract	(4) Frac. Contract
Panel A: Impact on Usage of Contract Workers				
Post X Treat	-0.31***	-0.03***	-0.34***	-0.03***
	(0.08)	(0.01)	(0.08)	(0.01)
R2	0.80	0.81	0.79	0.81
	Log Contract	Frac. Contract	Log Contract	Frac. Contract
Panel B: Impact on Payroll Workers and Firm Size				
Post X Treat	0.20***	0.06**	0.17***	0.03
	(0.03)	(0.02)	(0.03)	(0.02)
R2	0.84	0.94	0.85	0.93
N	105692	105692	105692	105692

Notes: The above table uses data from the Annual Survey of Industries between 2000-2007. The sample excludes neighboring states of Andhra Pradesh, which are Chhattisgarh, Karnataka, Maharashtra, and Odisha. It reports the impact of the policy reform on workers in Columns (1) and (2) and worker-Days in Columns (3) and (4). Panel A reports the impact of the policy on the usage of contract workers, while Panel B reports the impact on the usage of regular payroll workers and total workers. Frac. Contract is the number of contract workers (or worker-days) as a fraction of total workers (worker-days) in a firm. Post is defined as 1 for years after 2003, and 0 before that. Treat is defined as 1 for Andhra Pradesh, and 0 for other states. Source of the data is the Annual Survey of Industries between 2000-2007. The outcome variables are winsorized from the top and bottom at 1%. Robust standard errors are clustered at the state level and reported in parentheses. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

Table A6: Robustness of ASI Results to Alternate Fixed Effects

	Log Contract		Log Regular		Fraction of Contract	
	(1) Workers	(2) Worker-Days	(3) Workers	(4) Worker-Days	(5) Workers	(6) Worker-Days
Panel A: Year Fixed Effects						
Post X Treat	-0.51***	-0.50***	0.17***	0.17***	-0.05***	-0.05***
	(0.04)	(0.04)	(0.03)	(0.03)	(0.00)	(0.00)
R2	0.04	0.04	0.04	0.04	0.04	0.04
Panel B: Industry-Year Fixed Effects						
Post X Treat	-0.44***	-0.44***	0.21***	0.21***	-0.05***	-0.05***
	(0.08)	(0.08)	(0.03)	(0.02)	(0.01)	(0.01)
R2	0.08	0.08	0.09	0.10	0.09	0.09
Panel C: State-Industry Fixed Effects						
Post X Treat	-0.55***	-0.53***	0.20***	0.21***	-0.05***	-0.05***
	(0.05)	(0.05)	(0.03)	(0.03)	(0.00)	(0.00)
R2	0.15	0.14	0.14	0.15	0.16	0.16
Panel D: Industry-Year and State-Industry Fixed Effects						
Post X Treat	-0.47***	-0.45***	0.22***	0.22***	-0.05***	-0.05***
	(0.08)	(0.08)	(0.03)	(0.03)	(0.01)	(0.01)
R2	0.15	0.14	0.14	0.15	0.16	0.16
Panel E: Preferred Specification in the Main Paper						
Post X Treat	-0.38***	-0.40***	0.21***	0.17***	-0.03***	-0.03***
	(0.06)	(0.06)	(0.03)	(0.03)	(0.00)	(0.00)
R2	0.80	0.79	0.85	0.86	0.81	0.81
N	161552	161666	161552	161666	161552	161666

Notes: The above table uses data from the Annual Survey of Industries between 2000-2007. The outcome variables in Columns (1)-(2) are log contract worker and worker-days respectively; Columns (3)-(4) are log regular workers and worker-days respectively, and Columns (5)-(6) are the fraction of contract workers and worker-days respectively. Regressions in Panel A include year fixed effects and a state-time trend. Panels B and C additionally include 2-digit industry-year fixed effects, and state-2-digit industry fixed effects respectively. Panel D includes both state-2-digit-industry and 2-digit industry-year fixed effects. Lastly, Panel E includes firm fixed effects and 2-digit industry-year fixed effects, as in the specification in the paper. Post is defined as 1 for years after 2003, and 0 before that. Treat is defined as 1 for Andhra Pradesh, and 0 for other states. The outcome variables are winsorized from the top and bottom at 1%. Robust standard errors are clustered at the state level and reported in parentheses. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

A.4 Additional Tables from the National Sample Surveys

Table A7: Robustness of Results to Not Using Sampling Weights

	Probability of Working In			Log
	(1) Informal Contract	(2) Informal Firm	(3) Casual Work	(4) Daily Wage
Post x Treat	0.024***	0.037***	0.014**	-0.092***
	(0.008)	(0.011)	(0.007)	(0.016)
Control, Pre Mean	0.39	0.67	0.59	4.60
R2	0.43	0.41	0.44	0.52
N	90260	163192	216558	91510

Notes: The above table uses data from all rounds of the National Sample Survey between 1999-2005. Each outcome variable takes the value 1 if an individual reports working on an informal contract (Column 1), informal firm (Column 2), and casual work (Column 3). Daily log wages are reported in Column (4). Post is a binary variable that takes the value 1 for years after 2003, while Treat is a binary variable that takes the value 1 for Andhra Pradesh and 0 otherwise. All regressions contain state-2 digit industry and 2-digit industry-year fixed effects along with a state-time trend. Control, Pre Mean reports the mean of the outcome variable among the control states in the years before the policy reform. Robust standard errors are clustered at the state level and reported in parentheses below. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$. The sample size varies depending on the coverage of the underlying variable across rounds as well as whether a worker reports it. We do not use sampling weights.

Table A8: Excluding Self-Employed Individuals from the NSS Analysis

	Probability of Working In		
	(1) Informal Contract	(2) Informal Firm	(3) Casual Work
Panel A: Including Self-Employed Individuals			
Post x Treat	0.055***	0.045***	0.027***
	(0.009)	(0.014)	(0.009)
Control, Pre Mean	0.39	0.86	0.73
R2	0.43	0.39	0.41
N	90260	163192	216558
Panel B: Excluding Self-Employed Individuals			
Post x Treat	0.051***	0.057***	0.027**
	(0.009)	(0.017)	(0.011)
Control, Pre Mean	0.48	0.77	0.62
R2	0.44	0.41	0.47
N	87823	87951	140173

Notes: The above table uses data from all rounds of the National Sample Survey (NSS) between 1999-2005. Each outcome variable takes the value 1 if an individual reports working on an informal contract (Column 1), informal firm (Column 2), and casual work (Column 3). Table A1 provides detailed definitions for each outcome variable. Panel A includes self-employed individuals, who are classified as informal firms, and casual workers. Panel B excludes them. Post is a binary variable that takes the value 1 for years after 2003, while Treat is a binary variable that takes the value 1 for Andhra Pradesh and 0 otherwise. All regressions contain state-2 digit industry and 2-digit industry-year fixed effects along with a state-time trend. Control, Pre Mean reports the mean of the outcome variable among the control states in the years before the policy reform. Robust standard errors are clustered at the state level and reported in parentheses below. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$. The sample size varies depending on the coverage of the underlying variable across rounds as well as whether a worker reports it. The regressions include the NSS sample weights.

Table A9: Alternate Definitions of Informal Workers

	Prob. of Working on Informal Contract			
	(1) Daily/Weekly	(2) Daily	(3) Daily/Weekly	(4) Daily
Post x Treat	0.055***	0.058***	0.113***	0.106***
	(0.009)	(0.012)	(0.012)	(0.012)
Sample	Main	Main	All	All
Control, Pre Mean	0.65	0.57	0.65	0.57
R2	0.43	0.32	0.48	0.37
N	90260	90260	150989	150989

Notes: The above table uses data from multiple rounds of the National Sample Survey. Columns (1) and (2) restrict the sample to the main sample years between 1999-2005, whereas Columns (3) and (4) include all available years. The outcome variable takes the value 1 if an individual reports working on an informal contract, which is defined to be wages received on a daily or weekly basis in Columns (1) and (3), and daily basis in Columns (2) and (4). Post is a binary variable that takes the value 1 for years after 2003, while Treat is a binary variable that takes the value 1 for Andhra Pradesh and 0 otherwise. All regressions contain state-2 digit industry and 2-digit industry-year fixed effects, and a state-time trend. Control, Pre Mean reports the mean of the outcome variable among the control states in the years before the policy reform. Robust standard errors are clustered at the state level and reported in parentheses below. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

Table A10: Event Study using the NSS

	Log Wages	Casual Workers
	(1)	(2)
1993-94	0.017	0.051***
	(0.020)	(0.012)
1999-00	0.000	0.000
	(.)	(.)
2003-04	-0.103***	-0.001
	(0.012)	(0.006)
2004-05	-0.163***	0.040***
	(0.014)	(0.005)
R2	0.60	0.44
N	170781	377636

Notes: The above table uses data from multiple rounds of the National Sample Survey. The outcome variable in daily log wages are reported in Column (1), and takes the value 1 if an individual reports working in casual work in Column (2). Treat is a binary variable that takes the value 1 for Andhra Pradesh and 0 otherwise. All regressions contain state-2 digit industry and 2-digit industry-year fixed effects, along with a state-time trend. Robust standard errors are clustered at the state level and reported in parentheses below. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$. The sample size varies depending on the coverage of the underlying variable across rounds as well as whether a worker reports it.

Table A11: Heterogeneity by Manufacturing and All Sectors

	Probability of Working In			
	(1) Informal Contract	(2) Informal Firm	(3) Casual Work	(4) Wage
Panel A: Manufacturing Sector (Main Sample)				
Post x Treat	0.034	0.124***	0.075***	-0.073
	(0.022)	(0.015)	(0.021)	(0.046)
Control, Pre Mean	0.48	0.70	0.64	4.37
R2	0.34	0.21	0.27	0.46
N	21679	39644	49268	18822
Panel B: All Sectors (Main Sample)				
Post x Treat	0.055***	0.062***	0.027***	-0.046**
	(0.009)	(0.012)	(0.009)	(0.017)
Control, Pre Mean	0.39	0.67	0.59	4.60
R2	0.43	0.34	0.41	0.50
N	90260	144035	216558	91510
Panel C: Manufacturing Sector (All Years)				
Post x Treat	0.027*	0.116***	0.061*	-0.239***
	(0.014)	(0.027)	(0.030)	(0.047)
Control, Pre Mean	0.48	0.70	0.62	4.08
R2	0.39	0.29	0.32	0.60
N	32281	57756	83409	34384
Panel D: All Sectors (All Years)				
Post x Treat	0.113***	0.115***	0.029***	-0.191***
	(0.012)	(0.018)	(0.009)	(0.020)
Control, Pre Mean	0.39	0.67	0.73	3.94
R2	0.48	0.35	0.44	0.60
N	150989	239663	377636	170781

Notes: The above table uses data from multiple rounds of the National Sample Survey between 1999-2005 for the main sample, and between 1993-2009 for the "all years" sample. Each outcome variable takes the value 1 if an individual reports working on an informal contract (Column 1), informal firm (Column 2), and casual work (Column 3). Daily log wages are reported in Column (4). Panels A and B restrict the sample to the manufacturing and all sectors for the main sample years, while Panels C and D use all years available. Post is a binary variable that takes the value 1 for years after 2003, while Treat is a binary variable that takes the value 1 for Andhra Pradesh and 0 otherwise. All regressions contain state-2 digit industry and 2-digit industry-year fixed effects. Panels C and D additionally include a state-time trend. Control, Pre Mean reports the mean of the outcome variable among the control states in the years before the policy reform. Robust standard errors are clustered at the state level and reported in parentheses below. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$. The sample size varies depending on the coverage of the underlying variable across rounds as well as whether a worker reports it.

Table A12: Impact on LFP, Unemployment, and Agricultural Work

	LFP	Unemployed	Agri. Worker
	(1)	(2)	(3)
Post x Treat	-0.003	-0.002	-0.005
	(0.006)	(0.003)	(0.008)
Control, Pre Mean	0.46	0.02	0.17
R2	0.37	0.03	0.17
N	1444107	1444107	1444107

Notes: The above table uses data from multiple rounds of the National Sample Survey. Each outcome variable takes the value 1 if an individual reports working in the labor force (Column 1), is unemployed (Column 2), or works in agriculture (Column 3). Post is a binary variable that takes the value 1 for years after 2003, while Treat is a binary variable that takes the value 1 for Andhra Pradesh and 0 otherwise. All regressions contain state-2 digit industry and 2-digit industry-year fixed effects, and a state-time trend. Control, Pre Mean reports the mean of the outcome variable among the control states in the years before the policy reform. Robust standard errors are clustered at the state level and reported in parentheses below. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

A.5 Additional Tables from the Economic Census

Table A13: Robustness to Alternate Definitions of Large Firms

	Probability of Remaining Unregistered				
	(1)	(2)	(3)	(4)	(5)
Large Firm:	= 10	= 20	= 30	= 50	= 100
Post × Treat	0.276***	0.318***	0.365***	0.422***	0.385***
	(0.028)	(0.029)	(0.030)	(0.034)	(0.038)
Control, Pre- Mean	0.84	0.80	0.78	0.75	0.72
R2	0.37	0.38	0.38	0.38	0.40
N (millions)	1.36	0.54	0.32	0.18	0.08

Notes: The above table uses data from the 1998 and 2005 rounds of the Economic Census. The outcome variable is a binary variable that takes the value 1 if a firm is unregistered and 0 otherwise. We classify firms as Large if they employ more than 10 workers (Column 1), 20 workers (Column 2), 30 workers (Column 3), 50 workers (Column 4), and 100 workers (Column 5). All regressions include district, 2-digit industry, state-industry, and year fixed effects, along with controls for firm characteristics like whether the firm is owned by a female, whether the owner is from a lower-caste, and whether the firm uses electrical power. Control, Pre Mean reports the mean of the outcome variable among the control states in the years before the policy reform. Robust standard errors are clustered at the state level and reported in parentheses below. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

Table A14: Robustness to Restricting Results to the Manufacturing Sector Only

	Large	Small	All
	(1)	(2)	(3)
Post × Treat	0.292***	0.125***	0.127***
	(0.031)	(0.028)	(0.028)
Control, Pre Mean	0.71	0.95	0.94
R2	0.35	0.22	0.22
N (millions)	0.40	12.44	12.84

Notes: The above table uses data from the 1998 and 2005 rounds of the Economic Census. We restrict the firms to the manufacturing sector only. The outcome variable is a binary variable that takes the value 1 if a firm is unregistered and 0 otherwise. We classify firms as Large (Column 1) if they employ more than 10 workers, and Small (Column 2) if they employ less than 10 workers. Column 3 includes all firms. All regressions contain district, year, and state-2-digit industry fixed effects and control for firm characteristics like whether the firm is owned by a female, whether the owner is from a lower-caste, and whether the firm uses electrical power. Control, Pre Mean reports the mean of the outcome variable among the control states in the years before the policy reform. Robust standard errors are clustered at the state level and reported in parentheses below. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

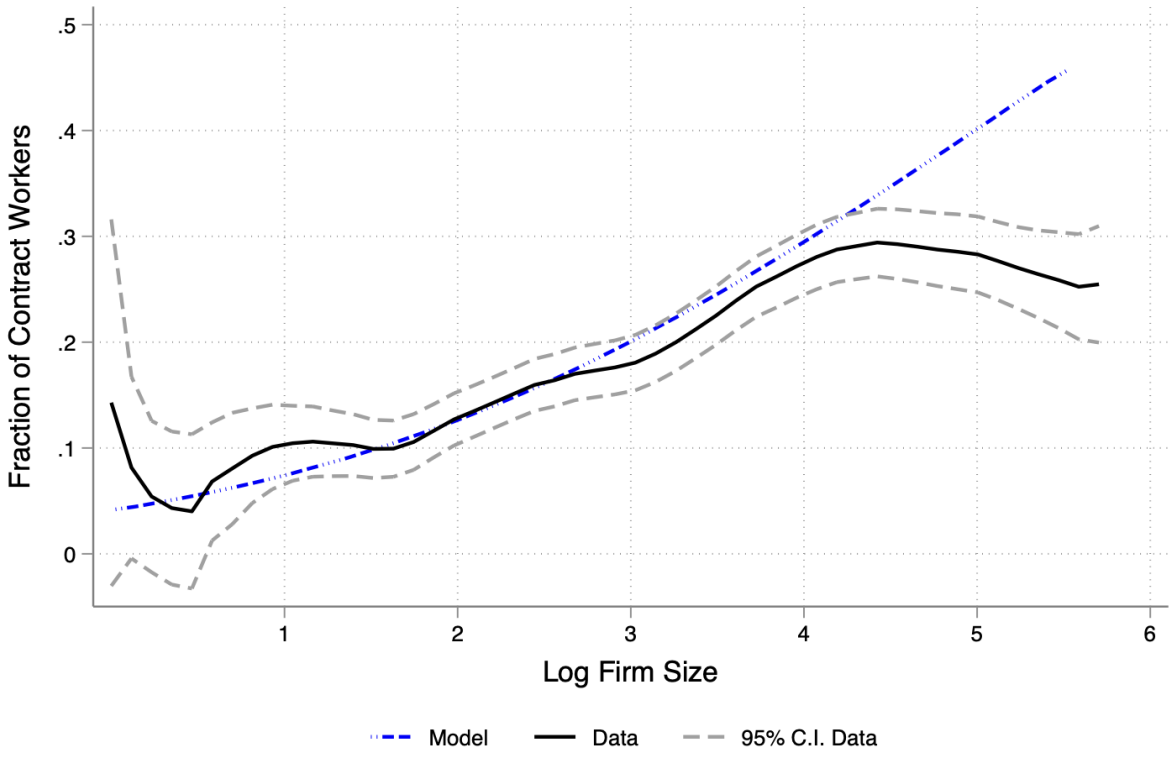
A.6 Model Fit

Table A15: Correlation between Firm-Size and Fraction of Contract Workers

	Fraction of Contract Workers	
	(1) Without F.E.	(2) Ind. F.E.
Log Firm Size	0.123***	0.119***
	(0.027)	(0.027)
Log Firm Size^2	-0.010***	-0.009**
	(0.004)	(0.004)
R2	0.05	0.12
N	2170	2170

Notes: The above data is at from the Annual Survey of Industries for Andhra Pradesh in 2002. Column (1) reports the regression estimates without any fixed effects, while Column (2) adds 2-digit industry fixed effects. Robust standard errors are reported in parentheses. * is $p < 0.1$, ** is $p < 0.05$, *** is $p < 0.01$.

Figure A4: Correlation between the Fraction of Contract Workers and Firm-Size in Andhra Pradesh



Notes: The above figure uses establishment-level data from the Annual Survey of Industries for Andhra Pradesh in 2002 (before the policy) and plots a non-parameteric correlation (solid black line) between log firm-size on the horizontal axis and the fraction of contract workers on the vertical axis. The dotted gray lines are the 95% confidence intervals. The dash-dotted blue line is the fit simulated from the model.

Table A16: Model Fit

		Model	Data
		(1)	(2)
No. of informal firms ('000)	Targeted	1608.01	1606.23
No. of formal firms ('000)	Targeted	16.37	16.34
Firm size in informal firms	Targeted	1.63	2.77
Firm size in formal firms	Targeted	152.46	112.08
Variance in log-firm size (Informal)	Targeted	0.58	0.58
Variance in log-firm size (Formal)	Targeted	1.33	1.33
Coefficient	Targeted	-0.01	-0.01
Frac. Informal labor	Untargeted	0.63	0.79
Frac. Informal firms	Untargeted	0.99	0.99

Notes: The above table reports the moments in the model (Column 1) with those in the data (Column 2).

Table A17: Derivatives of Moments to Parameter Changes

Moment	σ_x	σ_e	θ	ϕ	b_c	b_r
	(1)	(2)	(3)	(4)	(5)	(6)
Var. Firm-size (Informal)	0.08	1.58	5.34	0.13	0.16	0.05
Var. Firm-size (Formal)	0.40	1.84	0.09	0.97	0.51	0.62
Avg. Firm-size (Informal)	0.53	0.48	3.05	0.34	0.42	0.03
Avg. Firm-size (Formal)	0.69	0.38	0.86	0.59	1.40	1.49
Coeff. On $Ln(FirmSize)^2$	0.10	0.02	0.07	0.97	0.20	0.37
Frac. Contract Workers	0.12	0.03	0.10	0.26	0.31	0.43

Notes: This table reports the derivatives of each moment with respect to each parameter. Each row is a moment calculated from the model simulation. Each number in the table indexed by row R and column C, is the percent change in the moment in row R, when a parameter in column C is increased by 1 percentage.

B Permitted Non-Core Activities for Contract Workers

1. Sanitation works, including Sweeping, Cleaning, Dusting, and Collection and disposal of all kinds of waste.
2. Watch and ward services including security service.
3. Canteen and Catering services.
4. Loading and Un-loading Operations.
5. Running of Hospitals, Educational & Training Institutions, Guest Houses, Clubs and the like where they are in the nature of support services of an Establishment
6. Courier Services which are in nature of support services of an Establishment.
7. Civil and other constructional works, including maintenance.
8. Gardening and maintenance of lawns etc.
9. Housekeeping and laundry services etc., where they are in nature support services of an Establishment.
10. Transport services including Ambulance Services.
11. Any activity of intermittent in nature even if that Constitutes a core activity of an Establishment and
12. Any other activity which is incidental to the core activity.

ONLINE APPENDIX

C1 Theory Details and Mathematical Proofs

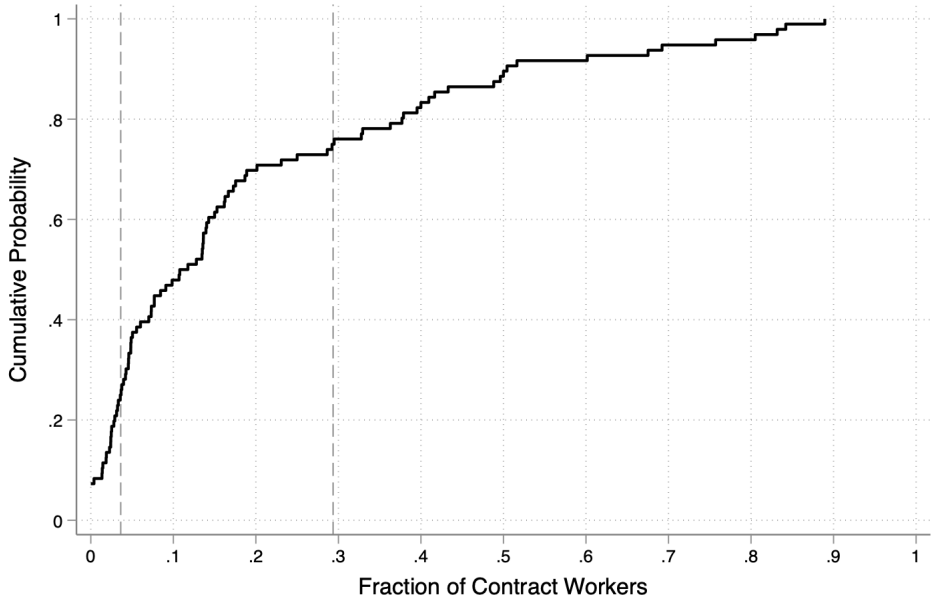
C1.1 Modeling the Usage of Contract and Regular Workers in the Formal Sector

In this section, we provide evidence to support our assumption of modeling contract and regular workers as imperfect substitutes in the production process. To that effect, it is possible that *within* a task, workers could be substitutes between contract and payroll. However, firms might allocate workers differently *across* tasks. The ASI data does not provide us with details on tasks and occupations that workers are assigned to. However, we leverage multiple rounds of the NSS data from our paper, where we can measure the occupations that formal sector workers work in, based on the nature of their contract (payroll vs contractual).

Specifically, for formal sector workers in the NSS, we calculate the fraction of contract workers in a 2-digit occupation. Pooling across all years, Figure C1 reports the CDF of this distribution, where we find a substantial variation — the bottom 25% of occupations have less than 5% their workers on contracts, while the top 25% of occupations have at least 30% or more of their workers on contracts. Therefore, there appears to be a substantial variation in how firms allocate workers across occupations based on the status of their contracts.

To be able to precisely measure the variation in contract status that can be attributed to within occupations, as opposed to across them, we construct a Theil Index. The Theil Index, similar to a Gini coefficient, allows us to measure the extent of inequality in the distribution of contract workers. For example, an index of 0 would imply that the fraction of contract workers does not vary across occupations. On the other hand, an index of 1 would imply complete seggregation. In addition, an advantage of the Theil Index is that it can further be decomposed to examine what fraction of this aggregate inequality is driven by inequality *within* an occupation, as opposed to across occupations. We calculate the Theil Index (and its

Figure C1: Distribution of Contract Workers Across Occupations in Formal Firms



Note: The above figure pools the sample of workers working in the formal sector across multiple rounds of the NSS between 1999-00 and 2009-10. The figure plots the fraction of contract workers across occupations. The dotted gray lines report the 25th and 75th percentiles.

decomposition) separately for each round of the NSS. We find that on average, across all rounds, 42% of the variation in the distribution of contract workers is driven *across* occupations as opposed to within them.[45] Put together, the above exercise suggests that while workers (irrespective of contracts) could potentially be substitutes within a task/occupation, the status of their contract non-trivially determines what occupations they are allocated to. If these occupations are imperfect substitutes in the production process, then the way that we model it as a CES framework could be a reasonable approximation.

[45]This ranges from 36% to 48% across rounds.

C1.2 Firms in the Formal Sector

Production Decision of Incumbents: The profit maximization problem of the incumbents in the formal sector can be broken into a two-step problem, where in the first step, firms maximise:

$$\pi_f = \max_{l_f}(1-t)pzl_f^{\rho} - w_f l_f$$

Define $w_f = \left[\tilde{b}_r^{1-\sigma} + b_c^{1-\sigma}\right]^{\frac{1}{1-\sigma}} w$, where $\tilde{b}_r = b_r z^{-\phi}$. Taking the first order condition and solving, we get:

$$l_f^*(z) = \left[\frac{\rho(1-t)p}{w_f}\right]^{\frac{1}{1-\rho}} \times z^{\frac{1}{1-\rho}}$$
$$y_f^*(z) = \left[\frac{\rho(1-t)p}{w_f}\right]^{\frac{\rho}{1-\rho}} \times z^{\frac{1}{1-\rho}}$$
$$\pi_f^*(z) = (1-\rho)\left[\frac{\rho}{w_f}\right]^{\frac{\rho}{1-\rho}} \times ((1-t)pz)^{\frac{1}{1-\rho}}$$

Given this, the firm then choose l_r and l_c in a dual problem, given by:

$$\min w\tilde{b}_r l_r + w b_c l_c$$
$$\text{s.t. } l_f = \left[l_r^{\frac{\nu-1}{\nu}} + l_c^{\frac{\nu-1}{\nu}}\right]^{\frac{\nu}{\nu-1}}$$

Let w_f, defined above, be the Lagrangian constant so that:

$$L = wb_r l_r + wb_c l_c - w_f\left[l_f - \left[l_r^{\frac{\nu-1}{\nu}} + l_c^{\frac{\nu-1}{\nu}}\right]^{\frac{\nu}{\nu-1}}\right] \tag{13}$$

Taking the first-order conditions with respect to l_c, l_r and w_f and solving, we get:

$$wb_p l_p = \frac{b_p^{1-\nu}}{b_p^{1-\nu} + b_c^{1-\nu}} \times w_f l_f^*(z) \tag{14}$$
$$wb_c l_c = \frac{b_c^{1-\nu}}{b_p^{1-\nu} + b_c^{1-\nu}} \times w_f l_f^*(z)$$

From Equation 14, the ratio of contract to permanent workers is given by:

$$\frac{l_c}{l_p} = \left(\frac{b_c}{b_p}\right)^{-\nu}$$

Entry Decision: Let E_f be the fixed cost of entry (in units of output) into the formal sector. A firm can enter the formal sector as long as $\pi_f \geq pE_f$. This implies that there is a threshold productivity z_f^E where $\pi_f(z_f^E) = pE_f$. Rearranging the variable profit equation from above, we get:

$$z_f^E = \frac{\left[E_f/(1-\rho)\right]^1 - \rho}{(1-t)\left(\rho p/w_f\right)^{\rho}}$$

C1.3 Incumbents in the Informal Sector

The profit maximization problem of the incumbents in the informal sector is straightforward and can be given by:

$$\pi_i = \max_{l_i} pzl_i^{\rho} - w_i l_i^{\theta}$$

Taking the first order condition and solving, we get:

$$l_i^*(z) = \left[\tilde{\rho} \times \frac{p}{w} \times z\right]^{\frac{1}{\theta-\rho}}$$

$$y_i^*(z) = \left[\frac{\tilde{\rho}p}{w}\right]^{\frac{\tilde{\rho}}{1-\tilde{\rho}}} \times z^{\frac{1}{1-\tilde{\rho}}}$$

$$\pi_i^*(z) = (1-\tilde{\rho})\left[\frac{\tilde{\rho}}{w}\right]^{\frac{\tilde{\rho}}{1-\tilde{\rho}}} \times (pz)^{\frac{1}{1-\tilde{\rho}}}$$

where $\tilde{\rho} = \rho/\theta$.

C1.4 Cost Function and Probability of Detection

An alternate way to present the model is to allow for a size-dependent penalty of operating in the informal sector. Let $\tau(l)$ be the penalty function such that $\tau(l) > 0$, $\tau'(l) < 0$ and $\tau(\infty) \to 0$. Therefore maximization problem of the firm can be written as:

$$\max_l \tau(l) pz l^\rho - wl$$

Taking the first order condition and rearranging:

$$\left[\rho\tau(l) + l\tau'(l)\right] pz l^\rho = wl \tag{15}$$

Comparing it to the baseline model, we have:

$$\left[\frac{\rho}{\theta} l^{\frac{\rho}{\theta}-\rho}\right] pz l^\rho = wl \tag{16}$$

Equations 15 and 16 are therefore connected through the $\tau(l)$ function. This implies:

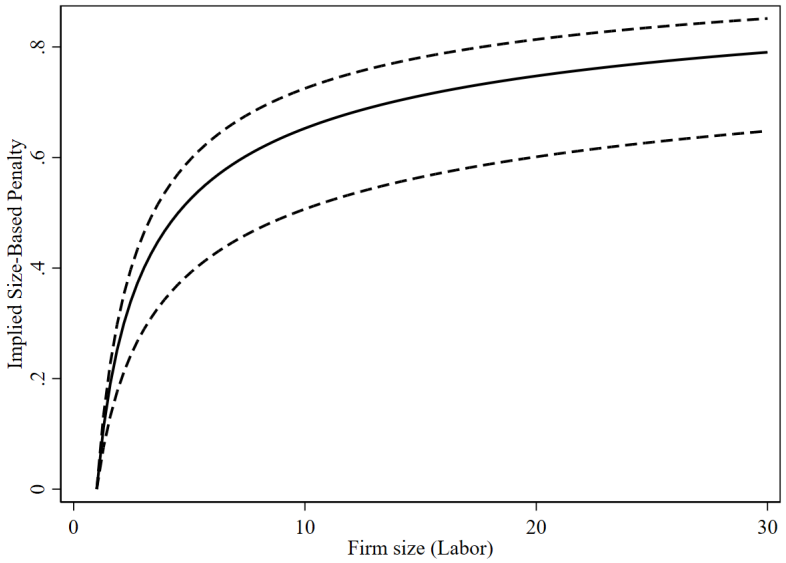
$$\rho\tau(l) + l\tau'(l) = \frac{\rho}{\theta} l^{\frac{\rho}{\theta}-\rho} \tag{17}$$

This is a differential equation of the form $ay + xdy/dx = bx^c$. This has a general solution of the form $y = \frac{bx^c}{a+c} + \frac{k}{x^a}$ where k is an integration constant. Therefore:

$$\tau(l) = \left[l^{\frac{\theta-1}{\theta}} + k \right] l^{-\rho} \tag{18}$$

Assume $\tau(l) = xl/(1+xl)$, where x is a parameter that governs the marginal change in the probability detection as firm size increases. A larger x implies a higher probability of being detected conditional on the same firm-size (l). From the calibration exercise in the paper, we have $\rho = 0.738$ and $\theta = 1.3$ and from Equations (15) and (16), this implies a value of $x = 0.088$ and a probability detection function shown in Figure C2.

Figure C2: Size-based penalty function



Notes: The above graph plots the size-based penalty function of operating in the informal sector, as a function of firm size. The solid black line show the median value across all industries, while the dashed lines show the 25th and 75th percentiles.

C2 Details of Estimation Algorithm

The estimation algorithm proceeds as follows:

Step 1: Take a guess for an initial set of parameters in φ (defined in the paper). Equation 1 allows us to calibrate the value of $b_{r,j}$ such that $b_{r,j} = R_j^{1/\nu} \times b_{c,j}$, where R_j is the ratio of contract to payroll workers that is observable in the data.

Step 2: Using the distributional assumptions on the productivity distribution as well as the potential entrants (M), we compute the threshold productivity in the informal and formal sectors given by:

$$\frac{N_f}{M} = 1 - F(x_f^*) \tag{19}$$

$$\frac{N_f + N_i}{M} = 1 - F(x_i^*) \tag{20}$$

where: the number of potential entrants, M, is calculated by subtracting the total number of entrepreneurs and workers in the formal and informal firms in Andhra Pradesh from the total adult workforce of Andhra Pradesh as reported in the 2001 Population Census of India.

Step 3: From Equations 2 and 4, the variance in log-labor in the formal and informal sectors will be given by:

$$Var(\ln l_f) = \frac{\sigma_f^2 + \sigma_\varepsilon^2}{(1-\rho)^2} \tag{21}$$

$$Var(\ln l_i) = \frac{\sigma_i^2 + \sigma_\varepsilon^2}{(\theta-\rho)^2} \tag{22}$$

where, $\sigma_f^2 = Var(\ln x | x \geq x_f^*)$ and $\sigma_i^2 = Var(\ln x | x \in \{x_i^*, x_f^*\})$.

Step 4: In an inner loop, we take a guess for the price vector prices and wages. Conditional on the parameter values in the outer loop, we use Equations 2, 3, 4 and 5, to calculate the total demand and supply in the goods and labor markets.

We then update the price vector in an iterative loop until the goods and labor markets clear (we set the threshold tolerance for convergence to 0.5%).

Step 5: After calculating the equilibrium price and wages, we calculate the average firm size in the formal and informal sectors using Equations 2 and 4.

Step 6: Using the simulated moments from above, we then choose the parameter vector that minimizes the distance between the data moments and simulated moments such that $\widehat{\varphi} = \text{argmin}_{\varphi}\ g(\varphi)'\ W\ g(\varphi)$, where W is a $5J + 2 \times 5J + 2$ weighting matrix that is selected efficiently using a two-step process. In the first step, we set W to be the identity matrix I and calculate $\widehat{\varphi}_1$. Using $\widehat{\varphi}_1$, we can then compute $W^{TwoStep} = \frac{1}{N}[g(\varphi_1)g(\varphi_1)']^{-1}$, which we use as the weighting matrix and redo the exercise to calculate $\widehat{\varphi}^{SMM}$.

Step 7: We then use the zero-profit conditions in the informal and formal sectors (Equation 7) to calibrate the fixed cost of entry and formalization in these industries.

Step 8: Duffie and Singleton (1993) provide weak and strong consistency conditions that are satisfied in the model. Similar to Gourinchas and Parker (2002) and Ulyssea (2018), asymptotic normality holds as long as (i) φ_0 and φ^{SMM} are in the interior parameter space; (ii) the simulator used is continuously differentiable w.r.t. φ in the neighborhood of φ_0 and (iii) $G = E[\frac{\partial g(\varphi^{SMM})}{\partial \varphi}]$ exists, is finite and $G'W'G$ is non-singular. Standard errors are then calculated using the empirical gradient of G i.e., using small changes to φ^{SMM}, and given by $\sqrt{\frac{1}{N}(G'WG)^{-1}}$.