Small by Choice? Reassessing the Aggregate Implications of Size-Based Regulations

Kalyani Padmakumar*

November 2022

Please click here for the latest version

Abstract

Firms with employment exceeding a threshold level are more regulated than smaller firms in many countries. Current work gauges such rules' impact based on whether firms bunch at the threshold, with more bunching taken to imply higher regulatory costs. In this paper, I show that size-based rules can be highly restrictive for firms even with no bunching. I argue using India's employment protection legislation (EPL) that applies to manufacturing plants with 100+ full-time workers. Although plants do not bunch at 100 workers, I show that EPL restricts employment growth of plants by lowering their transitions from below to above the regulatory threshold. Plants near the threshold also substitute towards contract workers and capital, on which EPL does not apply. To quantify aggregate implications, I develop and estimate a structural dynamic heterogeneous firm model with multiple factors of production and labor search costs. Counterfactuals suggest that removing EPL would make more firms enter and hire full-time workers faster, raising full-time employment and output by 9% and 4% in the long run.

*Department of Economics, The Pennsylvania State University, State College, Pennsylvania, USA 16801. Email: kxp338@psu.edu.

1 Introduction

Large firms are more regulated than smaller firms in many countries. For instance, labor market regulations in France apply to firms with 50+ workers, and the Indian employment protection legislation (henceforth, EPL) applies to manufacturing plants with 100+ full-time workers. Since these rules are size-based, one would expect firms to avoid falling under their purview by choosing to bunch at the threshold, as in the French firm size distribution in Figure 1. This approach to evaluate size-based rules has been adopted extensively in existing work (Guner et al. [2008], García-Santana and Pijoan-Mas [2011], Gourio and Roys [2014], Garicano et al. [2016] and Amirapu and Gechter [2020]), with the implication that the more bunching there is at the threshold, the more costly the rule is. However, there is no such bunching at 100 workers in the Indian plant size distribution, as shown in Figure 2 below[1]. Should this be taken to mean that this legislation is not restrictive? This paper says no, and I offer a new approach to gauge the restrictiveness and effects of such size-based rules.

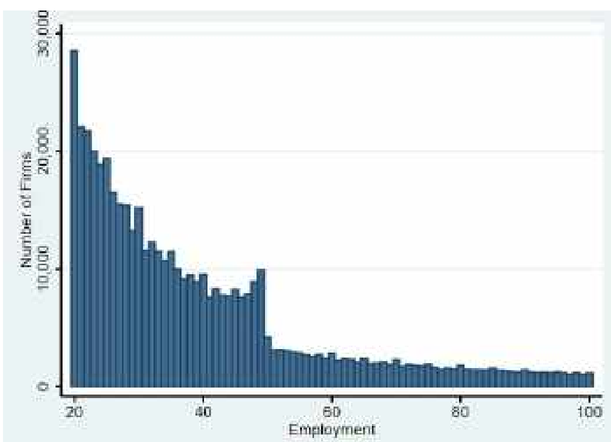


Figure 1: French firm size distribution
Source: Gourio and Roys [2014]

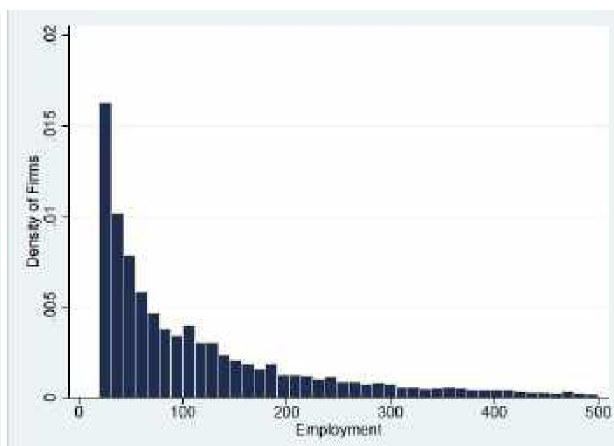


Figure 2: Indian plant size distribution
Source: Hsieh and Olken [2014]

In this paper, I show that the Indian legislation slows down employment growth of plants by lowering their transitions from below to above the regulatory threshold of 100 full-time workers. I also find that plants near the threshold substitute towards other factors of production on which the legislation does not apply, such as temporary contract workers and capital goods. Further, to explain the lack of bunching by Indian plants and to shed light

[1] This fact has been rigorously established in Hsieh and Olken [2014] and Amirapu and Gechter [2020].

on aggregate implications of the legislation, I develop a dynamic heterogeneous firm model with multiple factors of production and labor search costs. The model links size-based rules to decisions on labor search by plants and identifies conditions under which plants would bunch or not bunch in response to these rules. The model is then fit to plant-level data from India. Lastly, I quantify the legislation's output and employment consequences by comparing steady states under alternative regulatory regimes.

The estimated model implies that removing the burden of this legislation would increase aggregate full-time employment and output in the treated states by 9% and 4%, respectively, in the long run. Removing this legislation would impact aggregate outcomes through two channels. First, it would allow more plants to enter by increasing the value of entry. Second, these heterogeneous plants hire more full-time workers, increasing aggregate output. These findings imply that an exclusive focus on bunching to evaluate size-based rules is likely to yield misleading results.

What makes the Indian EPL costly for plants? This legislation requires manufacturing plants with 100+ full-time workers to obtain government authorization and provide severance pay before firing even one worker. The legislation imposes two kinds of costs on plants. First, anecdotal evidence suggests that plant owners face considerable delays in securing government authorization, thereby increasing the uncertainty they face during an adverse shock. These delays increase firing costs for plants with 100+ full-time workers. Second, government officials enforce labor regulations in India through frequent inspections of plant premises. These inspectors have the power to extract bribes and tighten the administrative burden that plant owners face. Amirapu and Gechter [2020] cites several instances of inspectors exploiting these laws' arbitrary and antiquated nature to extract bribes. Plant owners incur these costs not just during adverse shocks, and the bribes they pay typically increase with the number of workers. Following Amirapu and Gechter [2020], this paper interprets the enforcement regime as increasing the marginal cost of production for plants with 100+ full-time workers.

The main challenge in analyzing the causal relationship between this legislation and plant outcomes is in credibly measuring regulatory costs since they cannot be observed in the data. Previous work, such as Besley and Burgess [2004], Hasan et al. [2007], Ahsan and Pagés [2009], studied the legislation's impact on plants by exploiting the fact that the power to amend this legislation lies with the states. For instance, by aggregating all state-level amendments from 1958 to 1992, Besley and Burgess [2004] classified states as pro-worker/

pro-employer and analyzed how outcomes like output and employment varied between pro-worker and pro-employer states. However, Bhattacharjea [2006] and Bhattacharjea [2020] highlight issues with how existing work classifies states as pro-worker and pro-employer.

In this paper, I establish the legislation's causal effect on plants by exploiting a natural experiment implemented in several Indian states. Starting in 2014, several Indian states increased the threshold of this legislation from 100 to 300 full-time workers in a staggered manner. This policy varied regulatory costs for plants that employed between 100 and 300 full-time workers in the treated states. Using panel data on Indian manufacturing plants from 1999 to 2018 and exploiting the staggered roll-out of this policy, I document two main findings. First, treated plants are more (less) likely to transition around 100 (300) full-time workers post-policy. Second, treated plants around 100 (300) full-time workers employ fewer (more) temporary contract workers and capital goods in production post-policy. These estimates are robust to a rich set of controls, including year fixed-effects, industry-year-size fixed effects, and plant fixed effects. Furthermore, event studies confirm that these significant impacts start only from the year of the reform and not earlier. They rule out the possibility that these post-policy impacts are a mere continuation of past trends.

But there are two main questions that still need to be answered. First, it is unclear why plants do not bunch in response to the Indian EPL. Second, what are the output and employment consequences of the legislation? In order to answer these questions, this paper develops a continuous-time dynamic heterogeneous firm model with multiple factors of production and labor search costs.

The model incorporates several features from existing work on firm dynamics (Hopenhayn [1992], Hopenhayn and Rogerson [1993], Melitz [2003], Das et al. [2007], Coşar et al. [2016]), such as monopolistically competitive markets, plants heterogeneous in their productivity which evolves over time and an endogenous mass of plants operating in the economy. I add to this setup in two ways. First, plants employ heterogeneous labor (full-time and contract workers) and capital goods in production. The production function exhibits different degrees of substitutability between these inputs based on the causal evidence documented above. It is isomorphic to the ones used in Krusell et al. [2000], Parro [2013].

Second, given the Indian institutional context, plants face different costs while adjusting these inputs. In particular, I assume that plants can adjust contract workers and capital without any costs. However, adjusting the size of full-time workers is costly and time-consuming for the plant. What is crucial here is the nature of these adjustment costs. Hiring these workers requires effort by the plant; more effort means a faster arrival rate of workers on average. However, high levels of effort involve convex search costs. Similarly, each full-time worker gets separated from the plant at an exogenous rate. If plants want to expedite the firing of workers, they can increase the separation rate at a cost. These formulations of hiring and firing costs build on existing work in the labor search literature with multi-worker firms (Acemoglu and Hawkins [2014], Kaas and Kircher [2015]) and network formation in international trade (Eaton et al. [2016], Eaton et al. [2022]). I then link size-based rules with plants' hiring and entry decisions.

There are two key takeaways from this framework. First, costly size-based rules lower transitions of plants around the threshold. As plants near the threshold, their expected marginal benefit from having more workers is lower, with higher regulatory costs. Hence, plants' search effort decreases as they approach the threshold. The decrease in search effort lowers the rate with which they get additional workers and move forward. The larger the magnitude of regulatory costs, the lower their transition probability around the threshold.

Second, in steady state, whether or not the distribution of full-time workers exhibits bunching depends on the separation rate of full-time workers and the convexity of search cost. To see this, consider plants that slow down their search because of the regulation. If the separation rate of full-time workers and the convexity of search cost are both high, the marginal benefit these plants receive from employing at the threshold would not be worth the marginal costs they have to pay. As a result, they would be less likely to bunch at the threshold. If either of these parameters is low, their marginal benefit from employing at the threshold would be worth the search costs. This would make them bunch at the threshold.

The above framework is then fit to plants in the treated states using simulated method of moments. In order to estimate policy invariant parameters and regulatory costs, this procedure leverages information contained in the policy change's causal impacts documented earlier, the pre-policy distribution of full-time workers, and the pre-policy share of contract workers employed by different-sized plants. Although the above framework is fit to only labor data, the estimated model replicates the pre-policy share of output contributed by different-sized plants reasonably well.

The quantified model implies that removing the burden of this legislation would increase aggregate full-time employment and output in the treated states by 9% and 4%, respectively, in the long run. Therefore, policymakers can use this legislation to address several distortions in Indian manufacturing that the misallocation literature has highlighted.

For instance, as pointed out by Hsieh and Klenow [2014], full-time employment by Indian manufacturing plants does not grow with age compared to Mexico and the US. The above findings suggest that relaxation of this legislation would make plants expand their full-time workforce faster, thereby improving the correlation between age and full-time employment. This policy change would also make plants less reliant on temporary contractual workers, who are not protected by these laws, for their operations. Bertrand et al. [2021] argues that plants hire contract workers via contractors or staffing companies. These companies charge a high fee to the plants making the latter less competitive. Besides, with more contract workers, plants have fewer incentives to invest in their skills which would adversely impact productivity. Therefore, by changing the composition of employment towards more full-time workers, relaxation of this legislation could increase the comparative advantage of manufacturing plants and reduce the thick left tail of the productivity distribution documented in Hsieh and Klenow [2009].

The rest of the paper is organised as follows. In section 2, I discuss the Indian EPL in greater detail and highlight the two channels through which it imposes costs on plants. Section 3 describes the data source that I use for my analysis. Section 4 contains my main empirical findings. In subsection 4.1, I show that the policy change impacted transition probabilities of plants around the 100 worker and 300 worker threshold, which is the main result of this paper. In subsection 4.2, I show that contract labor share and capital labor ratio of plants close to the 100 and 300 worker threshold changed significantly in the expected directions after the policy change. Section 5 describes the structural model that I use to quantify the magnitude of regulatory costs. In subsections 5.1 to 5.3, I lay out the model setup and describe the equilibrium conditions. Subsection 5.4 delineates the conditions under which costly size based rules lead to bunching and no bunching in the plant size distribution. Section 6 discusses the strategy with which I use the model to quantify the magnitude of regulatory costs. Section 7 describes the estimates, model fit and counterfactual results. Section 8 concludes.

2 Institutional Background

A complex web of 463 Acts, 32542 compliances, and 3048 filings govern employer-employee relationships in Indian industries. Amongst these, the piece of legislation that has attracted the most attention in policy circles is chapter VB of the Industrial Disputes Act (IDA) of 1947. Chapter VB of the IDA mandates industrial plants that employ more than 100 workmen to acquire permission from a government authority before laying off even one worker. IDA defines a workman as ‘any person employed in any industry, to do any skilled or unskilled manual or clerical work, for hire or reward.’ This definition includes full-time or permanent workers the plant hires directly for production. However, it excludes managerial staff and temporary contract workers engaged in production since the latter set of workers is hired through staffing companies and not by the plant directly.

The authority in question is usually a labor court or an industrial tribunal whose primary objective is to adjudicate industrial disputes between employers and employees. The labor court would grant or refuse permission after conducting an inquiry and allowing the employer and worker concerned to present their case. Once the employer secures permission from the labor court, they are mandated to pay the affected worker severance pay of 15 days per year of service. If either party is not satisfied with the labor court’s verdict, they can approach the higher courts (the High Court or the Supreme Court).

It is often argued that IDA constrains plants from laying off workers even during adverse shocks, thereby increasing the plant’s firing cost after employing more than 100 full-time workers. More specifically, the opportunity cost of time and effort for employers in following the due process as laid down by these rules is very high because Indian state capacity is weak. As of 1998, there existed only one labor court for 437868 workers. The number of cases pending before Labor Courts is over 0.1 million as of 2020. 35% of these 0.1 million cases have been unresolved for over a year and 37% pending for more than three years as shown in table 1[2]. The Union Labor Minister gave three reasons for the large pendency of cases in these labor courts. "One, the absence of affected parties at the time of the hearing; two, seeking of frequent adjournments by the parties to file documents; and three, parties approaching the Courts challenging orders of reference issued by the appropriate government as well as orders issued by the Tribunals on preliminary points;. . . ." (Shyam Sundar(2020), Debroy and Bhandari [2008]).

Additionally, government officials enforce labor regulations in India through frequent inspections of plant premises. These inspectors have the power to extract bribes and tighten the administrative burden that plant owners face. Amirapu and Gechter [2020] cites several instances of inspectors exploiting these laws’ arbitrary and antiquated nature to extract

[2] The number of labor-related cases unresolved before the High Court and Supreme Court also run into lakhs as can be seen in table 2.

bribes. Plant owners incur these costs not just during adverse shocks, and the bribes they pay typically increase with the number of workers (Debroy and Bhandari [2008], Amirapu and Gechter [2020]).

Although one cannot observe the magnitude of regulatory costs in the data, one would expect forward-looking plants to consider these costs when deciding whether or not they should enter or employ more than 100 full-time workers in the future. These costs will also be reflected in their demand for other inputs like contract workers and plant & machinery[3]. Therefore, one can infer if the regulation is costly for plants by analyzing the impact of an exogenous variation in these regulatory costs on transition probabilities around the 100-worker threshold and the plants' intake of contract labor and plant & machinery.

Given this objective, a series of policy changes implemented in several Indian states after 2014 acquire significance. In order to improve the ease of doing business, several Indian states increased the legislation's threshold from 100 to 300 full-time workers. This varied regulatory costs for plants that employed between 100 and 300 full-time workers in these states[4]. If this legislation was costly for plants, then one would expect to see changes in transitions of plants around the 100 and 300 worker thresholds, and in their demand for contract labor and plant & machinery near these thresholds. Therefore, the subsequent section's objective is to study whether the policy changes implemented in these treated states impacted these outcome variables.

3 Data Source

In this paper, I use the Annual Survey of Industries dataset (henceforth, ASI), published by the Ministry of Statistics and Program Implementation of the Government of India, for analysis. The unit of enumeration in ASI is at the establishment/factory/plant level, and the sampling frame includes all registered manufacturing plants in India. All manufacturing plants with more than ten workers (if they use power) and those with more than twenty

[3]Plants hire contract workers through staffing companies on a temporary contractual basis. They are not full-time employees of the establishment, they get paid 20% less than formal workers, and they do not fall under the purview of IDA (Bertrand et al. [2021]).

[4]Figure 3 in the Appendix contains the list of treated states and their year of treatment. India follows a federal system of governance - a Central Government for the entire country and a Government for each state. The Constitution of India has conferred the authority to amend this regulation on the Central and State Governments.

workers (if they do not use power) have to be registered by law.

In ASI, the sample to be surveyed each year is drawn from the sampling frame as follows. Plants in the sampling frame belong to the census sector and the sample sector. Plants that employ more than 100 workers regardless of location and those located in the six least industrially developed states, regardless of their employment, belong to the census sector. The remaining plants, those not located in the six least industrially developed states and employing not more than 100 workers, belong to the sample sector. Since grouping plants into census and sample sectors depends on size, which varies with time, this classification gets updated regularly. Plants belonging to the census sector would get surveyed yearly until the next revision occurs. Those belonging to the sample sector in each state are arranged into different groups based on their 4-digit industry classification, and 1/5th of plants get drawn from each (state, 4-digit industry) combination based on stratified circular systematic sampling.

This paper uses the panel version of the ASI dataset from 1998 to 2018. ASI has information about the state in which plants operate, and the analysis covers manufacturing units in all major states of India. They also report detailed plant-level information - such as the industries they belong to, employment of full-time workers/contract workers/managerial staff, sales, profits, current assets, current liabilities, and the book value of fixed assets. The availability of detailed plant-level data enables me to incorporate a rich set of controls while studying how plant dynamics responded to the policy change.

ASI is better suited than two other datasets on the Indian manufacturing sector usually used by researchers- the Economic Census and the CMIE Prowess Database - to answer the central question of this paper. Although the Economic Census covers the universe of all manufacturing plants in India, they publish their data only once in 5 years. Hence, it is impossible to track individual plants over time. CMIE Prowess database reports large firms' balance sheet information; however only 10% of firms provide information on the number of workers they employ. Moreover, Indian labor laws are applicable at the plant level, not at the firm level.

4 Reduced Form Evidence

4.1 Policy Impact on Plant Size Transitions

This section investigates whether the policy change affected plant transitions around the 100-worker and 300-worker threshold. In other words, among plants that employed less than or equal to 100 (300) full-time workers in one period, are they more (less) likely to employ above 100 (300) workers in the subsequent period after policy? If the legislation was costly for plants, then one would expect the policy change to have the maximum impact on transitions around 100 and 300 workers for the following reason. Before policy, plants would have faced a steep increase in labor costs if they had transitioned from below 100 full-time workers to above 100 full-time workers. After policy, plants would no longer face a steep increase in labor costs while making this transition. However, with 300 full-time workers being the new threshold, plants would face a larger increase in labor costs if they had to employ more than 300 workers.

This paper tests the above hypothesis in the following manner. Since the policy impact on plant size transitions would vary depending on its size, I divide the distribution of full-time workers (l_f) into five equally-sized bins. The variable Bin_{ijst} takes values from 1 to 5, depending on the number of full-time workers employed by plant i in industry j in state s at time t.

	$Bin_{ijst}=1$		$Bin_{ijst}=2$		$Bin_{ijst}=3$		$Bin_{ijst}=4$		$Bin_{ijst}=5$
l_f 0		19		100		161		300	

With this definition of Bin_{ijst}, this paper measures the policy impact by estimating the below difference-in-differences regression.

$$\begin{aligned} Y_{ijst} &= \gamma_t + \alpha_s + \alpha_j + \gamma' X_{ijst} + \sum_{k>1} \beta_k^{pre} \mathbb{1}[Bin_{ijst-1} = k] \\ &+ \delta \mathbb{1}[Post_{st}] + \sum_{k>1} \beta_k^{post} \mathbb{1}[Post_{st}] \mathbb{1}[Bin_{ijst-1} = k] + \epsilon_{ijst}; \end{aligned} \tag{1}$$

The outcome variable Y_{ijst} is defined as below, and it captures transitions of plants across bins. To see this, suppose $Bin_{ijst-1} = 2$. Then, the dependent variable would take the value one if the plant employs more than 100 workers in the next period. Analogously, suppose $Bin_{ijst-1} = 4$. Then, the dependent variable would take the value one if the plant employs more than 300 workers in the subsequent period. Bins 2 and 4 are the policy-relevant bins, and the other bins serve as placebos. Additionally, Y_{ijst} is defined only if plant i is sampled in both t and t-1.

$$Y_{ijst} = \begin{cases} 1 & \text{if } Bin_{ijst} > Bin_{ijst-1} \\ 0 & \text{if } Bin_{ijst} \leq Bin_{ijst-1} \end{cases}$$

In specification (1), $\mathbb{1}[Post_{st}]$ is a dummy variable that takes the value one if state s implemented the policy by year t and 0 otherwise. k refers to the value of Bin_{ijst-1}. Here, $Bin_{ijst-1} = 1$ is the base category, so all outcomes are relative to this category. In order to minimize omitted variable bias, I incorporate a rich set of covariates that are both time-invariant and time-varying in the regression specification. In particular, industry fixed effects (α_j) and state fixed effects (α_s) account for time-invariant unobserved heterogeneity at the industry level and state level, while year fixed effects (γ_t) control for aggregate shocks. X_{ijst} includes industry-year-bin fixed effects, which accounts for differential trends by industry-firm size, and time-varying plant-specific characteristics such as the previous period share of contract workers, managerial staff, sales, and capital.

With these controls, β_k^{pre} measures the differential response of transition probabilities by plants in $Bin_{ijst-1} = k$ relative to the base category in the pre-policy period. δ captures the average impact of the policy on transition probabilities from the base category. The coefficient $\delta + \beta_k^{post}$ captures the policy impact on transition probabilities by plants in $Bin_{ijst-1} = k$. Since the set of controls includes industry-year-bin fixed effects and previous period share of contract workers, $\delta + \beta_k^{post}$ is identified by comparing two plants with the same previous period contract worker share and within the same industry-year-size category. Standard errors have been clustered at the plant level.

Figure 4 reports estimates of $\delta + \beta_k^{post}$ for each bin k from equation (1). Following the implementation of the policy, there is a 5 percentage point increase in the probability with which plants in the treated states transition from below 100 full-time workers to above 100 full-time workers. Analogously, there is a 5 percentage point decrease in the probability with which plants transition from below 300 full-time workers to above 300 full-time workers post-policy. There is no statistically significant impact of the policy on transitions around any other threshold, which suggests that these results are not driven by reporting errors or due to any other shock in the treated states. These results are consistent with what one would expect. As discussed in the beginning of the section, one would expect the policy change to have an impact only around 100 and 300 full-time workers, and not around any other threshold.

Threats to Identification: One concern is whether the above result is driven by inherent differences between plants in the treated and control states. I re-estimate specification (1) with plant fixed effects (α_i) as an additional control variable to address this. As long as differences in plants between the treated and control states are time-invariant, the selection will be controlled by plant-fixed effects. The coefficient of interest is still $\delta + \beta_k^{post}$, but now it captures within-plant variation in transition probability over time. Figure 5 reports the results from estimating equation (1) with plant fixed effects. The results indicate a similar percentage point change in transition probability around the policy-relevant thresholds as before.

Another concern is whether the significant effects documented above started before the year of the reform. If that is the case, then the significant impacts in Figure 4 would be a mere continuation of past trends and cannot be attributed to the policy change. To test this, I estimate another version of specification (1) in which I replace the single post-policy indicator ($1\{Post_{st}\}$) with state-specific individual year dummies. For each state, the year before the reform is the base category, so all outcomes are relative to the base category.

Figures 6, 7 and 8 report results from this event study. The estimated coefficients, relative to the base year, were not statistically significant before the reform. Hence, the outcome variable for plants in the treatment and control group followed similar trends before the policy change. However, following the policy, there is a significant increase in the probability with which plants transition around 100 full-time workers in the treated states. Analogously, there is a significant decrease in the probability with which plants transition around 300 full-time workers in the treated states post-policy. There is no statistically significant impact of the policy change on transition probabilities around any other threshold.

4.2 Policy Impact on Plant's Substitution Patterns

The above results suggest that the policy change impacted plant size transitions around the old and new regulatory thresholds. Did it also impact plant's intake of other factors of production on which the regulation does not apply? In particular, this paper considers the policy impact on plants' share of contract workers and their intake of plant & machinery[5][6].

The policy impact on plant's substitution patterns would vary based on its size and the elasticity of substitution between full-time workers, contract workers and capital. Before the policy change, those close to the 100-worker threshold would have been more constrained from employing full-time workers than plants farther away. If contract workers and capital are substitutable with full-time workers, then those close to 100 workers would have been more likely to substitute full-time workers with these factors. After the policy reform, however, we would expect plants around the 100-worker threshold to lower their intake of these factors. Conversely, we would expect plants around the new regulatory threshold of 300 full-time workers to increase their intake of contract workers and capital.

In order to ascertain the policy impact on plant's substitution patterns, I estimate the same difference-in-differences regression as in specification (1). Figures 9 and 10 plot the policy impact, as captured by the coefficient $\delta + \beta_k^{post}$, after estimating equation (1). Results suggest that capital stock and contract labor share of plants decreased (increased) significantly as they approached the old (new) regulatory threshold of 100 (300) full-time workers after the policy. As with transition probabilities, there is no statistically significant impact of the policy change on substitution patterns of plants around any other threshold.

To summarise, using a natural experiment involving the Indian employment protection legislation, I show in this section that this legislation restricts employment growth of plants by lowering their transitions around the regulatory threshold. Since this legislation applies only to full-time workers, I also show that plants use part-time contract workers and capital as margins of adjustment to avoid falling under this law.

However, some questions remain: (1) Under what conditions would a costly size-based rule lead to bunching? Under what conditions would it not lead to bunching? (2) From a policy perspective, it is necessary to understand the driver behind these regulatory costs. As argued earlier, anecdotal evidence suggests that the Indian employment protection legislation is costly for manufacturing plants in two ways. First, plants face significant delays

[5]As described earlier, contract workers are hired by plants through staffing companies on a temporary contractual basis. They are not full-time employees of the plant and they do not fall under the purview of IDA (Bertrand et al. [2021]).

[6]Plant i's capital stock at time t is defined as: (Net closing value of plant and machinery by plant i at time t/Net closing value of all fixed assets by plant i at time t). Plant i's contract labor share at time t is defined as: (No. of contract workers employed by plant i at time t)/(Total number of workers employed by plant i at time t).

in securing government authorization from labor courts thereby increasing the uncertainty that they face during an adverse shock. Second, the frequent inspections that take place to check for compliance imposes costs on plants. Therefore, are these regulatory costs primarily due to the delays in securing government authorization during an adverse shock? Or is it because of the enforcement regime? (3) What are the output and employment consequences of the legislation? To answer these questions, I build a rich quantitative model of plant behavior below. I show that this framework can replicate key features in Indian data, and I estimate the model in Section 6 to conduct policy counterfactuals.

5 Model

Motivated by the Indian institutional context and the reduced-form evidence documented above, I build a dynamic heterogeneous model with multiple factors of production and labor search costs. I now describe the model setup in greater detail.

5.1 Setup

I assume that time is continuous and plants in the economy discount the future at rate r $\geq$ 0. The economy comprises of an endogenous mass of active plants. Plants are engaged in monopolistic competition and each plant produces a unique variety. Hence, the total mass of plants active in the economy at any given point in time determines the total number of varieties that are available in the economy at that time.

Plants in this economy can choose to become active by paying an entry cost worth c_ϵ units of the composite industrial good. Once they become active, they start their operations by drawing a productivity from a productivity distribution F(.). A newly active plant begins it's life with zero workers and zero capital. I now describe the production technology and hiring process of plants that have started their operations.

Production Technology: Let q be the value of output produced by a plant with productivity ϵ. Output q is produced using full-time workers (l_f), contract workers (l_c) and capital goods (k). I allow the production function to exhibit different degrees of substitutability between these three factors of production based on the evidence presented in the previous section. More specifically, output q is produced by aggregating full-time workers, contractual workers and capital goods using a nested CES function.

In the first level of aggregation, as shown in equation (2), contract workers (l_c) and full-time workers (l_f) are combined in a CES function with elasticity of substitution denoted by $v > 1$. Output from the first level of aggregation, h, is then combined with capital goods k in another CES function in equation (3) with elasticity of substitution $\rho > 1$.

$$h = (l_f^{\frac{v-1}{v}} + l_c^{\frac{v-1}{v}})^{\frac{v}{v-1}}. \tag{2}$$

$$m = (h^{\frac{\rho-1}{\rho}} + k^{\frac{\rho-1}{\rho}})^{\frac{\rho}{\rho-1}}. \tag{3}$$

Output of an establishment with productivity ϵ is then given by equation (4) below. Here, $\alpha \in$ (0,1). Thus, output is strictly increasing in ϵ and m, and strictly concave in m. Moreover, for any ϵ, the INADA conditions hold with respect to m: (1) $\lim_{m\to 0} q = \infty$, and (2) $\lim_{m\to\infty} q = 0$.

$$q = \epsilon m^{\alpha} \tag{4}$$

For a plant that produces output as per equation (4), it's sales in equilibrium is given by $R = q^{\frac{\mu-1}{\mu}} D^{\frac{1}{\mu}}$. Here, $\mu > 1$ is the constant elasticity of substitution across the different varieties available in the market and D is an aggregate variable that includes the Dixit-Stiglitz price index and aggregate expenditure in the economy.

Producer Dynamics: I allow productivity of plants to evolve over time. More specifically, following Shimer [2005] and Eaton et al. [2022], I specify an Ehrenfest diffusion process for ϵ by discretizing it's log into $2n_\epsilon + 1$ points with $n_\epsilon \geq 1$ and with each grid point being $\Delta_\epsilon > 0$ apart[7]: $\log(\epsilon) \in \{-n_\epsilon\Delta_\epsilon, -(n_\epsilon - 1)\Delta_\epsilon, ..., 0, ...(n_\epsilon - 1)\Delta_\epsilon, n_\epsilon\Delta_\epsilon\}$.

For a plant with productivity ϵ, productivity shocks arrive at the Poisson rate $\lambda_\epsilon > 0$ and the new productivity ϵ' moves up or down by one grid point as follows:

$$\epsilon' = \begin{cases} \epsilon + \Delta_\epsilon & \text{w.p. } 0.5(1 - \frac{\epsilon}{n_\epsilon\Delta_\epsilon}) = p \\ \epsilon - \Delta_\epsilon & \text{w.p. } 0.5(1 + \frac{\epsilon}{n_\epsilon\Delta_\epsilon}) = 1 - p \end{cases}$$

Here, larger values of p indicate faster mean reversion. In other words, for plants with larger values of the initial productivity shock, their probability of receiving an even higher

[7]As I describe in the subsequent sub-sections, the state space of the establishment's problem is discrete.

productivity shock is lower. Hence, the parameters of the productivity process can be summarized by $(n_\epsilon, \Delta_\epsilon, \lambda_\epsilon)$. I consider the state space of productivity to be discrete because it significantly simplifies computation. Following the discussion in Shimer [2005], it can be shown that as the fineness of the grid (Δ_ϵ) increases, the above diffusion process converges to an Ornstein- Uhlenbeck process with mean 0 of the following form[8]:

$$d\epsilon = -\frac{\lambda_\epsilon}{n_\epsilon}\epsilon dt + \sqrt{\lambda_\epsilon}\Delta_\epsilon dW \tag{5}$$

Here, dW is a standard Wiener process, $\frac{\lambda_\epsilon}{n_\epsilon}$ is the auto-correlation parameter and $\sqrt{\lambda_\epsilon}\Delta_\epsilon$ is the instantaneous standard deviation parameter. As I argue in the next section, I calibrate plant level productivities from the data as a Solow residual from the revenue production function following Hsieh and Klenow [2009]. The plant level productivities thus calculated are then considered to be annual observations generated from the productivity evolution process described above.

Hiring Technology: Plants face different sets of costs while hiring full-time workers, contract workers and capital goods as I explain below. Suppose the regulatory environment is such that plants fall under the purview of labor regulations if their formal workforce l_f exceeds R. Given this regulatory environment, the total wage bill of a plant with l_f formal workers is given by:

$$\begin{cases} \mathrm{w}_f l_f & \text{if } l_f \leq R \\ \mathrm{w}_f(1+\tau_f)l_f & \text{if } l_f > R \end{cases}$$

Here, $R > 0$ is the threshold beyond which regulations apply to plants and w_f is the wage rate per full-time worker that the plant takes as given. $\tau_f \geq 0$ captures the increase in marginal cost of production for a plant with more than l_f full-time workers because of the regulation. As argued in Section 1, $\tau_f \geq 0$ can be interpreted as the costs that plants face because of the frequent inspections that take place to check for their compliance with the Indian EPL.

In addition to the wages that they have to pay full-time workers, the process of hiring them is also costly and time-consuming for plants. In order to hire workers, a plant with productivity ϵ and l_f full-time workers has to choose an effort level or search intensity defined by $\sigma(\epsilon, l_f) \geq 0$. Then, the number of workers that come to work for them is Poisson

[8]The detailed derivation has been provided in Appendix III(A).

distributed with arrival rate $\lambda(\sigma(\epsilon, l_f)) > 0$. If a plant chooses a higher $\sigma(\epsilon, l_f)$, then smaller is the time that they will have to wait on average to get the next worker.

However, plants cannot always choose high levels of $\sigma(\epsilon, l_f)$ because it is costly. Let $C(\sigma, \epsilon, l_f)$ be the cost function associated with search intensity σ for a plant with productivity ϵ and l_f formal workers. Here, C(.) is a twice continuously differentiable, strictly increasing and strictly convex function in search intensity that satisfies $\lim_{\sigma \to +\infty} C_\sigma(\sigma, \epsilon, l_f) = +\infty$ and $C_l(\sigma, \epsilon, l_f) < 0$. The latter inequality implies that larger plants face smaller costs than smaller plants for the same search effort. I assume the following functional form for $C(\sigma, \epsilon, l_f)$ and $\lambda(\sigma(\epsilon, l_f))$.

$$C(\sigma, \epsilon, l_f) = b_\epsilon \sigma^{\eta(\epsilon, l_f)+1} \quad \& \quad \eta(\epsilon, l_f) = \eta_\epsilon (l_f + 1)^{k_\epsilon} \quad \forall (l_f, \epsilon). \tag{6}$$

$$\lambda(\sigma(\epsilon, l_f)) = \sigma(\epsilon, l_f) + p_\epsilon (l_f)^{q_\epsilon}, \qquad p_\epsilon, q_\epsilon \in R. \tag{7}$$

The total wage bill of a plant with l_c contract workers is $w_c(1+\tau_c)l_c$. $w_c > 0$ is the wage rate per contract worker, while $\tau_c \geq 0$ denotes the fee that they have to pay the staffing company or the differential productivity of contract workers in production. Analogously, the total expenditure incurred by a plant with k capital goods is pk, where p denotes the differential productivity of capital goods in production. Plants take w_c, τ_c and p as given. Unlike full-time workers, plants do not face any other costs while hiring contract workers and capital goods.

Firing Costs: Plants face different set of costs while firing full-time workers, contract workers and capital goods as well. Every employed full-time worker l_f quits from a plant with productivity ϵ at a natural rate $\delta_0(\epsilon, l_f) > 0$. On top of that, if the plant wants to fire a full-time worker, then the plant can choose $\delta \geq \delta_0(\epsilon, l_f)$. If these separation rates are independent across full-time workers, then these separation shocks will arrive for a plant with l_f workers at the Poisson rate $\delta l_f > 0$. Thus, for a plant with l_f full-time workers, larger the value of δl_f, smaller is the time that it has to wait on average until they lose a full-time worker. However, plants always cannot choose high values of δ because there are costs associated with it. Let $C(\delta, \epsilon, l_f)$ denote the cost for a plant with l_f full-time workers and productivity ϵ that chooses separation rate δ. I assume the following functional form for $C(\delta, \epsilon, l_f)$.

$$C(\delta, \epsilon, l_f) = \begin{cases} (\delta - \delta_0(\epsilon, l_f))^{\kappa_\epsilon} & \text{if } l_f \leq R \\ \gamma_R (\delta - \delta_0(\epsilon, l_f))^{\kappa_\epsilon} & \text{if } l_f > R \end{cases}$$

Here again, $R > 0$ is the threshold beyond which regulations apply to plants. $\gamma_R \geq 1$ captures the increase in firing cost for plants with more than R workers because of the regulation. $\gamma_R \geq 1$ can be interpreted as the costs that plants face because of the uncertainty that they face when they are hit by an adverse shock. Also, $\delta_0(\epsilon, l_f) = \delta_{0\epsilon}(l_f + 1)^z$ for all (l_f, ϵ).

I assume that plants do not have to pay any firing costs while reducing the size of their contract labor force and capital goods. Since there are no adjustment costs while adjusting contract workers and capital goods, contract workers and capital goods are variable inputs in the production process that gets determined every period, depending on a plant's productivity and inherited full-time workers l_f.

Plant Exit: Every plant gets destroyed at an exogenous Poisson rate $s > 0$. I assume that an exiting plant receives a scrap value of 0.

5.2 Characterization of the Incumbent's Problem

At any given instant, there are two state variables for an incumbent plant: productivity ϵ and l_f full-time workers. Suppose the following processes are independent of each other: (1) the Poisson process that governs the arrival rate of full-time workers, (2) the process that governs the separation rate of full-time workers, (3) the process that determines if the plant gets destroyed or not, and (4) the process that governs the arrival of productivity shocks.

With these, the goal is to derive the plant's problem in continuous time. The choice problem of a plant with productivity ϵ and l_f full-time workers is as follows:

$$\begin{aligned} (r + \lambda(\sigma) + \delta l_f + s + \lambda_\epsilon)V(l_f, \epsilon) = \\ max_{\sigma \geq 0, \delta \geq \delta 0}\{\Pi(l_f, \epsilon) - w_f l_f - C(\sigma, \epsilon, l_f) - C(\delta, \epsilon, l_f) + \delta l_f V(l_f - 1, \epsilon) \\ + \lambda(\sigma)V(l_f + 1, \epsilon) + \lambda_\epsilon p_\epsilon V(l_f, \epsilon + \Delta_\epsilon) \\ + \lambda_\epsilon(1 - p_\epsilon)V(l_f, \epsilon - \Delta_\epsilon)\} \end{aligned} \tag{8}$$

with

$$\Pi(\epsilon, l_f) = max_{l_c, k \geq 0}\{R(\epsilon, l_f) - w_c(1 + \tau_c)l_c - pk\} \tag{9}$$

Equation (8) can be interpreted as follows. Given productivity level ϵ and l_f full-time workers, plant earns flow payoff of $\Pi(\epsilon, l_f) - C(\sigma, l_f) - C(\delta, l_f)$ until the next event occurs. $\Pi(\epsilon, l_f)$ denotes flow profits earned by the plant and is pinned down by the optimal value of

l_c and k that maximises the right hand side of equation (9). Since there are no adjustment costs for l_c and k, the optimal value of l_c and k is such that their marginal product equals their respective factor price.

Time until the next event depends on the Poisson arrival rates s, δl_f, $\lambda(\sigma)$ and λ_ϵ. Larger these arrival rates, faster the next event will occur on average. The next event is either: (1) destruction of the plant (with probability $\frac{s}{s+\delta l_f+\lambda(\sigma)+\lambda_\epsilon}$), or (2) separation of a worker (with probability $\frac{\delta l_f}{s+\delta l_f+\lambda(\sigma)+\lambda_\epsilon}$), or (3) arrival of a new worker (with probability $\frac{\lambda(\sigma)}{s+\delta l_f+\lambda(\sigma)+\lambda_\epsilon}$), or (4) arrival of a productivity shock (with probability $\frac{\lambda_\epsilon}{s+\delta l_f+\lambda(\sigma)+\lambda_\epsilon}$).

While plants take s and λ_ϵ as given, they can expedite the arrival and separation of workers by choosing σ and δ respectively. The optimal search intensity of a plant with productivity ϵ and l_f full-time workers is captured by the policy function $\sigma(l_f, \epsilon)$, and it satisfies the following first order condition:

$$\underbrace{C_\sigma(\sigma(l_f,\epsilon),\epsilon,l_f)}_{\text{Marginal cost of search}} \geq \underbrace{\lambda'(\sigma(l_f,\epsilon))[V(l_f+1,\epsilon)-V(l_f,\epsilon)]}_{\text{Expected marginal benefit from search}} \tag{10}$$

Therefore, the optimal effort that plants put towards hiring at each stage depends on their marginal cost of effort and their expected marginal benefit from search. The latter comprises of the increase in likelihood of obtaining workers from increasing their effort (captured by $\lambda'(\sigma(l_f, \epsilon))$) and increase in their present value of profits from having more workers (captured by $V(l_f + 1, \epsilon) - V(l_f, \epsilon)$). Larger the degree of convexity in the search cost function, lower will be the responsiveness of search effort to expected marginal benefit.

Analogously, the optimal separation rate chosen by a plant with productivity ϵ and l_f full-time workers, $\delta(l_f, \epsilon)$, depends on the marginal firing cost and the expected marginal benefit from firing as captured by the following first order condition:

$$\underbrace{C_\delta(\delta(l_f,\epsilon),\epsilon,l_f)}_{\text{Marginal firing cost}} \geq \underbrace{l_f[V(l_f-1,\epsilon)-V(l_f,\epsilon)]}_{\text{Expected marginal benefit from firing}} \tag{11}$$

Equation (10) offers insights about the life-cycle of an individual plant. To fix ideas, suppose productivity stays constant throughout the life of plants. As a plant is born, it will have a high marginal product of labor. Accordingly, it's expected marginal benefit from search will be higher. With increasing marginal cost of search, it will choose a high level of search and will grow quickly. However, as it grows, it's marginal product of labor falls. Accordingly, it's expected marginal benefit from search will be smaller. Therefore, it will choose a lower level of search and will grow slowly. Due to decreasing returns to labor, after it reaches a certain size, it's optimal search effort would be just so that the workers who quit can be replaced or to just do replacement hiring. Then the plant is said to have reached it's target size. Larger the degree of convexity in the search cost function, smaller is it's target size and longer it will take to reach it's target number of workers.

5.3 Characterization of the Entrant's Problem

At each instant, a large mass of potential entrants decide whether they should enter or not. I assume plants have to pay c_e units of the final good to enter. After paying $c_e > 0$, plant draws it's productivity level (ϵ) from the ergodic distribution $f(\epsilon)$ implied by (5).

A potential entrant decides to become active only if their expected marginal benefit from entry is at least as high as c_e. Therefore, in equilibrium, the below free entry condition should hold at all times:

$$\int_{\epsilon} V(0,\epsilon) f(\epsilon) d\epsilon \leq c_e \tag{12}$$

The left hand side of equation (12) denotes the expected marginal benefit from entering for a potential entrant and $V(0,\epsilon)$ is as per equation (8) above. If equation (12) holds with strict inequality, then mass of entrants $M_e = 0$ in equilibrium. However, if equation (12) holds with equality, then mass of entrants $M_e > 0$ in equilibrium. To put it differently, equation (12) pins down the mass of entrants in equilibrium whenever that mass is positive. Larger the mass of entrants, larger will be the measure of active plants in the economy, smaller will be the aggregate price index which makes entry less valuable for potential entrants. Analogously, smaller the mass of entrants, smaller will be the measure of active plants in the economy, larger will be the aggregate price index in the economy which makes entry more valuable for potential entrants.

5.4 Steady State Plant Size Distribution

Based on $V(l_f,\epsilon)$, $\sigma(l_f,\epsilon)$, $\delta(l_f,\epsilon)$ and the mass of entrants calculated above, I derive the steady state plant size distribution. Here, size is defined in terms of the number of full-time workers employed by the plant. Let M(l_f,ϵ) denote the mass of plants at (l_f,ϵ) for $l_f \in$

{0,1,..,$\bar{L}$} and ϵ. The law of motion describing M(l_f,ϵ), for $l_f \in \{1, 2, .., \bar{L}\}$ and ϵ is:

$$\dot{M(l_f, \epsilon)} = \overbrace{\begin{array}{l}[M(l_f-1,\epsilon)\lambda(\sigma(l_f-1,\epsilon)) + M(l_f+1,\epsilon)\delta(l_f+1) + ... \\ ... \text{ M } (\mathrm{l}_f, \epsilon - \Delta_\epsilon)\lambda_\epsilon p_{\epsilon-\Delta\epsilon} + M(l_f, \epsilon+\Delta_\epsilon)\lambda_\epsilon(1-p_{\epsilon+\Delta\epsilon})]\end{array}}^{\text{Mass of plants that enter state } (l_f,\epsilon)} - \underbrace{[M(l_f,\epsilon)(\lambda(\sigma(l_f,\epsilon)) + \delta l_f + s + \lambda_\epsilon)]}_{\text{Mass of plants that leave state } (l_f,\epsilon)}$$

The mass of plants that enter state (ϵ,l_f) includes: (1) those from state (ϵ,$l_f - 1$) that get a worker with hazard rate $\lambda(\sigma(l_f-1,\epsilon))$, (2) those from state ($\epsilon$,$l_f+1$) that lose a worker because of exogenous or endogenous separation with hazard rate $\delta(\epsilon, l_f)(l_f+1)$, (3) those from state ($\epsilon$ - Δ_ϵ,l_f) that received a higher productivity shock with hazard rate $\lambda_\epsilon p_{\epsilon-\Delta\epsilon}$, and (4) those from state ($\epsilon + \Delta_\epsilon$,$l_f$) that received an adverse productivity shock with hazard rate $\lambda_\epsilon(1-p_{\epsilon+\Delta\epsilon})$. Similarly, the mass of plants that leave state (ϵ,l_f) includes those from state (ϵ,l_f) that either get a worker, lose a worker due to separation, receive a productivity shock or get destroyed with hazard rates $\lambda(\sigma(l_f,\epsilon)), \delta(\epsilon)l_f, \lambda_\epsilon$ and s_ϵ respectively.

The mass of plants with productivity ϵ and 0 workers evolves according to the following. The mass of plants that enter state (ϵ,0) includes: (1) those from state (ϵ,1) that lose a worker because of exogenous or endogenous separation with hazard rate $\delta(\epsilon, 1)$, (2) those from state (ϵ - Δ_ϵ,0) that receive a higher productivity shock with hazard rate $\lambda_\epsilon p_{\epsilon-\Delta\epsilon}$, (3) those from state ($\epsilon + \Delta_\epsilon$,0) that receive an adverse productivity shock with hazard rate $\lambda_\epsilon(1-p_{\epsilon+\Delta\epsilon})$, and (4) those entrants that received a productivity draw ϵ from the ergodic distribution f(ϵ). Similarly, the mass of plants that leave this state includes those that either get a worker or receive a different productivity shock with hazard rates $\lambda(\sigma(0,\epsilon))$ and λ_ϵ respectively.

$$\dot{M(0, \epsilon)} = \overbrace{\begin{array}{l}[M(1,\epsilon)\delta(\epsilon,1) + M_e f(\epsilon) + ... \\ ... \text{ M } (0,\epsilon - \Delta_\epsilon)\lambda_\epsilon p_{\epsilon-\Delta\epsilon} + M(0, \epsilon+\Delta_\epsilon)\lambda_\epsilon(1-p_{\epsilon+\Delta\epsilon})]\end{array}}^{\text{Mass of plants that enter state } (0,\epsilon)} - \underbrace{[M(0,\epsilon)(\lambda(\sigma(0,\epsilon)) + \lambda_\epsilon)]}_{\text{Mass of plants that leave state } (0,\epsilon)}$$

At steady state, by definition, $\dot{M(l_f, \epsilon)} = 0$ for $l_f \in$ {0,1,..,$\bar{L_f}$} and any ϵ. By solving the above system of $\bar{L_f}+1$ equations for $\bar{L_f}+1$ unknowns, we can derive closed form solutions for the steady state distribution of M(l_f,ϵ) for $l_f \in \{0,1,..,\bar{L_f}\}$ and any ϵ.

$$M(\bar{L},\epsilon) = \frac{M(\bar{L}-1,\epsilon)\lambda(\sigma(\bar{L}-1,\epsilon)) + M(\bar{L},\epsilon-\Delta_\epsilon)\lambda_\epsilon p_{\epsilon-\Delta\epsilon} + M(\bar{L},\epsilon+\Delta_\epsilon)\lambda_\epsilon(1-p_{\epsilon+\Delta\epsilon})}{s+\delta\bar{L}+\lambda(\sigma(\bar{L},\epsilon))+\lambda_\epsilon} \tag{13}$$

$$M(l_f,\epsilon) = \frac{M(l_f-1,\epsilon)\lambda(\sigma(l_f-1,\epsilon)) + \delta(l_f+1)M(l_f+1,\epsilon) + M(l_f,\epsilon-\Delta_\epsilon)\lambda_\epsilon p_{\epsilon-\Delta\epsilon} + M(l_f,\epsilon+\Delta_\epsilon)\lambda_\epsilon(1-}{s+\delta l_f+\lambda(\sigma(l_f,\epsilon))+\lambda_\epsilon} \tag{14}$$

$$M(0,\epsilon) = \frac{\delta M(1,\epsilon) + M_e f(\epsilon) + M(0,\epsilon-\Delta_\epsilon)\lambda_\epsilon p_{\epsilon-\Delta\epsilon} + M(0,\epsilon+\Delta_\epsilon)\lambda_\epsilon(1-p_{\epsilon+\Delta\epsilon})}{s+\lambda(\sigma(0,\epsilon))+\lambda_\epsilon} \tag{15}$$

The system of equations characterised by equations (13)-(15) imply the following. Given any (l_f,ϵ) pair, if plants lower $\sigma(l_f,\epsilon)$, then they will spend more time in (l_f,ϵ) on average. As a result, M(l_f,ϵ) at steady state will be higher.

5.5 Discussion of Mechanisms

Impact of the regulation on transitions and substitution patterns: How do these size-based rules impact transition probabilities and substitution patterns of plants around R workers? Here, the benchmark case is one in which there are no size-based rules ($\tau_R = 0$, $\gamma_R = 1$). With $\tau_R > 0$, $\gamma_R > 1$ and other things remaining the same, the optimal search effort by plants around R workers falls relative to the benchmark case. This is because the increase in present value of profits from having more workers is lower for plants around R workers because of additional regulatory costs. Therefore, other things remaining equal, two things happen when there are size-based rules relative to the benchmark case. First, plants around R workers spend more time on average in their respective states. Second, their probability of getting an additional worker and moving ahead falls. Thus, transition probabilities around the R worker threshold falls. Larger the magnitude of regulatory costs, larger will be the decline in search effort by plants around the regulatory threshold and larger will be the decline in transition probabilities around the regulatory threshold.

As a result, at the end of say one year, their full-time labor intake will be lower than what they would have had if the regulation had not been there. Depending on the magnitude of the elasticity of substitution between full-time workers and contract workers, their share of contract workers will be accordingly higher. Similarly, depending on the magnitude of the elasticity of substitution between full-time workers and capital, their capital stock will also be accordingly higher. Larger the magnitudes of these elasticity parameters, larger will be the contract labor share and capital stock of plants around the regulatory threshold of R workers.

I illustrate these with the help of an example. I first set τ_ϵ to be 0, γ_ϵ to be 1 and other parameters fixed at arbitrary values as per table 8. I then calculate the policy function and the steady state plant size distribution implied by these parameter values from (10)-(11) and (13)-(15) respectively. Using this steady state distribution as the initial condition, I simulate transition probabilities within a year for 10,000 plants. I then repeat the above exercise by varying τ_ϵ and keeping all other parameters fixed as before. The detailed algorithm is given in Appendix III(E).

Figure 11 shows how search effort and transition probabilities of plants around the regulatory threshold varies with τ_R. As τ_R increases, the search intensity of plants when they reach 100 workers and transition probability of plants around the 100 worker threshold steadily falls. Although the below figure has been drawn for $\gamma_R = 1$, this relationship holds for any $\gamma_R \geq 1$. There exists a similar relationship between transition probabilities around the regulatory threshold and γ_R given any $\tau_R > 0$.

For a given level of regulatory costs, Figure 12 plots the contract labor share by plants around 100 workers in steady state, for different values of the elasticity of substitution between full-time and contract workers. Figure 12 suggests that as the elasticity of substitution between full-time and contract workers increases, the average contract labor share is higher for the same magnitude of regulatory costs.

Conditions for Bunching and No Bunching: From above, one can see that the steady state distribution may or may not exhibit bunching in response to costly size based rules and this depends on the marginal adjustment cost of search. As discussed before, consider plants around the regulatory threshold of 100 full-time workers that slow down their search. With high separation rate of full-time workers, plants are more likely to fall back instead of moving forward. If the marginal adjustment cost of search effort is low, then plants will be able to adjust their search intensity accordingly and replace their lost workers quickly. Therefore, plants will stay around 100 workers and this is bunching. On the other hand, if the marginal adjustment cost of search effort is high, then plants will not be able to adjust their search intensity much in response to exogenous shocks. They are more likely to stay below 100 workers because they will not be able to replace their lost workers quickly.

Figure 13 confirms the intuition described above. In the steady state distribution with regulation, when search cost convexity and exogenous separation rate are both low, there is a spike in the density of plants at the 100 worker threshold because they slow down their search effort in response to the regulation. However, as we increase both these parameters, the magnitude of the spike falls even when regulatory costs are the same as before. For higher values of the search cost convexity and the exogenous separation rate, the spike in the density of plants at 100 workers is almost non-existent.

6 Taking the Model to the Data

In order to quantify the magnitude of regulatory costs and to recover aggregate implications of the regulation, I fit the above framework to plant level panel data from the treated states. Recall that the regulatory threshold was 100 formal workers prior to the reform and I assume that the establishment size distribution of treated states during the year before the reform was in steady state. The policy reform that was implemented in these states increased the regulatory threshold from 100 to 300 formal workers. The causal impacts that were derived in Section 4 above are informative about how transition probabilities, contract labor share and capital stock of establishments in the treated states around 100 workers changed because of the policy reform.

Now coming to the model, suppose regulatory threshold (R) is 100 workers. I proceed as follows.

Elasticities of Substitution: I first back out elasticities of substitution between formal labor and contract labor (v), and between labor and capital (ρ). From the discussion above, recall flow profits of an establishment with productivity ϵ and l_f formal workers is given by:

$$\Pi(\epsilon, l_f) = max_{l_c,k\geq 0}\{\epsilon m^{\alpha} - w_c l_c - pk\} \tag{16}$$

Since contract labor and capital are assumed to be static inputs in production, their marginal product equals their factor cost in equilibrium. Hence, taking the first order condition of (16) with respect to l_c and k and some algebraic manipulation yields the following 2 equations in ρ and v.

$$log(\frac{\alpha\epsilon m^{\alpha} - w_c l_c - pk}{w_c l_c}) = \frac{v-1}{v} log(\frac{l_f}{l_c}) \tag{17}$$

$$log(\frac{\alpha\epsilon m^{\alpha}-pk}{pk})=\frac{\rho-1}{\rho}log(\frac{[l_f^{\frac{v-1}{v}}+l_c^{\frac{v-1}{v}}]^{\frac{v}{v-1}}}{k}) \tag{18}$$

In equation (17), equilibrium wage of formal workers relative to contract workers is expressed as a function of their relative quantities. Similarly, in equation (18), equilibrium relative price of labor and capital is expressed as a function of their relative quantities. Following Krusell et al. [2000] and Parro [2013], v is identified from the percentage change in equilibrium relative wage of formal and contract workers, due to a 1 percentage change in their relative quantities. Column 1 of table 4 reports the result from estimating equation (20) with OLS and without any controls. According to column 1, the elasticity of substitution between formal and contract workers is 2.7. However, this estimate of v could be biased because the establishment's optimal choice of contract labor may be correlated with unobservables. For instance, there could be a third factor, such as an industry level or a state level demand shock, that impacts both relative wages and relative quantities of formal and contract workers. I account for this simultaneity issue by comparing establishments within the same industry-year and state-year cell. As long as these unobservables affect plants within the same industry-year and state-year cells equally, v will not be biased by these aggregate shocks. Columns (2)-(4) of table 4 report estimates of v after controlling for these confounding factors at the aggregate level. Finally, column 5 of table 4 includes establishment fixed effects, thereby estimating within-establishment relationship between relative wages and relative quantities. As argued in Ulyssea [2018], if there is positive assortative matching in the economy, establishment fixed effects will control for skill differences in workers across establishments. Even after accounting for a wide range of fixed effects in columns (2)-(5), the estimate of v is fairly stable at 2.7. This estimate of the elasticity of substitution between formal and contract workers in the Indian context is similar in magnitude to the elasticity of substitution between formal and informal labor in some of the Latin American countries.

I identify the elasticity of substitution between labor and capital analogously from equation (18). Column 1 of table 5 reports results from estimating equation (18) with OLS and without any additional controls. Column 1 estimates the elasticity of substitution between labor and capital as 3.3. However, this estimate could also be biased due to the simultaneity issues discussed above. Columns (2)-(5) of table 5 indicate that the elasticity of substitution between labor and capital in the Indian context ranges between 3.5 and 4.5 after accounting for the aforementioned simultaneity issues.

Productivity and Productivity Evolution Parameters: I now proceed to construct estimates of productivity at the establishment level. In doing this, I follow the mis-allocation literature, in particular Hsieh and Klenow [2009]. With elasticities of substitution estimated from above, and by utilizing the observed choices of l_f, l_c and k, I recover m using equations (2) and (3). With this estimated value of m, I back out establishment productivity as a Solow residual from the production function described in equation (4). In other words, log productivity of establishment i in industry j at time t is calculated as:

$$log(\epsilon_{ijt}) = log(Y_{ijt}) - \alpha_j log(m_{ijt}) \tag{19}$$

Here, ϵ denotes productivity, Y denotes log real sales and α_j denotes the median of real expenditure on labor and capital goods over real sales for all establishments in industry j [9]. As shown in the Appendix, establishment productivities follow a log normal distribution as in Hsieh and Klenow [2009].

I consider establishment level productivities from above to be annual observations generated from the Ornstein-Uhlenbeck process described in equation (5). By assuming that productivities follow this process, the next step is to back out the rate of persistence of productivity levels over time and the standard deviation of productivity shocks. The empirical specification to estimate these parameters is given by:

$$\epsilon_{ijt} - \epsilon_{ijt-1} = -\frac{\lambda_\epsilon}{n_\epsilon}\epsilon_{ijt-1} + \sqrt{\lambda_\epsilon}\Delta_\epsilon[W(t) - W(t-1)] + \alpha_j + \gamma_t + \omega_{ijt} \tag{20}$$

Here, ϵ_{ijt} denotes productivity of establishment i at time t, α_j and γ_t are industry and year fixed effects respectively. On estimating equation (20) with OLS, the coefficient on ϵ_{ijt-1} pins down the rate of persistence of productivity levels, after accounting for time-invariant industry level characteristics and aggregate shocks. The standard deviation of error terms from this regression pins down the instantaneous standard deviation of productivity shocks, which is given by $\sqrt{\lambda_\epsilon}\Delta_\epsilon$. Therefore, after fixing the number of grid points (n_ϵ), I back out λ_ϵ and Δ_ϵ.

Other Parameters: The final step involves estimating the remaining policy invariant parameters and the magnitude of regulatory costs given in $\Omega = \{s_\epsilon, \eta_\epsilon, k_\epsilon, b_\epsilon, p_\epsilon, q_\epsilon, \delta_\epsilon, \tau_\epsilon, \gamma_\epsilon\}$.

[9] Productivity is not calculated for establishments with missing values of sales/ wage bill/ expenditure on materials/book value of cap equipment. I also calculate 99.9 and 0.1 percentile of the overall productivity distribution and drop those < 0.1 percentile and > 99.9 percentile.

Ω is chosen to minimize the following objective function.

$$G(\Omega) = [\hat{m} - m(\Omega)]' W [\hat{m} - m(\Omega)] \tag{21}$$

Here, $\hat{m}$ is the vector of empirical moments, $m(\Omega)$ are the model counterparts of $\hat{m}$ and W is the identity matrix. I now describe the construction of moment conditions in $G(\Omega)$. I fix arbitrary values of parameters in Ω and I compute the steady state establishment size distribution from equations (13)-(15). I then calculate two objects: (1) share of establishments belonging to different bins of formal workers in steady state, and (2) average share of contract workers by plants belonging to different bins of formal workers in steady state. These are then matched with their empirical counterparts from the year before the reform to form the first set of moment conditions.

I then draw employment and productivity levels of 10,000 establishments in such a way that the overall establishment size distribution matches the steady state distribution calculated previously. Using this steady state distribution as the initial distribution, I simulate transition probabilities around the regulatory threshold, contract labor share and capital labor ratio under two scenarios: (1) if the regulatory cutoff had stayed at 100 formal workers, (2) and if the regulatory cutoff is increased to 300 formal workers. The difference in outcome variables between scenarios (2) and (1) is then matched with their respective reduced form moments to form the second set of moment conditions.

Identification: While these moment conditions are jointly estimated, some moments are particularly informative about some parameters. For instance, based on the discussion in Section 5.8, changes in transition probabilities, contract labor share and capital labor ratio of establishments around the 100 worker threshold due to the reform are informative about the magnitude of regulatory costs $(\tau_{100}, \gamma_{100})$.

Similarly, the share of establishments belonging to different size bins in steady state is informative about some of the policy invariant parameters. For instance, the mass of establishments belonging to the smaller size bins in steady state is informative about the search cost convexity parameter $\eta(\epsilon, l_f)$ and the exogenous separation rate between the establishment and workers $\delta_0(\epsilon, l_f)$. The intuition behind this is as follows. $\eta(\epsilon, l_f)$ captures the degree of responsiveness of an establishment's search effort to it's expected marginal benefit from search. For the same expected marginal benefit from search, larger the value of $\eta(\epsilon, l_f)$, lower will be the optimal search effort put in by the establishment. As a result, establishments will take longer on average to get new workers and move forward. Thus, larger the search cost convexity parameter, larger will be the mass of establishments belonging to the smaller size bins in steady state. Analogously, if $\delta_0(\epsilon, l_f)$ is very high, then establishments with any given number of formal workers will have faster attrition on average. As a result, there will be a large mass of establishments belonging to the smaller size bins in steady state. As argued in Section 5.8, $\delta_0(\epsilon, l_f)$ and $\eta(\epsilon, l_f)$ also discipline whether or not establishments bunch at the regulatory threshold in steady state.

The share of establishments belonging to the larger size bins in steady state is informative about how the search cost convexity parameter varies with l_f. In other words, it is informative about k_ϵ. If k_ϵ is higher, then establishments with a large number of workers will face lower costs than their smaller counterparts for the same search effort. In this case, larger establishments will be able to afford higher levels of search, thereby lowering the time that they take on average to get new workers and to move forward.

Targeted Moments: Table 6 in the Appendix lists the set of moments that were used to estimate parameters in Ω in equation (24). Model predictions of these moments match closely with their data counterparts. In particular, the pre-policy distribution of formal workers implied by the model, as shown in Panel A of table 6, do not exhibit bunching at the old regulatory threshold of 100 workers as in the data.

Non Targeted Moments: The estimated model is also consistent with establishment level data that was not directly targeted by estimation. Table 7 in the Appendix lists the share of output contributed by plants belonging to different bins of formal workers in steady state. Model predictions of these moments also match fairly well with data. In particular, the model replicates the fall in output shares by plants belonging to 101-162 formal workers and its subsequent rise reasonably well. It must be noted that only moments pertaining to establishment's intake of formal workers and contract workers were targeted in estimation. In other words, the level of capital stock and intermediate input usage of plants were not explicitly targeted in estimation of parameters. As a result, there is no reason for the estimated model to necessarily fit these moments.

Parameter Estimates: To rationalize the policy driven change in transition probabilities, contract labor share and capital labor ratio that we see in panel C of table 6, IDA should have increased the unit labor cost of establishments with 100+ full time workers by 7% prior to the policy change. Regulatory costs due to IDA seem to be primarily driven by increase in marginal cost of production of plants after they employ more than 100 full time workers. This is policy relevant because the cost burden of this legislation on establishments can be reduced if we reduce the harassment bribery associated with labor inspections.

Aggregate Consequences of the Legislation: Finally, using the above parameter estimates, I calculate the aggregate output and aggregate full-time employment of the legislation. The estimated model implies that removing the burden of this legislation would increase aggregate full-time employment and output in the treated states by 9% and 4%, respectively, in the long run. Removing this legislation impacts aggregate outcomes through two channels. First, it would allow more plants to enter by increasing the value of entry. Second, these heterogeneous plants hire more full-time workers, increasing aggregate output.

7 Concluding Remarks

In this paper, I explore how size-based regulations, those that apply to firms after their employment exceeds a specific threshold, shape firm outcomes. I study this using the Indian employment protection legislation (EPL), which requires manufacturing plants with 100+ full-time workers to obtain government authorization and provide severance pay before firing even one worker. There are two parts to my paper. In the first part, using a natural experiment, I document causal evidence that EPL restricts the employment growth of plants by lowering their transitions from below to above 100 full-time workers. Since this legislation applies only to full-time workers, I also find that plants use part-time workers and capital as margins of adjustment to avoid falling under its purview. In the second part of my paper, I build and estimate a rich quantitative model of firm behavior to conduct policy counterfactuals. My counterfactuals suggest that overall output and employment of plants would increase significantly if we remove the legislation altogether.

I view this work as a methodological contribution to the literature on how to evaluate size-based rules. Current work gauges such rules' impact based on whether firms bunch at the threshold, with more bunching taken to imply higher regulatory costs. However, there is no bunching by Indian plants around 100 workers, and my paper shows that these rules can be highly restrictive for firms even with no bunching. My quantitative framework rationalizes why firms would not always bunch with costly size-based regulations and argues that these rules should be evaluated based on how they impact firm transitions around the threshold.

These results are also very policy-relevant because it offers insights into what levers to pull to reduce these regulatory costs. EPL is known to impose costs on plants through two channels. First, plant owners face considerable delays in securing government authorization, thereby increasing uncertainty during an adverse shock. Second, government officials responsible for enforcement often exploit the arbitrary and antiquated nature of this law to tighten the administrative burden that plants face and to extract bribes. Through the lens of the model, I quantify each channel's importance.

Going forward, I plan to extend this paper as follows. At present, I have only focused on the regulatory impact on plants. I intend to place this framework in a general equilibrium setting by building on existing work by Coşar et al. [2016] and Dix-Carneiro et al. [2021]. With a general equilibrium framework, I aim to quantify the legislation's impact on workers' job stability and, hence, overall welfare. I also aim to use the model to shed more light on the wage determination of full-time and contract workers.

8 Appendix

I: Institutional Background

Table 1: Percentage Distribution of Cases Pending before Labor Courts in India as of October 2020

	All India
0 to 1 year	35.24%
1 to 3 years	27.75%
3 to 5 years	15.78%
5 to 10 years	13.38%
10 to 20 years	6.90%
20 to 30 years	0.93%
30+ years	0.02%
Total	146660

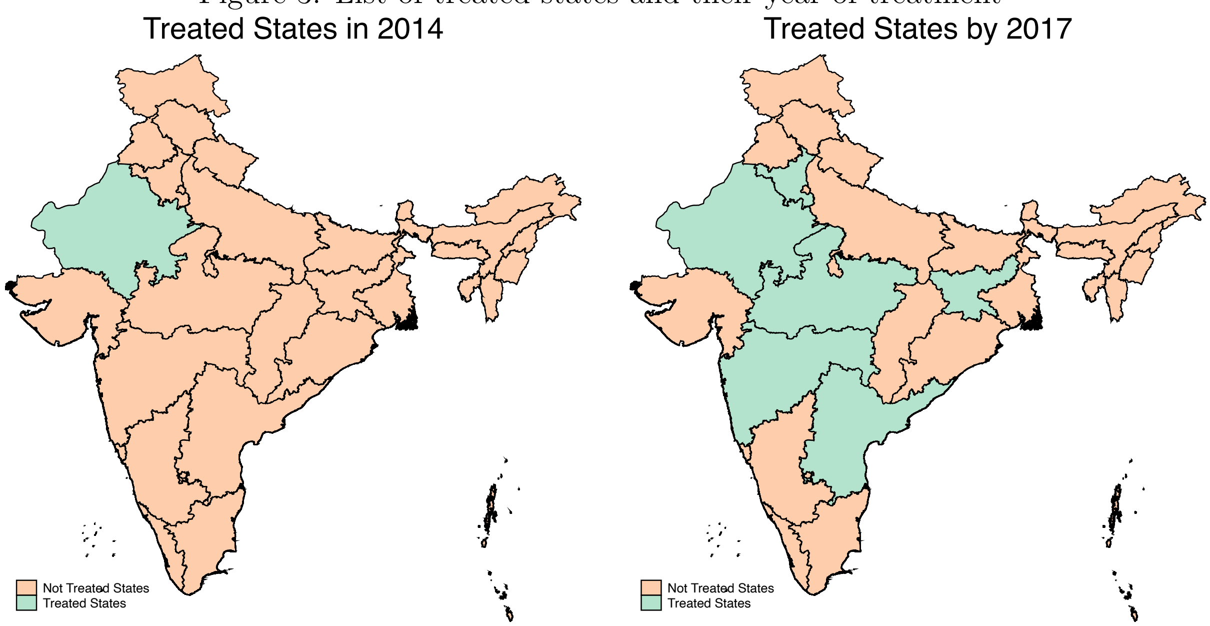
Source: National Judicial Data Grid, Shyam Sundar (2020)

Table 2: Pendency of Labor Cases in High Courts as of December 2013

	Pendency of Cases as on Dec 2013	% Share
Allahabad	10,43,398	23.38%
Madras	5,57,479	12.49%
Bombay	3,49,837	7.84%
Rajasthan	3,07,640	6.89%
Calcutta	2,80,006	6.27%
Punjab-Haryana	2,62,760	5.89%
Madhya Pradesh	2,61,611	5.86%
Andhra Pradesh	2,32,459	5.21%
Orissa	2,06,822	4.63%
Karnataka	1,96,972	4.41%
Kerala	1,32,159	2.96%
Patna	1,32,155	2.96%
Jammu-Kashmir	93,038	2.08%
Gujarat	91,953	2.06%
Jharkhand	72,958	1.63%
Delhi	64,652	1.45%
Himachal Pradesh	60,073	1.35%
Chhattisgarh	44,139	0.99%
Guwahati	40,912	0.92%
Uttarakhand	20,686	0.46%
Tripura+Manipur+Meghalaya	10,876	0.24%
Sikkim	120	0.00%
Total	44,62,705	100%

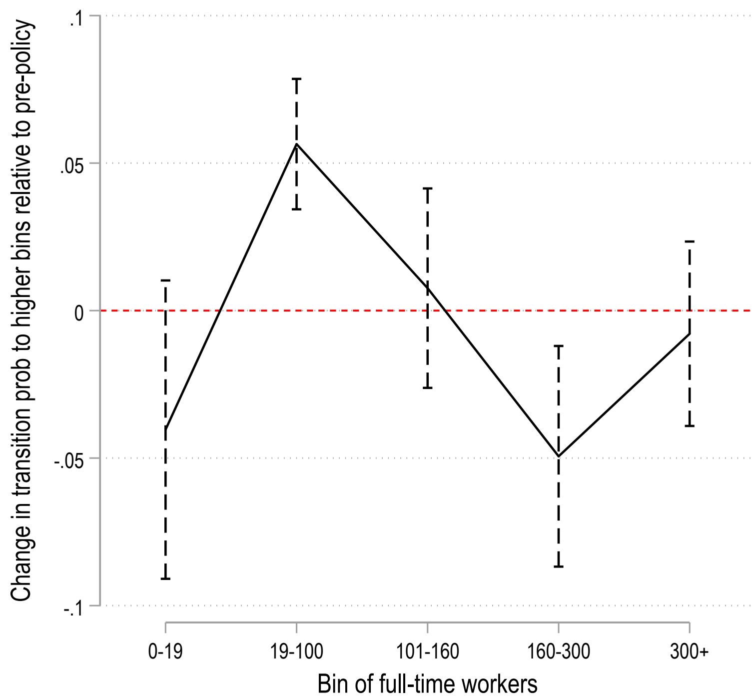
Source: Lok Sabha Unstarred Question No. 812 (November 24, 2014); Shyam Sundar (2020)

Figure 3: List of treated states and their year of treatment



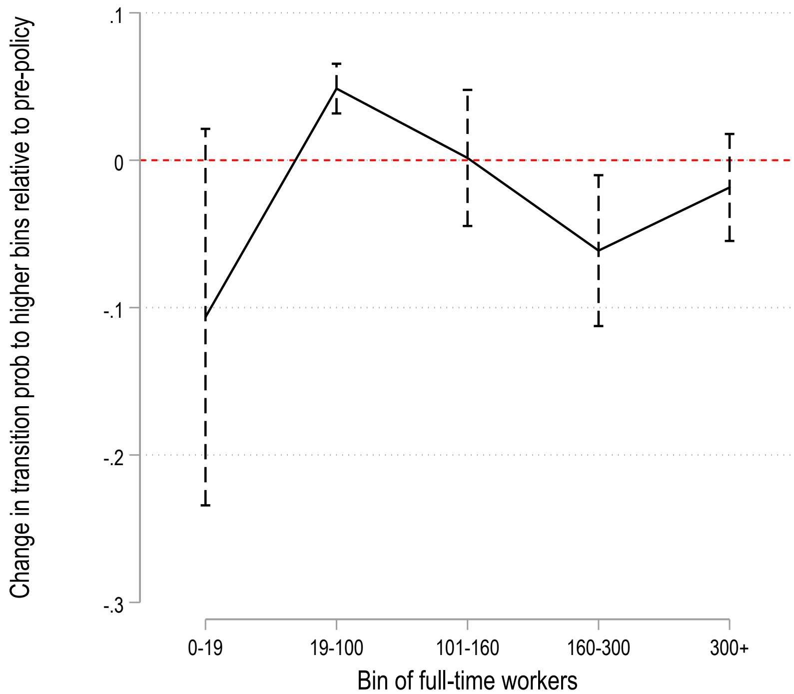
II: Reduced Form Results

Figure 4: Policy Impact on Transitions of Plants



Note: Each observation is at the plant-year level. Plants in each year are divided into bins based on their full-time employment: 0-19, 19-100, 101-160, 160-300 and 300+. The dependent variable is a dummy that takes one if the plant moves to a higher bin in the subsequent year. Bins 19-100 and 160-300 are policy relevant, while others serve as placebos. The figure reports coefficients and 95% confidence intervals on post-by-bin interactions from a regression of the dependent variable on post, bin, post-by-bin interactions and controls. Post is a dummy that varies by state and year: it takes one if the state changed the regulatory threshold to 300 full-time workers by the year. Bin 0-19 is the base category in the regression. Controls include state, industry, year, industry*year*bin fixed effects. Sample includes all major Indian states and covers 1999-2018.

Figure 5: Policy Impact on Transitions of Plants (with Plant Fixed Effects)



Note: Each observation is at the plant-year level. Plants in each year are divided into bins based on their full-time employment: 0-19, 19-100, 101-160, 160-300 and 300+. The dependent variable is a dummy that takes one if the plant moves to a higher bin in the subsequent year. Bins 19-100 and 160-300 are policy relevant, while others serve as placebos. The figure reports coefficients and 95% confidence intervals on post-by-bin interactions from a regression of the dependent variable on post, bin, post-by-bin interactions and controls. Post is a dummy that varies by state and year: it takes one if the state changed the regulatory threshold to 300 full-time workers by the year. Bin 0-19 is the base category in the regression. Controls include plant, year, industry*year*bin fixed effects. Sample includes all major Indian states and covers 1999-2018.

Event Study Graphs: Policy Impact on Transitions of Plants

Figure 6: Bin: 19-100 workers

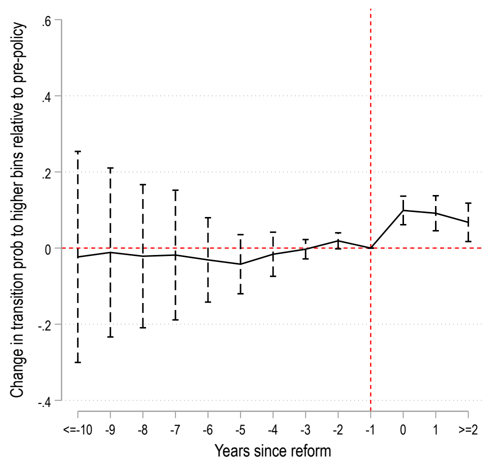


Figure 7: Bin: 101-160 workers

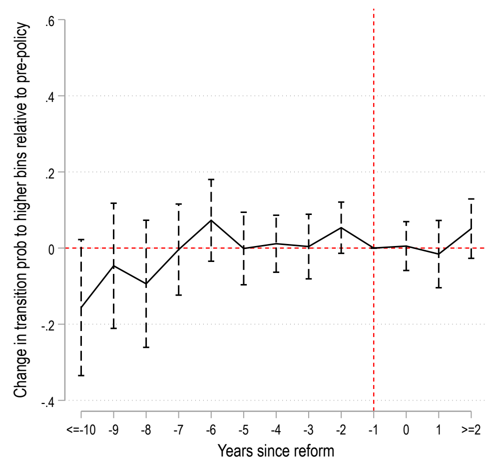
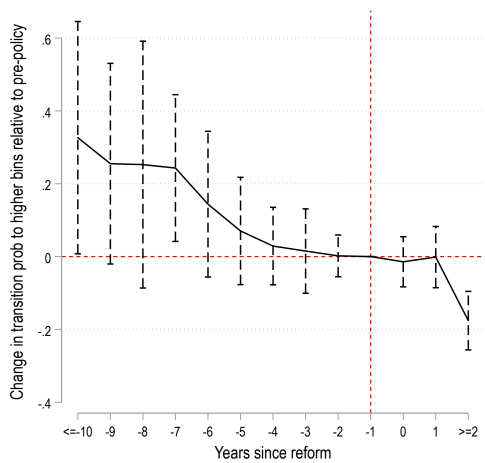
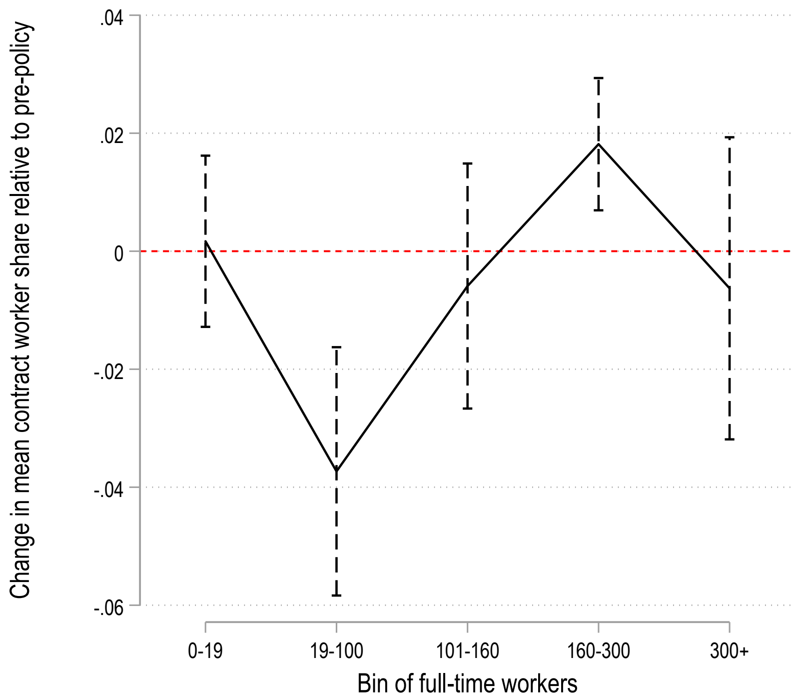


Figure 8: Bin: 161-300 workers



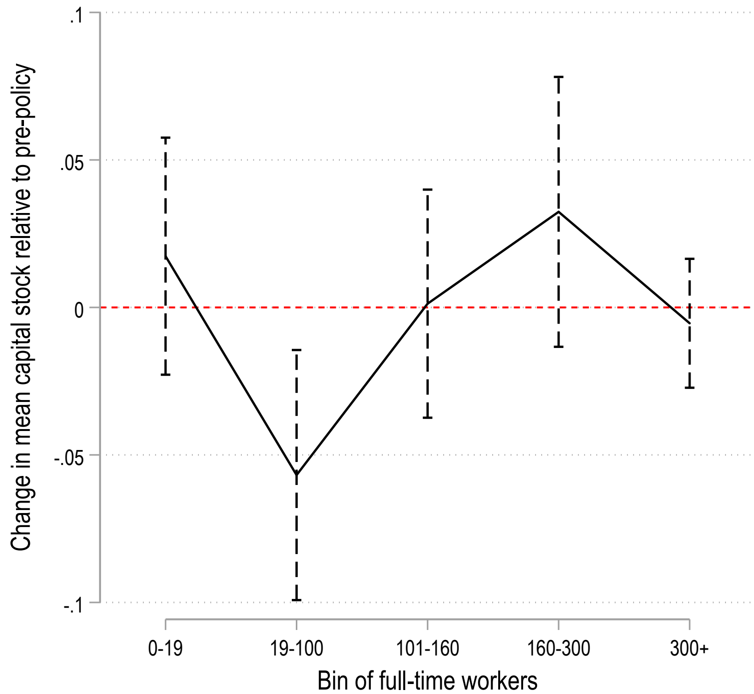
Note: Each observation is at the plant-year level. Plants in each year are divided into bins based on their full-time employment: 0-19, 19-100, 101-160, 160-300 and 300+. The dependent variable is a dummy that takes one if the plant moves to a higher bin in the subsequent year. Bins 19-100 and 160-300 are policy relevant, while others serve as placebos. The dependent variable is regressed on years since policy change, bins, interactions of bins and years since policy change, and control variables. For each bin, the above figure reports coefficients and 95% confidence intervals of its interaction with years since policy change. One year before the policy change is taken to be the base year and bin 0-19 is the base category in the regression. Controls include industry, state, year, industry*year*bin fixed effects. Sample includes all major Indian states and covers 1999-2018.

Figure 9: Policy Impact on Mean Contract Worker Share of Plants



Note: Each observation is at the plant-year level. Plants in each year are divided into bins based on their full-time employment: 0-19, 19-100, 101-160, 160-300 and 300+. The dependent variable is the share of contract workers employed by the plant in the subsequent year. Bins 19-100 and 160-300 are policy relevant, while others serve as placebos. The figure reports coefficients and 95% confidence intervals on post by bin interactions from a regression of the dependent variable on post, bin and post by bin interactions. Post is a dummy that varies by state and year: it takes one if the state changed the regulatory threshold to 300 full-time workers by the year. Bin 0-19 is the base category in the regression. Controls include industry, state, year, industry*year*bin fixed effects. Sample includes all major Indian states and covers 1999-2018.

Figure 10: Policy Impact on Mean Capital Stock of Plants



Note: Each observation is at the plant-year level. Plants in each year are divided into bins based on their full-time employment: 0-19, 19-100, 101-160, 160-300 and 300+. The dependent variable is plants' stock of plant and machinery, as a fraction of it's total fixed assets, in the subsequent year. Bins 19-100 and 160-300 are policy relevant, while others serve as placebos. The figure reports coefficients and 95% confidence intervals on post by bin interactions from a regression of the dependent variable on post, bin and post by bin interactions. Post is a dummy that varies by state and year: it takes one if the state changed the regulatory threshold to 300 full-time workers by the year. Bin 0-19 is the base category in the regression. Controls include industry, state, year, industry*year*bin fixed effects. Sample includes all major Indian states and covers 1999-2018.

III: Model

A: Details about the Ornstein Uhlenbeck Process

Derivation: For any given $\epsilon(t)$, what is $\epsilon(t+h)$ where h is an arbitrarily short period of time?

$$\epsilon(t+h) = \begin{cases} \epsilon(t) & \text{w.p. } 1 - h\lambda_\epsilon \\ \epsilon(t) + \Delta_\epsilon & \text{w.p. } h\lambda_\epsilon * 0.5\frac{1-\epsilon(t)}{n_\epsilon \Delta_\epsilon} \\ \epsilon(t) - \Delta_\epsilon & \text{w.p. } h\lambda_\epsilon * 0.5\frac{1+\epsilon(t)}{n_\epsilon \Delta_\epsilon} \end{cases}$$

Thus, $[\epsilon(t+h) - \epsilon_t|\epsilon_t]$ is a random variable. From the above pmf, we can derive $E[\epsilon(t+h) - \epsilon_t|\epsilon_t]$ and $Var[\epsilon(t+h) - \epsilon_t|\epsilon_t]$.

$E[\epsilon(t+h) - \epsilon_t|\epsilon_t] = -h\epsilon_t \frac{\lambda_\epsilon}{n_\epsilon}$.

$Var[\epsilon(t+h) - \epsilon_t|\epsilon_t] = h\lambda_\epsilon(\Delta_\epsilon)^2$.

Putting these together, define:

$$dx = \frac{d\epsilon - E[d\epsilon|\epsilon]}{Sd[d\epsilon|\epsilon]} \tag{22}$$

$$d\epsilon = -\frac{\lambda_\epsilon}{n_\epsilon}\epsilon dt + \sqrt{\lambda_\epsilon}\Delta_\epsilon dx \tag{23}$$

The parameters of the above process can be summarized by $(n_\epsilon,\ \lambda_\epsilon,\ \Delta_\epsilon)$.

If we change $(n_\epsilon,\ \lambda_\epsilon,\ \Delta_\epsilon)$ to $(n_\epsilon/\delta,\ \lambda_\epsilon/\delta,\ \Delta_\epsilon\sqrt{\delta})$, it does not change the auto-correlation parameter or instantaneous variance. Thus, as δ tends to 0, auto-correlation parameter and instantaneous variance remains intact but dx converges in distribution to a Gaussian (by Central Limit Theorem).

In the data, we know productivity. It is constantly evolving and state space is very fine. Given $\epsilon(t)$, productivity evolution can be summarized by:

$$d\epsilon = -\frac{\lambda_\epsilon}{n_\epsilon}\epsilon dt + \sqrt{\lambda_\epsilon}\Delta_\epsilon dW \tag{24}$$

Here, dW is a standard normal distribution. Discrete time version of the above Ornstein-Uhlenbeck process:

$$\epsilon(t+1) - \epsilon(t) = -\frac{\lambda_\epsilon}{n_\epsilon}\epsilon(t) + \sqrt{\lambda_\epsilon}\Delta_\epsilon[W(t+1) - W(t)] \tag{25}$$

B: Derivation of the Bellman Equation

Consider an establishment with productivity ϵ and L workers. Suppose we know that the transition happens after a time interval of length T.

$$\begin{aligned} V(L,\epsilon) = max_{\sigma\geq 0} \int_0^T & \{\epsilon F(L) - wL - C(\sigma, L, \epsilon)\}e^{-rh}\, dh \\ & + e^{-rT}[(Prob.\ of\ quitting) * V(L-1,\epsilon) \\ & + (Prob.\ of\ arrival) * V(L+1,\epsilon) \\ & - (Prob.\ of\ exit) * f(L)] \quad \forall (L,\epsilon) \end{aligned} \tag{26}$$

However, T is stochastic. Define X_a as the amount of time that the establishment has to wait until they get a worker. X_a is exponentially distributed with parameter $\lambda(\sigma(L,\epsilon))$. Define X_q to be the amount of time that the establishment has to wait until they lose a worker. X_q is exponentially distributed with parameter δL. Define X_e to be the amount of time until the establishment exits. X_e is exponentially distributed with parameter s. Suppose, X_a, X_q and X_e are independent random variables. Define X as the amount of time that the establishment spends with L workers.

$Pr(X > T) = \Pr(\min\{X_a, X_q, X_e\} > T)$
$= \Pr(X_a > T,\, X_q > T,\, X_e > T)$
$= \Pr(X_a > T)$*$\Pr(X_q > T)$*$\Pr(X_e > T)$ (since these are independent random variables)
$= e^{-(\lambda(\sigma)+\delta L+s)T}$.

Therefore, $f_X(T) = (\lambda(\sigma)+\delta L+s)e^{-(\lambda(\sigma)+\delta L+s)T}$. With $f_X(T)$, we can take the expectation of (10) with respect to T. Taking expectation of (10) with respect to T yields:

$$
\begin{aligned}
V(L,\epsilon) = max_{\sigma\geq 0}\int_0^\infty f_X(T)\int_0^T \{\epsilon F(L) - wL - C(\sigma,L,\epsilon)\}e^{-rh}\,dh\,dT \\
+ \int_0^\infty f_X(T)e^{-rT}[(Prob.\ of\ quitting) * V(L-1,\epsilon) \\
+ (Prob.\ of\ arrival) * V(L+1,\epsilon) \\
- (Prob.\ of\ exit) * f(L)]\,dT \quad \forall(L,\epsilon)
\end{aligned}
\tag{27}
$$

What is Prob. of quitting? Prob of arrival? Prob of exit?

Prob of quitting = $\Pr(X_q = \min\{X_a, X_q, X_e\}) \implies \Pr(X_q < \min\{X_a, X_e\})$

$\implies$

$$\int_0^\infty Pr(X_q < min\{X_a, X_e\}|X_q = t)Pr(X_q = t)\,dt$$

$\implies$

$$\int_0^\infty Pr(t < min\{X_a, X_e\})\,\delta L e^{-\delta L t}\,dt$$

$\implies$

$$\int_0^\infty Pr(t < X_a)Pr(t < X_e)\,\delta L e^{-\delta L t}\,dt$$

$\implies$

$$\int_0^\infty \delta L e^{-(\lambda(\sigma)+\delta L+s)t}\,dt$$

$\implies \frac{\delta L}{\delta L+s+\lambda(\sigma)}$.

Analogously, prob of arrival = $\frac{\lambda(\sigma)}{\delta L+s+\lambda(\sigma)}$.

Prob of exit = $\frac{s}{\delta L+s+\lambda(\sigma)}$.

On substituting the prob of arrival, prob of quitting and prob of exit in (11), we get the Bellman equation.

C: Algorithm to Solve the Bellman Equation and Policy Function

The Bellman equation can be defined on a compact state space L $\in$ [0,$\bar{L}$], where $\bar{L}$ is so large that it never binds (since F'(∞) = 0). Algorithm to solve the Bellman equation and policy function:

1. I first picked an arbitrary vector V(L,ϵ) for all ϵ and for all L $\in$ {0,1,...,$\bar{L}$} .

2. Given the initial guess of V(L,ϵ), I calculated $\sigma(L,\epsilon)$ $\forall$ (L,ϵ).

   - If V(L+1,ϵ) $>$ V(L,ϵ), then $\sigma(L,\epsilon)$ solves:

$$C_\sigma(\sigma(L,\epsilon), L, \epsilon) = \lambda'(\sigma(L,\epsilon))[V(L+1,\epsilon) - V(L,\epsilon)] \tag{28}$$

   - Otherwise, $\sigma(L,\epsilon) = 0$.

3. Using $\sigma(L,\epsilon)$ from step 2, I calculated the updated vector of V(L,ϵ) for all (L,ϵ) from the RHS of Bellman equation (2).

4. I then used the updated vector V(L,ϵ) obtained in step 3 as the new initial guess and repeated steps (2) and (3) until the maximum difference between the elements of the initial vector and updated vector was very small (less than 1e-6).

D: Algorithm to Solve for the Steady State Establishment Size Distribution

1. I first picked an arbitrary vector M(L,ϵ) for all (L,ϵ).

2. Given the initial guess of M(L,ϵ) for all (L,ϵ) from step 1, I calculated the updated M(L,ϵ) for all (L,ϵ) from (4)-(6).

3. I then used the updated vector M(L,ϵ) obtained in step 2 as the new initial guess and repeated the previous step until the maximum difference between the elements of the initial vector and updated vector was very small (less than 1e-6).

E: Simulation of Transition Probabilities from the Steady State Establishment Size Distribution

1. Generate artificial data for 10,000 establishments from the given steady state distribution.

2. Given the artificial data generated in step 1, the goal is to get the vector of establishment employment levels one year later. I proceed as follows:

   - Fix an establishment. Suppose it's initial condition is L as per step 1. Since it is continuous time, at each iteration, one has to determine: (1) time spent by the establishment in each state and (2) given that transition occurs, the state that they will transition to.

   - Starting from time 0, let X_1 denote the amount of time that the establishment has L workers. $\Pr(X_1 > h) = e^{-(\sigma(L,\epsilon)+\delta L+s)h}$. Draw a random number U_1 between 0 and 1. By equating U_1 with $e^{-(\sigma_1(L,\epsilon)+\delta L+s)h}$, one can back out h. This is the number of days in which the establishment has L workers.

   - Therefore, until the h'th day, the establishment stays at L workers. At time h, the establishment will go to state L+1 with probability $\frac{\sigma(L,\epsilon)}{\sigma(L,\epsilon)+\delta L+s}$. The establishment will go to state L-1 with probability $\frac{\delta L}{\sigma(L,\epsilon)+\delta L+s}$. The establishment will go to state 0 with probability $\frac{s}{\sigma(L,\epsilon)+\delta L+s}$. In order to determine which state the establishment will go to at time h, draw another random number U2 from [0,1]. If it is lower than $\frac{\sigma(L,\epsilon)}{\sigma(L,\epsilon)+\delta L+s}$, the establishment will go to state L+1. If it is greater than $\frac{\sigma(L,\epsilon)}{\sigma(L,\epsilon)+\delta L+s}$ but lower than $\frac{\sigma(L,\epsilon)+\delta L}{\sigma(L,\epsilon)+\delta L+s}$, the establishment will go to state L-1. Otherwise, it will go to state 0.

   - Therefore, now h days has passed and the establishment is in state L+1 or L-1 or 0 depending on the realization of U2. By repeating the above steps, we would know the state of the establishment 365 days later.

- Repeat the above steps for all establishments. What we will end up having is a vector of establishment employment levels 1 year later from when we started.

3. From the initial vector in step 1 and the new vector in step 2, one can calculate transition probability around the regulatory threshold as: (No of establishments that employed less than or equal to 100 workers based on the initial vector and above 100 workers one year later/No of establishments that employed less than or equal to 100 workers based on the initial vector).

F: Simulations

Table 3: Parameter values used for model simulations

Parameters	Description	Parameter Values
R	Regulatory threshold	100
r	Rate of time preference	0.05
Avg of log(ϵ)	Average establishment productivity	14.75
λ_ϵ	Annual arrival rate of productivity shocks	0.05
α	Prod function parameter	0.4
w_f	Annual average wage rate per formal worker	66000
w_c	Annual average wage rate per contract worker	38000
η	Convexity in cost of search	$\eta \in \{10, 10.25, 10.5\}$
δ_0	Annual worker attrition rate	$\delta \in \{0.1, 0.15, 0.2\}$
s	Annual establishment exit rate	0.5
κ	Convexity in firing cost function	5
ρ, v	Elasticities of substitution	1.6,3.3
τ,γ	Costs imposed by the regulation	0.2,0.1

Figure 11: Impact of regulatory costs (τ) on search and transitions around the threshold

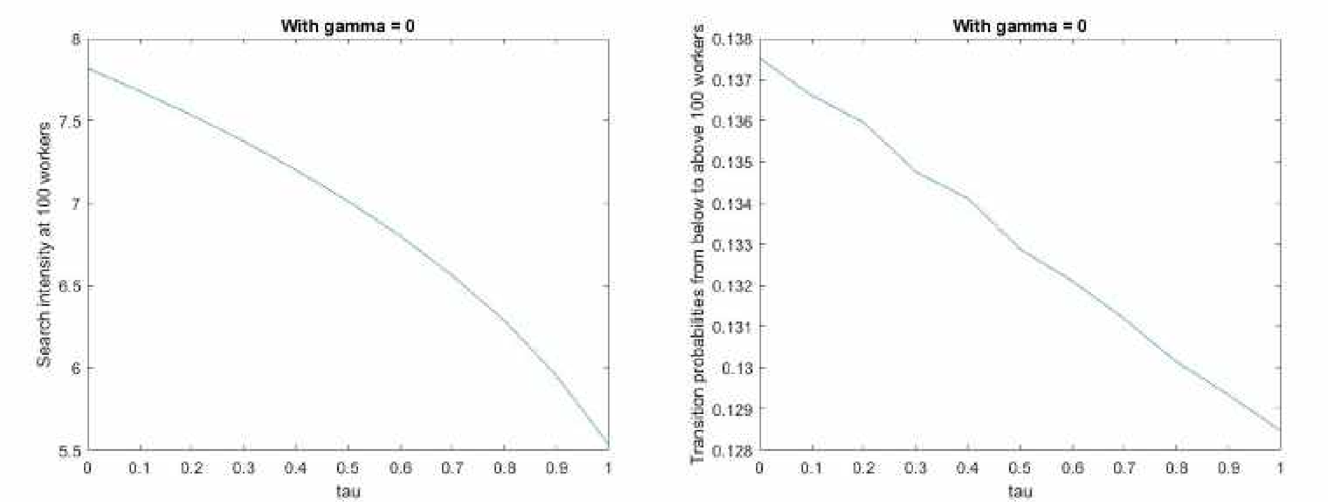


Figure 12: Impact of regulatory costs on substitutions near the 100-worker threshold

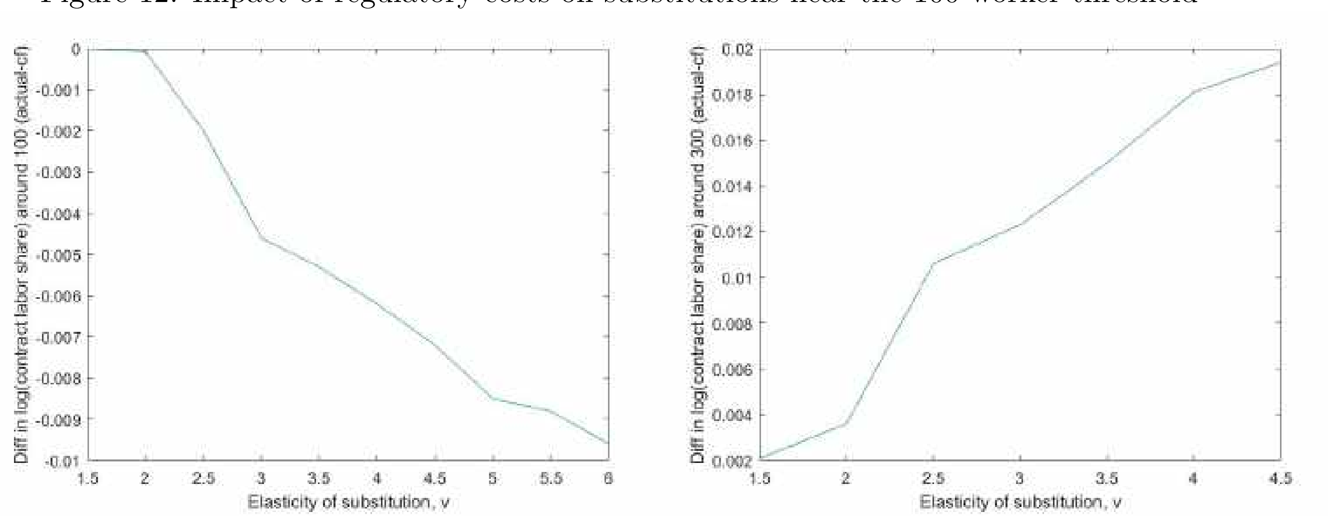
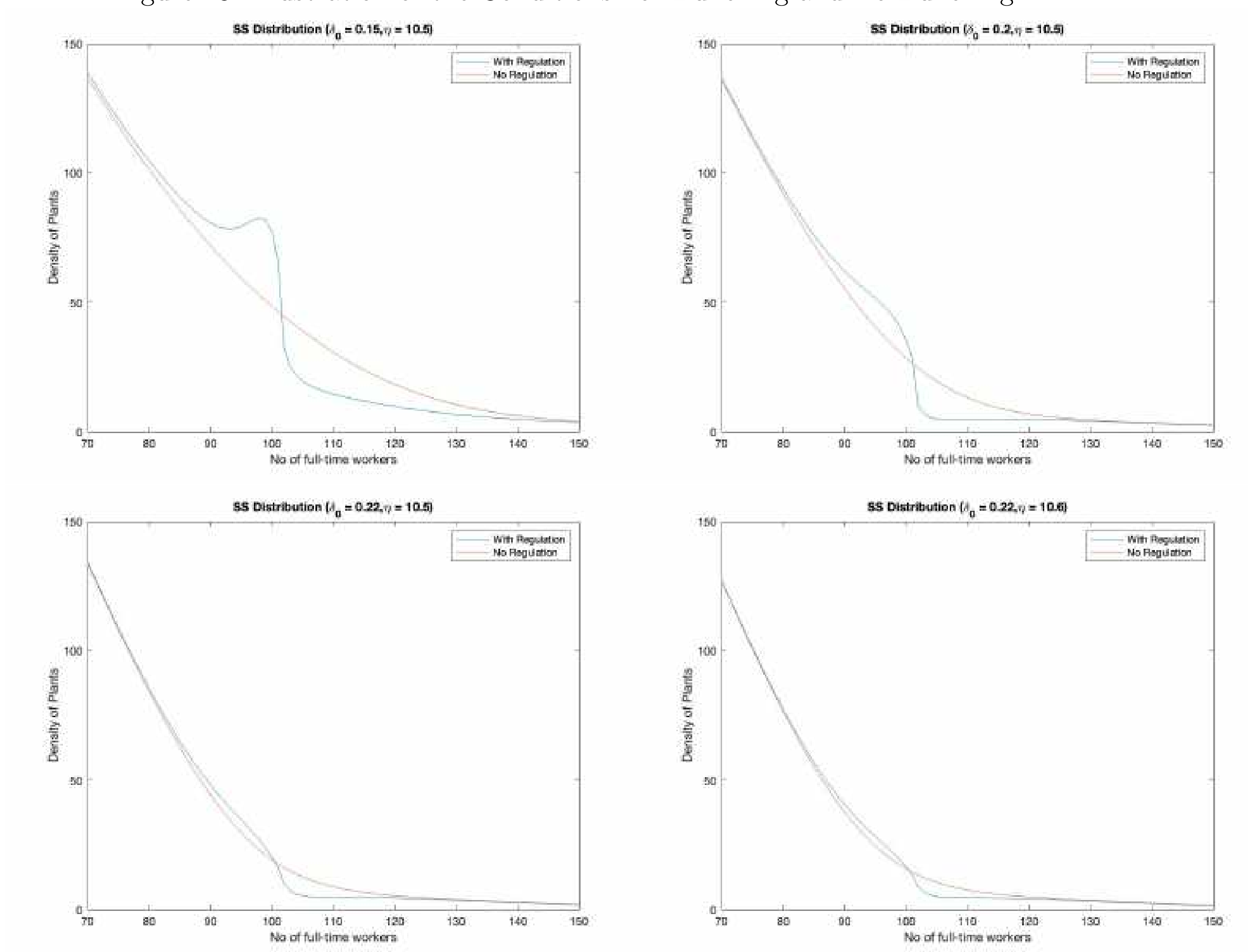


Figure 13: Illustration of the Conditions For Bunching and No Bunching



IV: Quantification of the Model

Table 4: Estimates of v: elasticity of substitution between full-time and contract workers

	(1)	(2)	(3)	(4)	(5)
Estimate of $\frac{v-1}{v}$	0.623^{***}	0.63^{***}	0.631^{***}	0.631^{***}	0.611^{***}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Constant	2.621^{***}	2.622^{***}	2.622^{***}	2.622^{***}	2.727^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Industry FE	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
State FE	No	Yes	Yes	Yes	Yes
Industry FE*Year FE	No	No	Yes	Yes	Yes
State FE*Year FE	No	No	No	Yes	Yes
Plant FE	No	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Estimates of ρ: elasticity of substitution between labor and capital

	(1)	(2)	(3)	(4)	(5)
Estimate of $\frac{\rho-1}{\rho}$	0.698***	0.719***	0.721***	0.721***	0.820***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Constant	8.555***	8.784***	8.805***	8.814***	10.024***
	(0.015)	(0.016)	(0.016)	(0.016)	(0.038)
Industry FE	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
State FE	No	Yes	Yes	Yes	Yes
Industry FE*Year FE	No	No	Yes	Yes	Yes
State FE*Year FE	No	No	No	Yes	Yes
Plant FE	No	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Targeted Moments: Model versus Data

Moments	Period	Model	Data
Panel A: Distribution of Formal Workers			
10^{th} percentile	Pre policy	6	5
20^{th} percentile	Pre policy	9	8
30^{th} percentile	Pre policy	14	13
40^{th} percentile	Pre policy	20	19
50^{th} percentile	Pre policy	31	31
60^{th} percentile	Pre policy	50	51
70^{th} percentile	Pre policy	81	83
80^{th} percentile	Pre policy	290	271
Panel B: Average share of contract workers by plants with			
0-20 formal workers	Pre policy	0.8	0.76
21-100 formal workers	Pre policy	0.64	0.61
101-162 formal workers	Pre policy	0.50	0.49
163-300 formal workers	Pre policy	0.40	0.42
300+ formal workers	Pre policy	0.32	0.32
Panel C: Impact of the policy change			
Change in transition probabilities around 100 workers	Post − Pre	0.013	0.015
Change in addition to capital goods around 100 workers	Post − Pre	-0.007	-0.01
Change in contract labor share around 100 workers	Post − Pre	-0.01	-0.02

Table 7: Non Targeted Moments: Model versus Data

Moments	Period	Model	Data
Share of output contributed by plants with:			
0-20 formal workers	Pre	0.2	0.16
21-100 formal workers	Pre	0.31	0.31
101-162 formal workers	Pre	0.09	0.12
163-300 formal workers	Pre	0.13	0.14
300+ formal workers	Pre	0.20	0.24

References

Daron Acemoglu and William B Hawkins. Search with multi-worker firms. *Theoretical Economics*, 9(3):583–628, 2014.

Ahmad Ahsan and Carmen Pagés. Are all labor regulations equal? evidence from indian manufacturing. *Journal of Comparative Economics*, 37(1):62–75, 2009.

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

Marianne Bertrand, Chang-Tai Hsieh, and Nick Tsivanidis. Contract labor and firm growth in india. Technical report, National Bureau of Economic Research, 2021.

Timothy Besley and Robin Burgess. Can labor regulation hinder economic performance? evidence from india. *The Quarterly journal of economics*, 119(1):91–134, 2004.

Aditya Bhattacharjea. Labour market regulation and industrial performance in india: A critical review of the empirical evidence. *Indian Journal of Labour Economics*, 49(2): 211–232, 2006.

Aditya Bhattacharjea. Labour market flexibility in indian industry: a critical survey of the literature. *A shorter version, incorporating referees' comments, has been accepted for publication in the International Labour Review*, 2020.

A Kerem Coşar, Nezih Guner, and James Tybout. Firm dynamics, job turnover, and wage distributions in an open economy. *American Economic Review*, 106(3):625–63, 2016.

Sanghamitra Das, Mark J Roberts, and James R Tybout. Market entry costs, producer heterogeneity, and export dynamics. *Econometrica*, 75(3):837–873, 2007.

Bibek Debroy and Laveesh Bhandari. India labour report 2008. *Teamlease Services*, 2008.

Rafael Dix-Carneiro, 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.

Jonathan Eaton, David Jinkins, James Tybout, and Daniel Xu. Two-sided search in international markets. In *2016 Annual Meeting of the Society for Economic Dynamics*, 2016.

Jonathan Eaton, David Jinkins, James R Tybout, and Daniel Xu. Two-sided search in international markets. Technical report, National Bureau of Economic Research, 2022.

Manuel García-Santana and Josep Pijoan-Mas. Small scale reservation laws and the misallocation of talent. 2011.

Luis Garicano, Claire Lelarge, and John Van Reenen. Firm size distortions and the productivity distribution: Evidence from france. *American Economic Review*, 106(11):3439–79, 2016.

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

Nezih Guner, Gustavo Ventura, and Yi Xu. Macroeconomic implications of size-dependent policies. *Review of economic Dynamics*, 11(4):721–744, 2008.

Rana Hasan, Devashish Mitra, and Krishnarajapet V Ramaswamy. Trade reforms, labor regulations, and labor-demand elasticities: Empirical evidence from india. *The Review of Economics and Statistics*, 89(3):466–481, 2007.

Hugo Hopenhayn and Richard Rogerson. Job turnover and policy evaluation: A general equilibrium analysis. *Journal of political Economy*, 101(5):915–938, 1993.

Hugo A Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150, 1992.

Chang-Tai Hsieh and Peter J Klenow. Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, 124(4):1403–1448, 2009.

Chang-Tai Hsieh and Peter J Klenow. The life cycle of plants in india and mexico. *The Quarterly Journal of Economics*, 129(3):1035–1084, 2014.

Chang-Tai Hsieh and Benjamin A Olken. The missing" missing middle". *Journal of Economic Perspectives*, 28(3):89–108, 2014.

Leo Kaas and Philipp Kircher. Efficient firm dynamics in a frictional labor market. *American Economic Review*, 105(10):3030–60, 2015.

Per Krusell, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante. Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, 68(5): 1029–1053, 2000.

Marc J Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica*, 71(6):1695–1725, 2003.

Fernando Parro. Capital-skill complementarity and the skill premium in a quantitative model of trade. *American Economic Journal: Macroeconomics*, 5(2):72–117, 2013.

Robert Shimer. The cyclical behavior of equilibrium unemployment and vacancies. *American economic review*, 95(1):25–49, 2005.

Gabriel Ulyssea. Firms, informality, and development: Theory and evidence from brazil. *American Economic Review*, 108(8):2015–47, 2018.