

Indian Labor Regulations and the Cost of Corruption: Evidence from the Firm Size Distribution*

Amrit Amirapu and Michael Gechter[†]

Preliminary and Incomplete

Abstract

This paper investigates the effects of an important but little-researched set of Indian labor and industrial regulations. We use a novel methodology to provide a) the first objective cost estimates of any Indian labor regulations and b) evidence of their impact on misallocation of resources between firms. Our methodology takes advantage of the fact that some regulations only apply to establishments which hire 10 or more employees. Using data from India's 2005 Economic Census, we observe that the distribution of establishments by size closely follows a power law, but with a significant drop in the distribution for establishments with 10 or more workers. Guided by a model based on Garicano, LeLarge and Van Reenen (2013) - but augmented to allow for the possibility of misreporting - we use this drop to estimate the implied costs of the regulation. We find that there is substantial variation in our estimated costs across states, industries and ownership types, and that our costs are more robustly correlated with measures of corruption than with any other factors, suggesting that poor state implementation may be as much or more to blame for the high costs than the regulations themselves. We find further that higher costs are associated with lower rates of future employment growth in registered (but not unregistered) manufacturing, suggesting that these costs may play a role in encouraging informality.

*We wish to thank Dilip Mookherjee, Arvind Subramanian, Hiroaki Kaido, Kevin Lang, Kehinde Ajayi, Sam Bazzi and Marc Rysman for their input, guidance and encouragement. We also thank Laveesh Bhandari, Aditya Bhattacharjea, Sharon Buteau, Areendam Chanda, Urmila Chatterjee, Nancy Chau, Bivas Chaudhuri, Bibek Debroy, Francois Gourio, Jayintha Harinath, Rana Hasan, Ravi Kanbur, Saibal Kar, Dan Keniston, Devashish Mitra, Sandip Mitra, G. Sajeevan, Johannes Schmieder, Pronab Sen, Anushree Sinha, Sandip Sukhtankar, Shyam Sundar, Balasekhar Sudalaimani, John Van Reenen and seminar participants at BU, Cornell University, ISI Kolkata and IGIDR for many useful comments. Representatives in the governments of Bihar, Kerala, Maharashtra, Tamil Nadu and West Bengal provided invaluable assistance in understanding the data collection process for the Economic Census. We gratefully acknowledge financial support from the Weiss Family Program Fund for Research in Development Economics. Stefan Winata and Haiqing Zhao provided expert research assistance, supported by the BU MA-RA mentor program.

[†]Boston University; aamirapu@bu.edu, mgechter@bu.edu

1 Introduction

India’s labor and industrial regulations have been blamed for many of the country’s ills, including low levels of aggregate productivity, slow growth of productivity, and lackluster job creation in the formal (i.e.: registered) sector (e.g. Hsieh and Klenow, 2009; Kochhar et al, 2006; Besley and Burgess, 2004; Hasan and Jandoc, 2012). Of particular note is Hsieh and Klenow’s (2009) landmark study, in which the authors argue that aggregate total factor productivity in India could be 40%-60% higher if not for significant misallocation of resources between firms. They go on to suggest that India’s labor regulations may be to blame for the observed misallocation, although they do not provide significant evidence for this link.¹ In fact, the view that labor regulations are of primary importance is not universally held. Many argue that the laws as written are rarely enforced so that, in practice, firms are effectively unconstrained.² Others argue that the existing evidence on the detrimental impact of labor regulations is flawed (Bhattacharjea, 2006 and 2009). Still others point out that the single regulation which seems to have received the greatest attention from economists and the press (Chapter VB of the Industrial Disputes Act), is not likely to constrain any but the very largest firms (Bardhan, 2014).³

This study uses a novel methodology and a uniquely well-suited dataset to address the above aspects of this debate while avoiding some of the criticisms that have been leveled at previous work. In short, we aim to study the behavior of firms in response to regulatory thresholds in order to determine whether and to what extent firms are in fact constrained by regulations.⁴ We proceed in the following steps. First, we generate the logged establishment-size distribution using data from the Economic Census of India (EC), which, importantly, aims to be a complete enumeration of *all* non-farm business units, regardless of size or status (formal or informal). From the logged distribution, we observe that it closely follows a power law, except for a discontinuous downshift in the distribution for establishments with 10 or more workers. This is precisely the threshold at which a multitude of regulations become legally binding, so we take this observation as evidence that firms with 10 or more workers

¹In fact, they end the paper with the following: “In addition to investigating these issues more fully, future work could try to relate differences in plant productivity to observable policy distortions much more than we have” (Hsieh and Klenow, 2009). We have taken this as a call to action!

²For instance, in a recent paper, Chaurey (2013) provides evidence that firms seem to hire contract workers as a way of avoiding certain regulations.

³Chapter VB of the Industrial Disputes Act (IDA) stipulates that firms in the industrial sector with 100 or more workers must obtain permission from the relevant governmental authority before laying off workers. Bardhan (2014) points out that 92 percent of firms in the garment sector have less than 8 workers.

⁴Note: for expositional purposes we occasionally refer to “firms”, although it would be more correct to refer to “factories” or “establishments”, since all of the data and most of the regulations are at the factory/establishment level rather than the firm level. Regardless, the distinction is almost moot: nearly all Indian firms are single factory/establishment firms.

do seem to be constrained in size by certain regulations - although these are not the same regulations that most others have focused on. We then develop a model of firm size choice under regulatory thresholds which is based on Garicano, LeLarge and Van Reenen (2014) (henceforth GLV), but augmented to explicitly allow for the possibility of misreporting.⁵ In particular, we model the regulations as causing an increase in the unit labor costs of those firms that report having exceeded the 10 worker threshold, and then use the observed distortion to estimate these costs. Under our primary estimation method at the All-India level, we find that firms behave as if operating at or above the 10-worker threshold entailed a 37% increase in their per-worker costs.

Our next step is to document substantial heterogeneity in the size of our estimated costs along several dimensions including state, industry and ownership type. For example, we find that the state with the highest estimated regulatory costs is Bihar and that privately-owned establishments have the highest costs, while government-owned establishments have the lowest. Exploring this variation further, we find that our estimated costs turn out to be correlated with some previous state-level measures of labor regulations (in particular, certain measures from Dougherty, 2009), though not with others (for example, the Besley-Burgess measure from Aghion, et al, 2008). Moreover, we find strong and robust correlations between our estimated costs and two quite distinct measures of corruption⁶, even after controlling for a number of factors including state GDP per capita. As further support for our state-level results, we show that industries with greater “regulatory dependence” have higher estimated costs but only when they are located in more corrupt states. We take these correlations to be suggestive of the fact that the true cost of the regulations may have more to do with bureaucracy and corruption, rather than the content of labor and industrial regulations themselves.

Finally, we turn to a brief discussion of the possible dynamic consequences of the costs we estimate. We show that higher costs are associated with slower growth in employment and productivity in the *registered* manufacturing sector (where the regulations apply), whereas the effect of higher costs on employment and productivity growth in the *unregistered* manufacturing sector (where the regulations do not apply) is more muted - or even in the opposite direction. This suggests that the costs we estimate may play a role in the “informalization” of the Indian economy, by pushing workers from the formal to the informal sector.

⁵Misreporting was a lesser concern in GLV’s original setting, as they had access to administrative data. In contrast, the data in the Economic Census are self-reported, which makes the threat of deliberate misreporting more significant in our case.

⁶These corruption measures include a subjective, perceptions-based measure of corruption from Transparency International and a measure of the percentage of electricity that is lost in transmission and distribution as reported by the Reserve Bank of India (this latter measure has been used as a proxy for government corruption and ineffectiveness in, for example, Kochhar et al, 2006).

This paper is related to at least three important strands of literature. The first, which we have already mentioned, is the literature on misallocation of resources and total factor productivity (TFP), as exemplified by Hsieh and Klenow (2009). Our contribution to this literature is to provide direct evidence that at least some of the misallocation of resources between firms in India *is* tied in some way to regulations.⁷ In particular, we show that size-based regulations (or at least the ways in which they are enforced) lead firms to fall short of their optimal scale, thus distorting the allocation of labor among firms in the economy and, likely, lowering TFP.

Another strand of literature to which we aim to contribute is the literature on corruption in the enforcement of government policies. Most previous studies (eg: Besley and McLaren, 1993; Mookherjee and Png, 1995) have modeled such corruption as *collusion* between inspectors and firms or citizens: corrupt inspectors allow firms to avoid the costs imposed by regulations in exchange for bribes - and hence corruption lowers costs associated with regulations. However, our results suggest that the costs associated with size-based regulations are higher in more corrupt environments, and hence are more in line with an alternative idea that corruption can take the form of *extortion* between inspectors and firms (i.e.: corrupt inspectors take advantage of bureaucratic regulations in order to extract higher rents from firms in the form of harassment bribes).⁸ We present a theoretical model and some interesting anecdotal evidence from “[ipaidabribe.com](#)” to support this interpretation.⁹ We view the support we provide for this alternative conception of corruption as another contribution of the paper and we hope that it will shift the present debate away from the arguments over the pro or anti-labor stance of labor regulations and towards arguments about clarity, bureaucracy and enforcement of regulations.

Lastly, this paper is also clearly related to the large literature that more generally investigates the impact of Indian regulations on economic outcomes. The literature dates back to at least Fallon and Lucas (1993), but seems to have really taken off after the work of Besley and Burgess (2004). In that paper, the authors first interpret state-level amendments to the Industrial Disputes Act (IDA) as either “pro-worker” or “pro-employer” and then aim to show that Indian states that amended the IDA in a “pro-worker” direction experienced slower growth in output, employment, investment and productivity in registered manufacturing. The paper, though extremely influential, has been criticized by Bhattacharjea (2006 and 2009) on a number of grounds. One of the major criticisms is that Besley and Burgess’s interpretations of amendments as “pro” or “anti-worker” are subjective and debatable (ie:

⁷In future work we hope to determine what portion of the TFP loss from misallocation estimated by Hsieh and Klenow (2009) can be attributed to the regulations we study.

⁸This finding echos Novosad and Asher (2014), in which it is argued that regulations can provide a means through which politicians can impose costs on businesses.

⁹We thank Andrew Foster for the suggestion to search for anecdotal evidence from this website.

different people might read and code them in a different way). This criticism affects most of the subsequent academic work on this topic, since most papers use the Besley-Burgess codings, but it is a criticism we are able to sidestep with our methodology. Since our analysis is based on size-thresholds stated explicitly in the laws themselves, it has the advantage of objectivity.

The second way in which we distinguish ourselves from the other papers in this literature is in the set of regulations that we study. Nearly all prior academic papers on the subject focus on analyzing components of the Industrial Disputes Act (IDA) exclusively - particularly Chapter VB of the IDA, which only applies to manufacturing firms with 100 or more workers. In contrast, the regulations we study affect all firms with 10 or more workers and are thus relevant for a much larger share of firms.¹⁰ This is particularly true given the remarkable extent to which the Indian firm size distribution is right-skewed (see Figure 1). The only other papers of which we are aware that study regulations that kick in at the 10-worker threshold are Dougherty (2009), Dougherty et al (2011) and Chatterjee and Kanbur (2013). The latter investigates the Factories Act, which applies to all manufacturing firms that use power and have 10 or more workers (or don't use power and have 20 or more workers), but their focus is to document non-compliance under the act, which we see as complementary to our approach of estimating the costs of the regulations.¹¹ The papers by Dougherty and co-authors employ state-level indices of labor reforms that differ from the Besley-Burgess codes in that they include consideration of non-IDA regulations such as the Factories Act, but they are constructed from surveys of industry experts and, as such, are by and large subject to similar concerns regarding subjectivity.

Another way in which we distinguish ourselves from the previous literature on Indian regulations is that we explore the effect of regulations in all non-farm segments of Indian economy - not just in registered manufacturing, on which nearly all previous academic studies have focused. A final contribution of the paper is to provide suggestive evidence that improper government enforcement of regulations may play a role in shifting employment from the registered to the unregistered sector.

In the next section (Section 2), we provide an overview of the relevant institutional details regarding Indian labor and industrial regulations. Section 3 introduces the data and covers some basics about the size distribution of enterprises in India. In Section 4 we go over the theoretical model and our corresponding empirical strategy. Section 5 provides the main

¹⁰We have also tried analyzing Chapter VB of the IDA using the same methodology we employ for the regulations with the 10 worker threshold, but find no effects. I.e.: there does not seem to be a downshift in the distribution for establishments with more than 100 workers. We also fail to observe "bunching" of firms at sizes just below 100, although the presence of rounding may make such bunching impossible to discern even if it exists.

¹¹Our estimated costs are robust to the possibility of noncompliance.

results. In Section 6, we interpret the findings, explore the multiple dimensions of variation in our results, and investigate the connection between our estimated costs and corruption. Section 7 concludes.

2 Institutional Background: Size-Based Regulations in India

In this paper we attempt to investigate the effect of employment size-based regulations in India. These are industrial and labor regulations that only apply to establishments that exceed a certain size - which can be measured in terms of a firm's revenue, the amount of fixed capital invested, or the number of workers employed. One of the most significant such thresholds occurs when establishments employ 10 or more workers, after which they must register with the government and meet various workplace safety requirements (under the Factories Act, for example¹²), pay social security taxes (under Employees' State Insurance Act), distribute gratuities (under the Payment of Gratuity Act) and bear a greater administrative burden (under, e.g., the Labor Laws Act).

Not only are the laws numerous, it has been argued that certain components of the laws are antiquated and/or arbitrary. For example we read in the "India Labour Report" that "Rules under the Factories Act, framed in 1948¹³, provide for white washing of factories. Distemper won't do. Earthen pots filled with water are required. Water coolers won't suffice. Red-painted buckets filled with sand are required. Fire extinguishers won't do... And so on" (TeamLease Services, 2006). Firm owners who choose not to comply with such regulations may face costs if discovered and convicted.¹⁴

In addition to - or in lieu of - the explicit costs of complying with the regulations, establishments with 10 or more workers may be subject to implicit costs associated with increased interaction with labor inspectors, et al, who may have the power to extract bribes and tighten (or ease) the administrative burden firms face. Indeed, inspectors in India have a large amount of discretion regarding the enforcement of administrative law. For example, in some cases, the definition of what constitutes a "day" is at the discretion of the inspector, and it is a commonly held view that "[w]hile grave violations are ignored, minor errors become a scope for harassment" (TeamLease Services, 2006).

This kind of behaviour has been referred to as "harassment bribery" (Basu, 2011). Anec-

¹²Technically the Factories Act applies for 10-plus worker establishments only if they use power. For establishments that do not use power, the Factories Act does not apply until they employ 20 workers.

¹³The Factories Act itself dates to 1948, but the origins of the law go back another 100 years at least, to Britain's first Factory Acts.

¹⁴These costs may include fines and/or prison sentences.

total evidence of inspectors using the complexity and sheer amount of paperwork as a way to extract bribes is easy to come by. For example, we have included a selection of citizen reports from “ipaidabribe.com” in Appendix 2, which demonstrate just this kind of behaviour.¹⁵ Interestingly, some of the reports suggest that the size of the bribe paid is a direct linear function of the number of employees - which will become relevant to our estimation procedure later.

As we alluded to earlier, the 10 worker threshold is not the only one relevant; there are other cutoffs at which different regulations become binding. For example, the one that seems to have received the most attention, both from academics and the press, is that of 100 workers, at which enterprises in most states become subject to Chapter VB of the Industrial Disputes Act, under which they must be granted government permission to lay off workers. There are other cutoffs still,¹⁶ but in this paper we will focus on estimating the costs and effects associated with the regulations that come into force at the 10 worker cutoff.

3 Data and the Size Distribution in India

3.1 Data

The data we rely on to investigate the 10-worker threshold comes from the Economic Census (EC) of India. The EC is meant to be a complete enumeration of all (formal *and* informal) non-farm business establishments¹⁷ in India at a given time, *regardless of their size*. It is this last clause that makes the EC different from every other data source available and precisely suited to our needs. Although we may occasionally refer to “firms”, the unit of observation in the data is actually a factory or an establishment, rather than a firm (i.e.: multiple establishments may belong to the same firm). Although the 2005 dataset contains a large number of observations (more than 40 million), there is not very detailed information collected on each observation. For each establishment in the data, there is information only on the total number of workers usually working, the number of non-hired workers (such as family members working alongside the owner), the registration status, the 4-digit NIC industry code, the type of ownership (private, government, etc) and the source of funds for the establishment. There is no information on capital, output or profits, and the data is cross-sectional.

¹⁵We thank Andrew Foster for this suggestion.

¹⁶For example, firms with 20 or more workers must abide by the Provident Funds Act. Firms with 50 or more workers must comply with Chapter VA of the Industrial Disputes Act, which requires them to provide compensation and notice to employees prior to lay-offs.

¹⁷Which the EC calls “entrepreneurial units” and defines as any unit “engaged in the production or distribution of goods or services other than for the sole purpose of own consumption.”

The EC has rarely been used in academic papers - possibly because it is cross-sectional and only contains information on the handful of variables just enumerated, so that better data sources exist for most purposes. The EC is ideal for our purpose, however, since it includes information on employment size and covers the entire universe of establishments. Other more commonly used datasets, such as the CMIE’s Prowess Database, the Annual Survey of Industries (ASI) or the National Sample Survey’s (NSS) Unorganized Manufacturing Surveys cover only certain parts of the distribution and thus cannot be used for our purpose. The ASI, for example, only covers factories in the manufacturing sector that have registered with the government under the Factories Act. However, registration under this Act is only required for establishments with 10 or more workers if the unit uses power (20 or more workers if the factory uses no power). Therefore, the selection of the ASI varies discontinuously at precisely the point of interest. Similar limitations on coverage make the other datasets - other than the EC - unsuitable.

Aside from the Economic Census, we also supplement our analysis with data from a variety of other sources. From the ASI we get employment and productivity in the registered manufacturing sector while we generate those variables for the unregistered sector with data from the Ministry of Statistics and Programme Implementation (MOSPI) and the Reserve Bank of India (RBI). We get data on state and industry level corruption from a) Transparency International’s “India Corruption Study 2005”, b) the RBI, and c) the World Bank Enterprise Survey for India (2005). Data on State-level regulatory enforcement come from the Indian Labour Year Book¹⁸. Other measures of state-level regulations come from Aghion, et al (2008) and Dougherty (2009), while industry-level measures of exposure to trade liberalization come from Ahsan and Mitra (2014).

3.2 The Size Distribution of Establishments in India

Figure 1 above shows the distribution of establishments by the number of total workers (hired plus non-hired workers) for establishments with up to 200 total workers in 2005. One striking feature of figure 1 is the fact that the densities for establishments with 10 or more workers are almost imperceptible.¹⁹ Figure 2 shows the drop in density for establishments with 10 or more workers in detail and figure 3 shows the full distribution of establishment size frequencies according to a log scale. Each point represents one bar in the earlier histograms.

Two things are most striking about figure 3. First, the natural log of the density is a linear

¹⁸We would like to thank Anushree Sinha and Avantika Prabhakar for their considerable and generous help in obtaining these data.

¹⁹The densities for establishments with more than 200 workers are also imperceptible. We have omitted them only for clarity in the figure.

Figure 1: Distribution of establishment size for establishments with 1-200 total workers, 2005

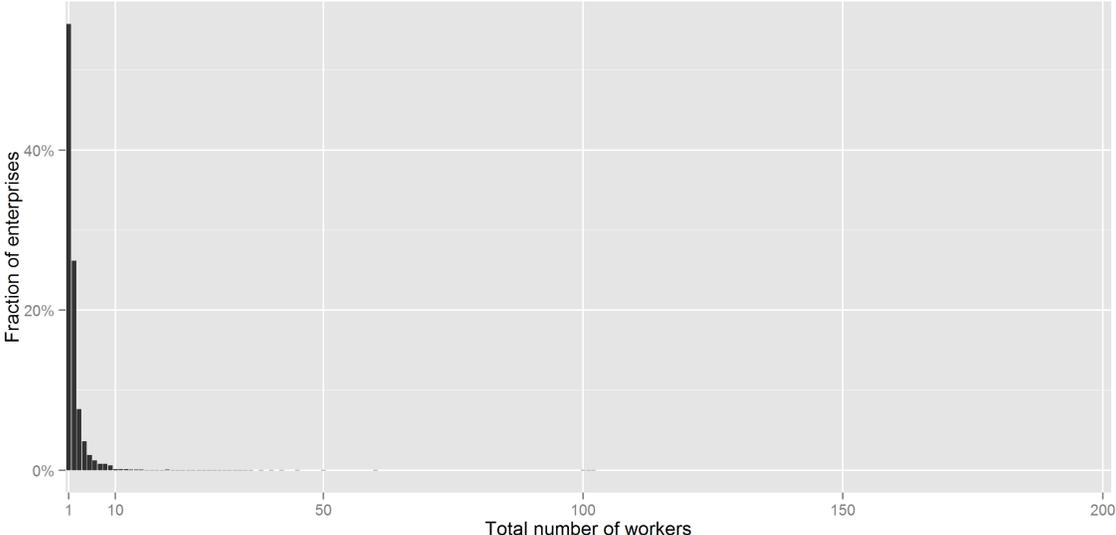


Figure 2: Distribution of establishment size for establishments with 5-25 total workers, 2005

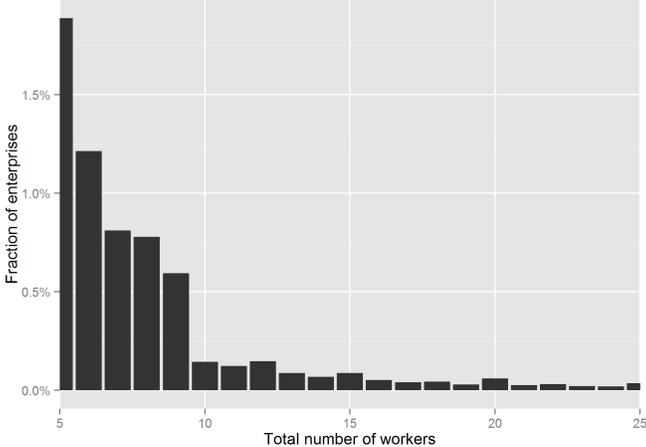
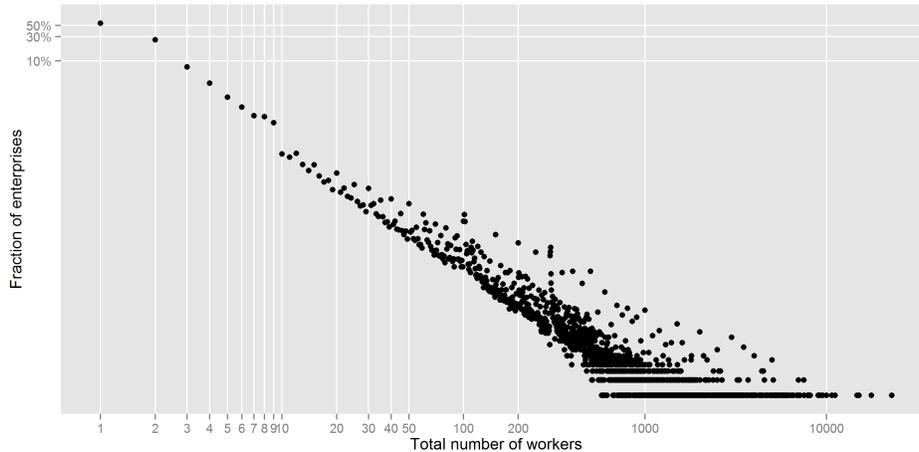


Figure 3: Distribution of establishment size, 2005, log scale



function of the natural log of the number of total workers. This implies that the unlogged distribution follows an inverse power law in the number of total workers. This pattern will be important for the analysis that follows but it is not very surprising in and of itself: power law distributions in firm sizes have been documented in many countries (e.g. Axtell, 2001 and Hernandez-Perez et al, 2006). The second and more unique feature of the distribution is that there is a level shift downward in the log frequency for establishment sizes greater than or equal to 10. Figure 4 shows this effect for establishments with fewer than 100 workers by running an OLS regression of the log density against log firm size and allowing the intercept to vary for firms with 10 or more workers. To the best of our knowledge, ours is the first paper to document this phenomenon in India.

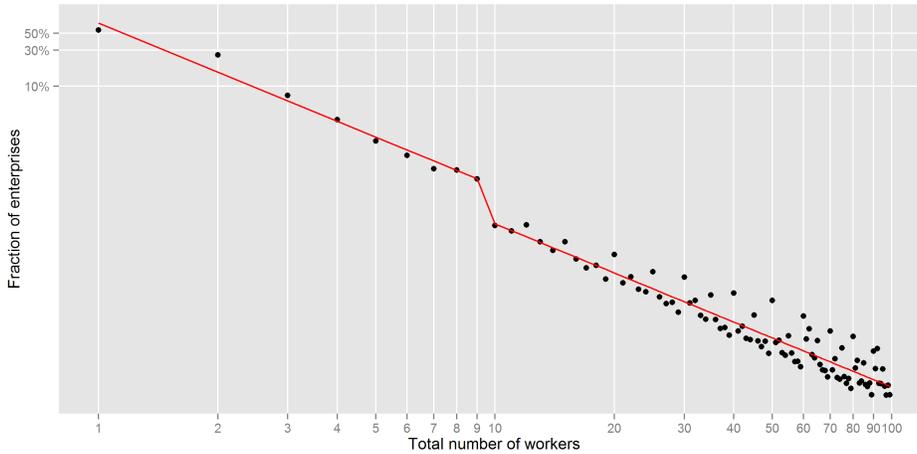
Also of note from the figures above is that there appears to be a significant amount of non-classical measurement error, seemingly due to rounding of establishment sizes to multiples of 5 and 10. The existence of rounding is not surprising given that the data are self-reported and that respondents are asked to give the “number of persons usually working [over the last year]”. Partially to alleviate concerns that the non-classical measurement error due to rounding might bias our results (and partially for other reasons to be made explicit shortly), we will employ an estimation procedure which first smooths the data non-parametrically.

4 Model and Empirical Strategy

4.1 Basic Model

To interpret the downward shift from Figure 4 in economic terms, we turn to the model in GLV. In their framework, size-based regulations are assumed to increase the unit labor costs

Figure 4: Downward shift at the 10-worker threshold in the distribution of establishment size, 2005, log scale (omitting establishments with more than 100 workers)



of firms that exceed the size threshold, which results in a downshift in part of the theoretical firm size distribution. From the magnitude of the downshift they observe in the empirical distribution they attempt to estimate the additional labor costs imposed by the regulations.

GLV begin with a distribution of managerial ability ($\alpha \sim \phi(\alpha)$) as the primitive object, following Lucas (1978). As is common in the literature (e.g. Eaton, 2011), they assume that the distribution of managerial ability follows a power law (e.g. $\phi(\alpha) = c_\alpha \alpha^{-\beta\alpha}$). It is this that will generate a power law in the theoretical firm size distribution. A firm with productivity or managerial ability α faces the following profit-maximization problem:

$$\pi(\alpha) = \max_n \alpha f(n) - w\bar{\tau}n$$

where n is the number of workers a firm employs, $f(n)$ is a production function (with $f'(n) > 0$ and $f''(n) < 0$), w is a constant wage paid to all workers, and $\bar{\tau}$ is a proportional tax on labor that takes the value 1 if $n \leq N$ and τ if $n > N$, where $\tau > 1$.

From the first order condition on this maximization problem, $\alpha = \frac{w\bar{\tau}}{f'(n)}$, one can see that higher productivity establishments/managers will employ more workers, and that firms which cross the threshold (N) and must therefore pay higher labor costs will hire fewer workers than they would otherwise. This latter feature is built to match the observed “downshift” in the actual firm size distribution to the right of the regulatory threshold.

One can informally characterize the solution as follows: one set of managers with particularly low productivity (below some threshold α_1) will be effectively unconstrained. These managers would have chosen to hire fewer than 10 workers whether or not the regulation was present. Another set of managers with slightly higher productivity (between some thresholds α_1 and α_2) would, in the absence of the regulation, have chosen to hire 10 or more

workers - but who, in the presence of the regulation, obtain higher profits by hiring only 9 workers to avoid the discontinuous increase in costs implied by crossing the threshold. These managers should be “bunched up” at 9. The last set of managers are those with high enough productivity ($\alpha > \alpha_2$) that it is not worth it to avoid the regulation and so they choose to exceed the threshold and pay the tax. However, these managers face higher marginal costs than they would in the absence of the regulation and therefore employ fewer workers by a constant proportion (resulting in a “downshift” in the logged firm size distribution).

An exact expression for the distribution of firm size, $\chi(n)$, can be recovered as a transformation of the distribution of managerial ability, $\phi(\alpha)$, since the first-order conditions on the firms’ maximization problems imply a monotonic relationship between α and n . The key result is that a function of the tax enters multiplicatively in the expression for the density of firms size n (for all $n > 9$). Therefore, the function of the tax enters additively in the *log density* for all firms large enough to be subject to the tax.

Formally, the density of firms with n total workers, $\chi(n)$ is given by:

$$\chi(n) = \begin{cases} \left(\frac{1-\theta}{\theta}\right)^{1-\beta} (\beta - 1)n^{-\beta} & \text{if } n \in [n_{\min}, N) \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta} (N^{1-\beta} - \tau^{-\frac{\beta-1}{1-\theta}} n_u^{1-\beta}) & \text{if } n = N \\ 0 & \text{if } n \in (N, n_u) \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta} (\beta - 1)\tau^{-\frac{\beta-1}{1-\theta}} n^{-\beta} & \text{if } n \geq n_u \end{cases}$$

where θ measures the degree of diminishing returns to scale, capturing both features of the production function and market power, β represents the negative slope of the power law and τ is the implicit per worker tax. Taking logs and combining the first and last cases²⁰ leads to:

$$\log(\chi(n)) = \log \left[\left(\frac{1-\theta}{\theta}\right)^{1-\beta} (\beta - 1) \right] - \beta \log(n) + \log(\tau^{-\frac{\beta-1}{1-\theta}}) 1\{n > 9\}$$

This leads to an estimating equation:

$$\log(\chi(n)) = \alpha - \beta \log(n) + \delta 1\{n > 9\} \tag{1}$$

We can identify τ according to:

$$\tau = \exp(\delta)^{-\frac{1-\theta}{\beta-1}}$$

τ is thus a function of θ , β and δ . We get estimates for α , β and δ from equation 1. Knowing α and β pins down θ , which allows us to identify τ .

²⁰In other words, we ignore the bunching at N and the valley directly after, since these are features that are not easily observable in the data. Instead we focus on the ranges $n \in [n_{\min}, N)$ and $n > n_u$.

4.2 Concerns Regarding Misreporting

Before proceeding further, we must consider how our results might be affected by the possibility of misreporting. This is important because one of the underlying assumptions of the analysis above is that the size distribution of firms as observed in the Economic Census is accurate. However, since the data are self-reported, it is possible that plant managers may misreport information to Economic Census enumerators. Specifically, if the managers are aware of the increased regulatory burden that is associated with employing 10 or more workers, and if they believe that the EC enumerators will relay information to government regulatory bodies, they may wish to hide the fact that their actual employment exceeds the threshold. To see how this type of behavior might affect our results, we model it explicitly in the following subsection.

A further reason to be concerned about the possibility of misreporting is due to the fact that Economic Census enumerators were required to fill out an extra form containing the address of any establishment that reported 10 or more workers. It is conceivable that enumerators might have found it preferable to under-report the number of workers for establishments with 10 or more workers in order to avoid the extra burden of filling in the “Address Slip”. Although we do not model this type of problem explicitly in what follows, the implications are nearly identical to those of the model we do explicitly analyze.²¹

4.3 A Theoretical Model of Misreporting

Our model of misreporting starts with the theoretical model from Section 4.1, and amends it to allow firms to choose not only their true employment (n), but also their *reported* employment (l). Then, a firm with productivity α faces the following profit-maximization problem:

$$\pi(\alpha) = \max_{n,l} \alpha f(n) - wn - \tau l * 1(l > 9) - F(n, l) * p(n, l)$$

where α , $f(n)$, w and τ are all defined as they were previously. The problem is identical except that now firms pay the extra marginal cost, τ , only on their *reported* employment, and not on their true employment. Furthermore, they only pay this cost if their reported employment exceeds the threshold.²² There is now an incentive for firms to misreport their employment in a downward direction (i.e.: to set $l < n$). Counteracting this incentive is

²¹The only difference is that higher fixed costs would replace higher marginal costs. It is, moreover, easy to show that if our estimation strategy is robust to the model of misreporting we do analyze, it is also robust to this second type of misreporting as well.

²²In point of fact it is most likely that firms’ answers to Economic Census enumerators have no impact on their regulatory burden, but it is possible that firms believe otherwise, and that is what is relevant.

that misreporting firms may be caught by the authorities with probability $p(n, l)$, and made subject to a fine, $F(n, l)$. As written above, both the probability of being caught and the magnitude of the fine may in general depend on n and l in an arbitrary way. However, if one is willing to make the assumption that the expected cost of misreporting ($F * p$) is an increasing and convex function of the degree of misreporting, $n - l$, it will be possible to use an estimation technique that will be only minimally biased by the presence of misreporting. Fortunately, based on our understanding of the context in which firms make these decisions,²³ we believe that this is the most reasonable assumption on the functional form of the expected cost that one could make.

One plausible way to obtain convex misreporting costs is to suppose that firms are caught with a probability that is linearly increasing in the degree of their misreporting (i.e.: $n - l$) and subject to a fine if caught which is also a linear function of their misreporting. Another possibility is that the probability of being caught is itself an increasing and convex function of the degree of misreporting and the fine if caught is fixed. In what follows we will assume the latter for clarity of exposition, but the analysis is identical for any assumption that yields convex costs of misreporting.

Specifically, suppose that misreporting firms are caught with probability $p(n, l) = \frac{(n-l)^2}{100}$, and pay a fixed fine, F , if caught. Then their profit maximization problem is:

$$\pi(\alpha) = \max_{n,l} \alpha f(n) - wn - \tau l * 1(l > 9) - F * \frac{(n-l)^2}{100}$$

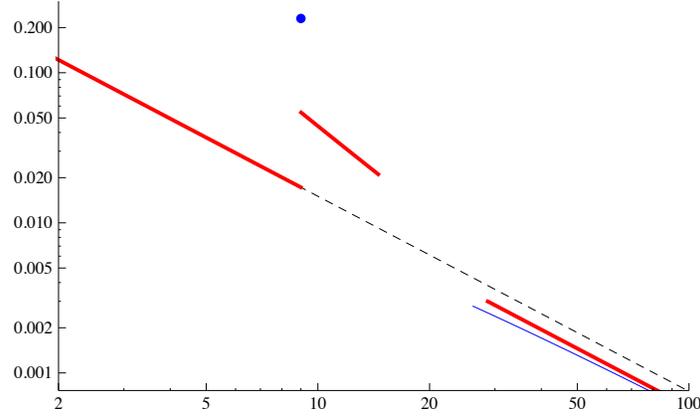
The solution to this problem can be informally characterized as follows. The lowest productivity firms (those with α below some threshold, α_1) will be unconstrained, choosing $n \leq 9$ and reporting truthfully ($l = n$). Higher productivity firms, with $\alpha \in [\alpha_1, \alpha_2]$, will choose $n > 9$, exceeding the regulatory threshold, but will find it profitable to misreport their employment, setting $l = 9$. These firms will only *appear* to be “bunched” up at 9, but will in fact have higher employment. The last category of firms are those with $\alpha > \alpha_2$, which are productive enough to warrant hiring work forces so large that they cannot avoid detection with reasonable probability and must report $l > 9$. Even these firms, however, with both $n > 9$ and $l > 9$ do not find it profit-maximizing to report truthfully. They can save on their unit labor costs by shading their reported employment, and will choose $l = n - \frac{50}{F}\tau$. Note that the degree of misreporting is by a constant amount, rather than a constant proportion.²⁴

More formally, the log of the density of firms with true employment n , $\log\chi(n)$, is given

²³This understanding is informed by informal interviews with small businesses in Chennai and our reading of the secondary literature.

²⁴This outcome is a result of the convex cost assumption.

Figure 5: Theoretical Model of Misreporting, log scale (Red = true distribution; Blue = reported distribution; Dashed Line = counterfactual distribution)



by:

$$\log\chi(n) = \begin{cases} \log A - \beta \log(n) & \text{if } n \in [n_{\min}, 9) \\ \log[\xi(n)] & \text{if } n \in [9, n_m(\alpha_2)] \\ 0 & \text{if } n \in (n_m(\alpha_2), n_t(\alpha_2)) \\ \log A'(\tau) - \beta \log(n) & \text{if } n \geq n_t(\alpha_2) \end{cases}$$

while the log of the density of firms with reported employment l , $\log\psi(l)$ is given by:

$$\log\psi(l) = \begin{cases} \log A - \beta \log(l) & \text{if } l \in [l_{\min}, 9) \\ \log(\delta_l) & \text{if } l = 9 \\ 0 & \text{if } n \in (9, l_t(\alpha_2)) \\ \log A'(\tau) - \beta \log(l + \frac{50}{F}\tau) & \text{if } l \geq l_t(\alpha_2) \end{cases}$$

where terms have been simplified and collected.²⁵ Both of these densities are graphically represented in Figure 5 under specific values of the parameters (the true distribution, $\chi(n)$, is in red, and the reported distribution, $\psi(l)$, is in blue). The key things to note are the following. First, for the range $l \leq 9$, the true distribution coincides with the reported/observed distribution. Second, there appears to be bunching at 9 in the reported distribution, but these firms in fact have greater than 9 workers. Third, compared to the distribution for $n < 9$, the true distribution *and the reported distribution* for $n \gg 10$ are downshifted ($A'(\tau) < A$), just as was the case in the model without misreporting.²⁶ Fourth - and most significantly - the reported distribution converges to the true distribution for large l : $\lim_{l \rightarrow \infty} \beta \log(l + \frac{50}{F}\tau) \rightarrow \beta \log(l)$. This is due to the fact that the misreporting is by a constant amount (as noted earlier), rather than by a constant proportion.

²⁵Derivation to be added in an appendix.

²⁶As before, the downshift is a function of τ .

To conclude this subsection on the theoretical implications of allowing for misreporting, we make two observations. First, misreporting may lead us to observe “bunching” in the firm size distribution when in reality there may be none. This is irrelevant for us since our estimation strategy does not rely on the bunching in any way. Second, misreporting may lead the reported/observed distribution to understate the true distribution *close to the cutoff*. However, because the reported distribution converges to the true distribution for large l/n , misreporting is not able to induce a downshift in the reported distribution that differs from the downshift in the true distribution at large values of l . Therefore, if we use an estimation strategy that focuses mostly on values far from the cutoff, our estimate of the downshift using the observed distribution is likely to reflect the real downshift and thus we are likely to avoid this source of bias. We develop such an estimation strategy in the following subsection.

Before proceeding, however, we should note again that the above analysis assumes that the expected costs of misreporting are strictly convex. There exist non-convex functional forms of the cost function which may lead one to observe a downshift in the reported distribution that is greater than the one in the true distribution, thus biasing any estimates of τ upwards.

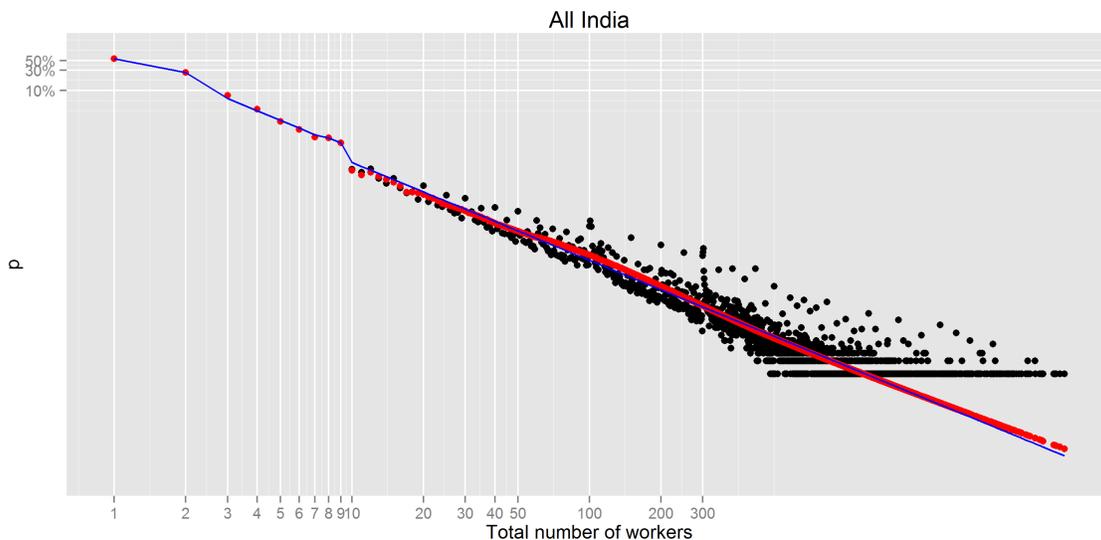
4.4 An Empirical Strategy Robust to the Possibility of Misreporting

Since convex misreporting costs imply that misreporting will only distort the distribution of reported establishment size versus the true distribution of establishment size close to the cutoff, we estimate the model on the *full* distribution of establishment size. Since estimating equation 1 treats each establishment size as one observation, using the full distribution of establishment size will mean that the model is primarily estimated using data far from the 10-worker cutoff.²⁷ However, estimating equation 1 on the full distribution of establishment size introduces two complications. First, we cannot perform the estimation on the empirical PMF for large firm sizes, since the empirical probability mass is truncated at the reciprocal of the number of observations (see figure 6 below), while the underlying density continues to diminish in establishment size. Second, respondents appear to round their reported number of workers to the nearest multiple of 5 (see figure 2), a phenomenon that is more pronounced for larger establishments and that could bias our results.

To address these two problems, we first estimate the density associated with each number of workers $\chi(n)$ non-parametrically using the method of Markovitch and Krieger (2000),

²⁷The largest establishment in the 2005 EC has 22,901 workers.

Figure 6: Downward shift at the 10-worker threshold in the distribution of establishment size estimated on nonparametric density estimates, 2005, log scale (including all establishments). Black = actual data; Red = smoothed data; Blue = linear fit.



which addresses the econometric issues arising in nonparametric density estimation of heavy-tailed data. We then use the nonparametric density estimates as a basis for fitting the model in equation 1, augmented by dummy variables for having 1, 2, 8 and 9 workers.²⁸ Figure 6 depicts the strategy. The black dots show the raw data. The red dots represent the result of the first step: nonparametric density estimates associated with each establishment size. The blue line shows the fit of the model in equation 1, augmented by the dummy variables, to the nonparametric density estimates.

Figure 6 above provides some evidence for the model described in section 4.3. The observed establishment size distribution appears to converge back to a power law with the same slope as for establishments with fewer than 10 workers, but deviates slightly from that slope at sizes just above the 10-worker cutoff. In the next section we report the results of the estimation.

²⁸The rationale for flexibly modeling the density at 1 and 2 workers is that own account enterprises and 2-worker enterprises are likely to be household enterprises and may therefore differ fundamentally in character from their larger counterparts. The rationale for flexibly modeling the density at 8 and 9 workers is that the theory above predicts that the reported density just below the cutoff will be biased upwards by any misreporting effects. Similarly, the theory also predicts that values above - but close to - the threshold may also be biased (downwards). It is also possible to flexibly control for such values as well, but doing so has only a very small effect on the estimates: as explained above, the estimates are driven mostly by observations relatively far from the threshold.

5 Preliminary Results

In this section we report the results of the estimation procedure described above applied to the 2005 Economic Census of India. Standard errors obtained by replicating the procedure on 200 draws from the empirical distribution function of the establishment size data are given in parentheses.

In the tables below, we first report estimates for $\tau - 1$ (as well as all other parameters estimated) at the All-India level. We then provide estimates for $\tau - 1$ for a selection of States, Industries and Ownership Types. For States, estimates are provided using the whole set of enterprises in the state, and also using only the subset of the state's establishments that are in the manufacturing sector (this latter estimation is denoted by τ_{manuf}). Estimates for *all* States and Industries will be added to the Appendix. The All-India estimate on $\tau - 1$ is .37 and is statistically significant. This means that, on average, establishments in India that hire more than 9 workers act as though they must pay additional labor costs of 37% of the wage per additional worker. In the tables and figures it can be seen that there is substantial variation in the magnitude of our estimates of the per-worker tax by State, Industry and Ownership Type. For example, the point estimate on $\tau - 1$ for the State of Kerala is .13, while the estimate for Bihar, on the other hand, is .97, implying that establishments in Bihar act as though they must pay a tax of 97% of the wage for each additional worker they hire past 9 workers.

We also observe substantial differences in the size of τ by industry: it appears that the effective tax is highest for establishments in construction and retail. As one might expect, the tax is nonexistent for establishments in the public administration sector (in fact it is negative, but this seems to result from the fact that the assumed power law does not fit the distribution of establishments in this sector well). Similarly, when looking at the differences by ownership type, we find that the estimates for τ are highest for private firms (especially unincorporated proprietorships), and nonexistent (or negative) for government-owned firms, where presumably the regulatory burden is less than in the private sector.

All-India Estimates Using the Full Distribution of Establishment Size

α	β	δ	θ	$\tau - 1$
-0.233	2.285	-0.679	0.407	0.368
(0.004)	(0.002)	(0.009)	(0.0005)	(0.003)

Estimates of τ by State Using the Full Distribution of Establishment Size

State	$\tau - 1$	$\tau_{manuf} - 1$
Bihar	.966 (0.077)	2.815 (0.425)
Gujarat	0.257 (0.013)	0.026 (0.010)
Kerala	0.127 (0.014)	0.704 (0.065)
Tamil Nadu	0.390 (0.012)	0.347 (0.019)
Uttar Pradesh	0.652 (0.022)	0.686 (0.034)
West Bengal	.198 (.010)	.476 (.018)

Estimates of τ by Industry Using the Full Distribution of Establishment Size (All India)

Industry	$\tau - 1$
Construction	.602 (0.061)
Public admin., social security	-.295 (0.002)
Hotels, restaurants	.520 (0.029)
Manufacturing	.324 (0.005)
Education	-.156 (0.003)
Wholesale, retail	.781 (0.020)

**Estimates of τ by Ownership Status Using the Full Distribution of
Establishment Size (All India)**

Ownership Type	$\tau - 1$
Government and PSU	-.087 (0.002)
Non Profit Institution	.021 (0.012)
Unincorporated Proprietary	.490 (0.005)
Corporate Non-Financial	-.134 (0.014)
Cooperative	.005 (0.014)

6 Discussion and Investigation of Mechanisms

6.1 Interpretation of Results

Thus far we have argued that the observed downshift in the distribution of establishments with 10 or more workers is related to the existence of certain labor and industrial regulations that become binding at that point. But if the regulations are responsible for the observed effect, then differences in the substance or application of the regulations should explain (at least part of) the great variation we observe across States and Industries.²⁹ In this section we explore these dimensions of variation with the goal of reaching a deeper understanding regarding the causes and consequences of the costs we have tried to estimate. The regressions

²⁹The variation across ownership types is straightforward to explain: the regulations are clearly not applied in the same way to privately owned enterprises and government enterprises. An additional explanation is that government establishments are not profit maximizing and thus would require a different motivational theory altogether to produce the observed power law distribution.

we do are cross-sectional and the variables used are endogenous, so the results cannot be given a causal interpretation, but we find them instructive nevertheless.

To preview our results, we do observe a correlation between our estimated costs (τ) and certain measures of the substance of the regulations. Moreover, we also find a robust and independent correlation between our estimated costs and several different measures of corruption/poor state governance, suggesting that it is not only the regulations themselves but also their enforcement and application that is responsible for the high costs we estimate. We also sketch a theoretical framework of bribery and extortion which casts light on the proper interpretation of our empirical results. Finally, we present some indicative evidence that our costs may have significant negative dynamic implications, as they are associated with lower growth in employment and productivity in registered manufacturing.

6.2 τ and Corruption: Evidence from the Interstate Variation

We start by regressing our state-level estimates of τ against other established measures of the regulatory environment (see Table 5).³⁰ These measures include the “Besley Burgess” measure of labor regulations (BB_97) from Aghion, et al (2008) and several measures from Dougherty (2009). BB_97 is a measure of the number of amendments that a state government has made to the Industrial Disputes Act in a “pro-worker” or a “pro-owner” direction, as interpreted by Aghion, et al (2008), who also update the measure to include amendments up to 1997.³¹ Positive values indicate more “pro-worker” amendments, which are assumed to imply a more restrictive environment for firms operating in those states. Dougherty (2009) also provides state level measures that reflect “the extent to which procedural or administrative changes have reduced transaction costs in relation to labor issues” (Dougherty, 2014). Higher values therefore indicate an improved environment for firms. Dougherty’s measures are unique in that they cover a wide range of labor-related issues - not just the IDA. In the analysis below, we will focus on measures from Dougherty (2009) that cover reforms regarding 1) inspectors, 2) the Factories Act and 3) an overall measure of reforms. All relevant variables in our analysis have been rescaled to have mean zero and standard deviation one, with the goal of allowing comparability between regression coefficients in different specifications.

In Table 5, correlations are reported between τ and all four measures both by themselves

³⁰Note that the estimates of τ we use in all the analysis below were generated using the procedure in Section 4.4 that we have argued is robust to possible misreporting and non-classical measurement error. Two estimates of τ were generated for each state: one using all the establishments in the state (“tau_npr” - *npr* refers to the non-parametric smoothing in the estimation process) and another using only establishments in the manufacturing sector (“tau_mnf_npr”).

³¹Since there have been few state-level amendments to the IDA between 1997 and 2005, this measure should be largely the same in 2005.

and while controlling for state GDP per capita and the state’s share of employment in manufacturing. The BB_97 measure does not seem to be correlated with τ (although our power is limited by the very small number of observations) while two of the measures from Dougherty (2009) are significantly correlated after applying controls (and have the “correct” sign).³² On the one hand the lack of correlation between τ and BB_97 is not surprising, as the latter capture variation only due to state amendments to the Industrial Disputes Act, which does not vary over the ten person threshold. On the other hand, if the Besley Burgess measure is meant to capture the general regulatory environment (which is how it is used in countless studies), we might well expect it to correlate with our measure of regulatory costs. That the correlation does not hold may therefore be of interest.

While the prediction regarding the correlation between τ and BB_97 may be ambiguous, that is not the case for Dougherty’s measure of transaction-cost reducing reforms related to the Factories Act (“Dougherty_FA”). We should expect our measure of τ to correlate negatively with “Dougherty_FA”, since the Factories Act does vary across the 10 worker threshold, and indeed we see that it does. τ is also correlated with “Dougherty_all”, the measure that aggregates reforms across all areas, although it does not appear to correlate with any other subcomponents (not all of which are depicted here).

Table 5: Tau vs Other Measures of Regulations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n
BB_97_n	-0.0863 (0.219)	0.00275 (0.185)						
Dougherty_all_n			-0.304 (0.235)	-0.437* (0.208)				
Dougherty_FA_n					-0.278 (0.213)	-0.397* (0.190)		
Dougherty_insp_n							-0.135 (0.226)	-0.337 (0.217)
sh_manuf_emp		-4.992 (4.706)		6.094 (3.929)		5.734 (3.916)		7.771 (4.464)
lnNSDPpc05		-0.539** (0.218)		-0.659** (0.247)		-0.661** (0.248)		-0.656** (0.266)
_cons	0.507** (0.219)	6.417** (2.086)	0.603** (0.218)	6.434** (2.350)	0.587** (0.216)	6.479** (2.358)	0.563** (0.224)	6.149** (2.483)
<i>N</i>	15	15	18	18	18	18	18	18

Standard errors in parentheses

Only including Major Indian States

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

³²In this and most of the analysis ahead, we focus on the 18 largest Indian States, for which data are most consistently available and which offer the most precise estimates of τ (the power law relationship breaks down in smaller states when there are not enough observations).

In addition to the above measures regarding state-level changes to the statutory, procedural and administrative aspects of the regulations, we also regress τ against certain other measures of the labor environment. Table 6 reports the results of τ regressed against per capita measures of strikes, man-days lost to strikes, lockouts and man-days lost to lockouts. One might imagine that strikes and lockouts capture relevant features of the regulatory and labor environment,³³ but we do not find them to be robustly correlated with τ .

One might also expect τ to be correlated with aspects of regulatory enforcement. To test this hypothesis we regress τ against state level variables related to enforcement such as the number of inspections, convictions, and fines levied under various regulations.³⁴ The results of the regressions for a subset of the enforcement related variables are shown in Table 7. In short, the only enforcement variable that is even close to being significantly correlated with τ is the percentage of factories registered under the Factories Act that have been inspected. However, as can be seen from the table, the enforcement data are only available for a small subset of the major states, leaving very little power in the regressions. Furthermore, the regressions shown exclude Uttar Pradesh, which is a substantial outlier in the enforcement data.

Table 6: Tau vs Strikes and Lockouts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n
strikes_pc_n	-0.349** (0.164)	-0.272 (0.168)						
mandays_lost_strikes_pc_n			-0.153 (0.181)	-0.223 (0.180)				
lockouts_pc_n					-0.0894 (0.168)	-0.157 (0.163)		
mandays_lost_lockouts_pc_n							-0.0778 (0.168)	-0.148 (0.160)
sh_manuf_emp		4.497 (4.128)		7.343 (4.730)		6.542 (4.568)		6.239 (4.501)
lnNSDPpc05		-0.410 (0.265)		-0.543* (0.270)		-0.562* (0.268)		-0.572* (0.270)
_cons	0.701*** (0.211)	4.200 (2.497)	0.618** (0.243)	5.149* (2.577)	0.575** (0.228)	5.379* (2.540)	0.571** (0.227)	5.516** (2.556)
N	18	18	17	17	18	18	18	18

Standard errors in parentheses

Only including Major Indian States

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

³³For example, some industrial regulations explicitly undermine or support the rights of parties to engage in strikes or lockouts.

³⁴These data were obtained from the 2005 Indian Labour Yearbook, which we were able to attain with the generous help of Anushree Sinha and Avantika Prabhakar of NCAER.

Table 7: Tau vs Enforcement of Regulations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n	tau_npr_n
perc_fac_insp_n	0.330 (0.202)							
num_conv_FA_tot_per_fac_n		-0.0315 (0.343)						
ln_num_pros_SEA_pc_n			0.107 (0.322)					
ln_fines_SEA_pc_n				0.413 (0.292)				
pros_per_inspec_n					0.00865 (0.232)			
fines_SEA_per_ins_n						3.004 (3.142)		
num_cases_dis_SEA_per_ins_n							-4.402 (15.03)	
cases_dis_per_pros_SEA_n								0.769 (1.003)
lnNSDPpc05	-0.122 (0.808)	-0.540 (1.056)	-1.938*** (0.546)	-2.572*** (0.669)	-1.850*** (0.491)	-2.222*** (0.598)	-1.762** (0.548)	-1.815*** (0.459)
sh_manuf_emp	-3.112 (6.243)	0.834 (10.53)	4.312 (4.376)	4.339 (3.882)	4.018 (4.467)	6.131 (4.663)	3.616 (4.429)	1.975 (4.898)
_cons	1.909 (7.708)	5.764 (9.917)	19.32*** (5.259)	25.78*** (6.638)	18.46*** (4.711)	22.62*** (6.172)	16.66* (7.508)	18.51*** (4.405)
N	10	9	13	13	13	13	13	13

Standard errors in parentheses

Only including Major Indian States (except UP) for which data exist.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To briefly summarize the results so far, when regressing τ against regulatory substance, enforcement or industrial disputes, one only observes correlations for certain specific types of measures of those phenomena. In contrast, in our remaining analysis, we will demonstrate that our measures of τ are strongly and robustly correlated with corruption, almost regardless of how it is measured. Table 8 reports the results of regressing τ against corruption as measured in a 2005 Transparency International (TI) Survey.³⁵ Column 1 includes all states for which there is data, while the remaining columns include only the 18 largest Indian states. Column 3 adds controls for state GDP per capita, share of manufacturing in employment and some others, while Column 4 adds the aggregate measure of regulatory reform from Dougherty (2009). With no exceptions, the coefficient on the TI corruption score is consistently significant and very large in magnitude: a one standard deviation increase in a state's corruption score is associated with a .65 standard deviation increase in τ . In particular, the fact that the coefficient remains significant in Column 4 even after controlling for Dougherty's measure of regulatory reform suggests that the relationship between

³⁵The TI corruption measure is based on a survey of perceptions and experience regarding corruption in the public sector.

τ and corruption is at least partly independent from the relationship with the regulations themselves.

In what follows, we will use the TI corruption score as our primary measure of corruption. However, we also regress τ against an alternative measure of corruption to check for robustness of the corruption relationship and to alleviate any concerns that the TI measure may be flawed as it is a result of individuals' perceptions and may therefore be biased. Table 9 reports the results of τ regressed against the percent of a state's available electricity that was lost in transmission and distribution in 2005/6. This variable has been used by other researchers as a proxy for corruption and poor state governance, and has the virtue of being a concrete and objective measure that does not depend on perceptions (Kochhar et al, 2006). As with the TI Corruption Score, the correlations between τ and this alternative measure of corruption is significant and large in magnitude regardless of sample or controls - including, again, the addition of the Dougherty measure of regulatory reform in Column 4.

Table 8: Tau vs Transparency International Corruption Score

	(1)	(2)	(3)	(4)
	tau	tau	tau	tau
TI Corruption Score	0.490** (0.180)	0.662*** (0.171)	0.659*** (0.157)	0.606*** (0.150)
log of Net State Domestic Product pc			-0.390* (0.196)	-0.445** (0.186)
share of employment in manufacturing			9.679** (4.100)	8.892** (3.863)
share of privately owned establishments			1.003 (4.917)	1.522 (4.607)
share of registered establishments			2.546** (1.131)	1.869 (1.134)
Regulatory Reforms from Dougherty 2009				-0.253 (0.153)
Constant	0.494** (0.175)	0.597*** (0.163)	1.664 (4.960)	2.077 (4.643)
Observations	20	18	18	18
States Included	All	Major	Major	Major

* p<0.10, ** p<0.05, *** p<0.01, Standard errors in parentheses

Table 9: Tau vs Transmission and Distribution Losses

	(1)	(2)	(3)	(4)	(5)
	tau	tau	tau	tau	tau
Electricity	0.329*	0.834***	0.886***	0.808***	0.894***
Transmission and Distribution Losses	(0.164)	(0.267)	(0.234)	(0.227)	(0.253)
log of Net State Domestic Product pc			-0.495**	-0.542**	-0.492**
			(0.205)	(0.196)	(0.216)
share of employment in manufacturing			10.76**	9.839**	10.61**
			(4.449)	(4.249)	(4.797)
share of privately owned establishments			0.843	1.415	0.872
			(5.249)	(4.979)	(5.484)
share of registered establishments			2.714**	2.011	2.725*
			(1.208)	(1.229)	(1.265)
Regulatory Reforms from Dougherty 2009				-0.256	
				(0.165)	
Electricity Available (GWH)					0.0204
					(0.171)
Constant	-3.09e-09	0.600***	2.676	2.988	2.620
	(0.162)	(0.179)	(5.313)	(5.029)	(5.565)
Observations	35	18	18	18	18
States Included	All	Major	Major	Major	Major

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, Standard errors in parentheses

Although the state-level correlations between τ and corruption appear to be robust, the regressions lack exogenous variation and are subject to the concern that our measures of corruption may be correlated with omitted variables that are also correlated with τ . To partially address these concerns, we attempt to take advantage of State X Industry level heterogeneity. In particular, inspired by Novosad and Asher (2014), we use 2005 World Bank Enterprise Survey (WBES) data to create an industry-level measure of “dependence on government bureaucracy”. Specifically, Indian firms in the 2005 WBES were asked how many times in a year they had an inspection or other required meeting with a government official. Averaging the firm-level responses by industry, we classify industries according to their average number of visits with officials (i.e., their dependence on government bureaucracy). If some industries have more meetings with officials, and if corruption takes place during some of these meetings, we would imagine that the costs of corruption would be highest for firms in those industries with the highest dependence on bureaucracy *and* in those states that have the highest levels of corruption. That is, we would expect that the *interaction*

between industry level dependence on bureaucracy and state level corruption is positive. If found to be the case, it would be harder to argue that the result is due to the presence of omitted variables.

The hypothesis is tested in Table 10. To do so we generate our measures of τ at the State X Industry level³⁶ and interact each of our three state level measures of corruption with a) the industry average number of visits from officials (“num_visits_total”) and b) the industry average duration of visits from officials (“dur_visits_total”). We include the interaction with average *duration* of visits as a placebo test: it is not clear that the duration of an inspection should be positively or negatively correlated with corruption.³⁷ We then regress our State X Industry measures of τ against the covariates including interaction terms. Our prior is that the interaction of corruption with duration of visits should be less significant than the interaction with number of visits. Indeed, this is mostly what we observe. The interaction between our measures of corruption and the number of visits is at least weakly significant (at the 10% level) for one of the two measures, while the interaction between corruption and average duration of visits is never significant.

To summarize our results from these investigations, we find:

1. a correlation between τ and certain aspects of the substance of regulations as measured in Dougherty (2009),
2. a nonexistent or inconclusive relationship between τ and measures of the labor environment and enforcement of regulations, and
3. a strong and robust relationship between τ and two distinct measures of corruption.

Although none of these results can be said to be causal, we find them suggestive of a relationship between corruption and high labor costs. Next, we turn our attention to the question of how and why greater corruption would lead to higher labor costs for firms. To this end, in the following subsection we outline a simple theoretical framework to elucidate the potential connection.

³⁶Industries here are categorized according to their groupings in the World Bank Enterprise Surveys, which distinguishes 24 distinct industry categories. Examples include “auto components”, “leather and leather products”, and “food processing”.

³⁷In particular, corruption may lead to longer inspections if the process of extracting the bribe takes time, or it may lead to shorter inspections, if there is no need to carry out the actual inspection.

Table 10: Tau vs State Level Corruption Interacted with Industry Level “Dependence on Regulation” (with Industry FEs)

	(1)	(2)	(3)	(4)
	tau_npr	tau_npr	tau_npr	tau_npr
corruption_score	0.132*	0.140*		
	(0.0668)	(0.0666)		
td_losses_05			0.0583	0.0493
			(0.114)	(0.110)
num_visits_total	0.127		0.130	
	(0.0829)		(0.0945)	
dur_visits_total		-0.00610		-0.109
		(0.124)		(0.143)
corrupt_scXnum_visits_tot	0.101*			
	(0.0545)			
corrupt_scXdur_visits_tot		0.0513		
		(0.0691)		
TDLsXnum_visits_tot			0.0112	
			(0.0848)	
TDLsXdur_visits_tot				-0.0698
				(0.107)
lnNSDPpc05	-0.185***	-0.185***	-0.197***	-0.194***
	(0.0364)	(0.0374)	(0.0428)	(0.0427)
_cons	-0.244	0.0235	-0.235	-0.217
	(0.187)	(0.442)	(0.187)	(0.498)
<i>N</i>	189	189	189	189

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, Standard errors clustered at the State Level, Industry FEs included.

6.2.1 A Theoretical Framework for Understanding Corruption Between Inspectors and Firms

We consider two distinct types of corruption that could take place between inspectors and firms: 1) Bribery and Collusion, and 2) Bribery and Extortion. Collusion takes place when inspectors allow firms to avoid costs due to regulation in exchange for bribes. However, poor state governance (which here would imply an inability to control corruption) would then lead to *lower* costs for firms, as greater corruption would make it easier to avoid the

full costs of regulation.³⁸ However, what we observed above in Section 6.2 was a robust *positive* correlation between effective costs (τ) and poor governance/corruption. To explain this phenomenon, we employ a model of *extortion*.

We start with the observation that all firms reporting at least 10 employees fall under the jurisdiction of certain regulations. Imagine, now, that the regulations are so complex so as to make it impossible (or prohibitively costly) for any firm to be fully in compliance with all aspects of the law as written.³⁹ Therefore, an inspector can, at any time, choose to subject a firm under his jurisdiction to a penalty e , which may include financial (e.g.: fines) and/or non-financial elements (e.g.: harassment, time needed to defend claims of violations, etc). We can think of the extent of the penalty (e) as a function of state governance: properly functioning governments hire and motivate inspectors to pursue substantive violations rather than minor ones, while inspectors in corrupt or dysfunctional governments can get away with threatening to arbitrarily impose high penalties if a bribe is not paid (i.e.: extortion).

In such an environment, firms reporting 10 or more employees (and hence under the jurisdiction of the inspector) may face a choice between exposing themselves to the penalty, e , or paying a bribe, b . Assume that inspectors face no costs or benefits from imposing the penalty on the firm, but naturally benefit from receiving the bribe. There is thus a surplus to be had from paying/receiving the bribe b and avoiding the penalty. If the inspector and firm Nash Bargain over the surplus with bargaining weights α and β , respectively, the problem is the following:

$$\max_b (b)^\alpha (e - b)^\beta$$

The solution is for the firm to pay a bribe $b = \frac{\alpha}{\alpha + \beta} * e$. The cost born by the firm is therefore increasing in α , the bargaining weight of the inspector, and in e , the maximum penalty to which the firm can be subjected. It is reasonable to imagine that this maximum level of extortion, e , is roughly proportional to the size of the firm, so that $e = e' * n$, where n is the number of workers in the firm and e' is the per worker level of extortion. In that case the bribe per worker, $\frac{b}{n}$, is equal to $\frac{\alpha}{\alpha + \beta} * e'$.

This framework can be embedded into the firm's choice of true and reported employment as modeled in Section 4.3. In particular, the firm now faces a choice between reporting

³⁸See, for example, a model of corruption such as the one in Khan, Khwaja and Olken (2014).

³⁹This does not not require much imagination. As we mentioned in Section 2, many of the laws have components that are antiquated, arbitrary, contradictory and confusing. That the laws may be impossible to fully comply with is suggested by the many anecdotes we provide in Appendix 2 as well as the following observation, which we re-quote: "Rules under the Factories Act, framed in 1948, provide for white washing of factories. Distemper won't do. Earthen pots filled with water are required. Water coolers won't suffice. Red-painted buckets filled with sand are required. Fire extinguishers won't do... And so on" (TeamLease Services, 2006).

employment greater than or less than 10, where reporting less than 10 allows it to avoid the costs of bribery, and reporting greater than 10 exposes it to the bribery costs. In that framework, τ corresponds to $\frac{b}{n}$, and is therefore increasing in the bargaining power of the inspector (α) and the corruption level of the state (e'). In this way, we can make sense of the empirical results above in terms of this basic framework. A more fully fleshed out model is provided in Appendix 3.

6.3 Possible Consequences of τ

In the subsections above, we tried to argue that our estimated costs (τ) are most likely due, not only to the substance of the regulations themselves, but also to high levels of corruption. In this subsection we will indicate possible consequences of high values of τ . Again, the results cannot be given a causal interpretation, but we find them compelling nevertheless. Table 11 displays the results of employment growth in the manufacturing sector at the State Level regressed against our measures of labor market distortions (τ and τ_{manuf}) as well competing measures (“BB_97” and “Dougherty_all”). For each of the four measures, we observe its performance as a predictor of future employment growth in *registered* manufacturing as well as its correlation with employment growth in *unregistered* manufacturing. Interestingly, in the regressions of employment growth in registered manufacturing against τ and τ_{manuf} , the coefficient on τ is negative and at least weakly significant, while the coefficient for employment growth in unregistered manufacturing is positive - significantly so in the case of τ_{manuf} . This result makes sense: we should expect higher costs to negatively effect the sectors to which the costs apply - in this case the registered sector, since that is under the ambit of labor regulations while the unregistered sector is not. If these correlations reflect a causal chain, it would mean that high levels of corruption (as measured by τ) are pushing employment from the registered to the unregistered sector.

Also included in Table 11 are the results of employment growth in manufacturing regressed against BB_97 and Dougherty_all. Neither regressor has a coefficient that is statistically significant or of meaningful magnitude.⁴⁰ Given the caveat that none of these results has the virtue of exogenous variation, it does appear to be the case that our measures of labor market distortions do a better job of predicting future employment growth (or the lack thereof) than the established alternatives. This is also true when considering future growth in manufacturing *productivity* rather than employment, as shown in Table 12. The magnitudes of the coefficients on τ do not change when looking at growth in the unregistered

⁴⁰It is not quite fair to regress growth between 2010 and 2005 on a regressor that uses data only up until 1997 (as is the case for BB_97), but a) we have duplicated these results using from growth from 1997 to 2002 and the results are the same, and b) the Besley Burgess measure from Aghion et al (2008) should be largely the same in 2008 due to the lack of state level reforms between 1997 and 2005.

sector as opposed to the registered, but the significance levels are lower, as we might expect. Other than that, the results are quite the same as in Table 11.

Table 11: Growth in Manufacturing Employment Between 2005 and 2010 vs Tau and Other Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	reg_mf_e	unreg_mf_e	reg_mf_e	unreg_mf_e	reg_mf_e	unreg_mf_e	reg_mf_e	unreg_mf_e
tau_npr_n	-0.0212 (0.0150)	0.0154 (0.0186)						
tau_mnf_npr_n			-0.0471** (0.0217)	0.0623** (0.0256)				
BB_97_n					-0.00525 (0.00731)	0.00979 (0.0142)		
Dougherty_all_n							0.0226 (0.0130)	-0.0143 (0.0159)
lnNSDPpc05	0.00353 (0.0175)	0.0250 (0.0207)	0.0107 (0.0145)	0.0192 (0.0161)	0.00413 (0.00863)	0.0140 (0.0168)	0.0212 (0.0154)	0.0136 (0.0195)
sh_manuf_emp	-0.357 (0.263)	-0.0402 (0.326)	-0.708** (0.258)	0.435 (0.325)	0.0194 (0.186)	-0.559 (0.362)	-0.515* (0.245)	0.0525 (0.323)
_cons	0.0855 (0.167)	-0.246 (0.197)	0.0372 (0.139)	-0.229 (0.152)	0.0209 (0.0825)	-0.0675 (0.160)	-0.0861 (0.147)	-0.131 (0.182)
N	18	17	18	17	15	15	18	17

* p<0.10, ** p<0.05, *** p<0.01, Standard errors in parentheses, Only including Major Indian States

Table 12: Growth in Manufacturing Productivity Between 2005 and 2010 vs Tau and Other Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	reg_mf_p	unreg_mf_p	reg_mf_p	unreg_mf_p	reg_mf_p	unreg_mf_p	reg_mf_p	unreg_mf_p
tau_npr_n	-0.0270* (0.0126)	-0.00918 (0.0194)						
tau_mnf_npr_n			-0.0512** (0.0181)	-0.0567* (0.0275)				
BB_97_n					-0.00266 (0.0122)	-0.00372 (0.0154)		
Dougherty_all_n							0.0167 (0.0122)	0.00973 (0.0166)
lnNSDPpc05	-0.0148 (0.0147)	-0.0152 (0.0217)	-0.00478 (0.0121)	-0.0120 (0.0174)	-0.00455 (0.0143)	-0.00723 (0.0182)	0.00438 (0.0145)	-0.00793 (0.0204)
sh_manuf_emp	0.248 (0.221)	0.0696 (0.340)	-0.154 (0.215)	-0.349 (0.350)	0.453 (0.310)	0.526 (0.392)	0.0728 (0.230)	0.0104 (0.338)
_cons	0.119 (0.140)	0.190 (0.206)	0.0454 (0.116)	0.198 (0.163)	-0.0208 (0.137)	0.0530 (0.174)	-0.0678 (0.138)	0.119 (0.190)
N	18	17	18	17	15	15	18	17

* p<0.10, ** p<0.05, *** p<0.01, Standard errors in parentheses, Only including Major Indian States

6.4 Inter-temporal Variation: A Final Puzzle

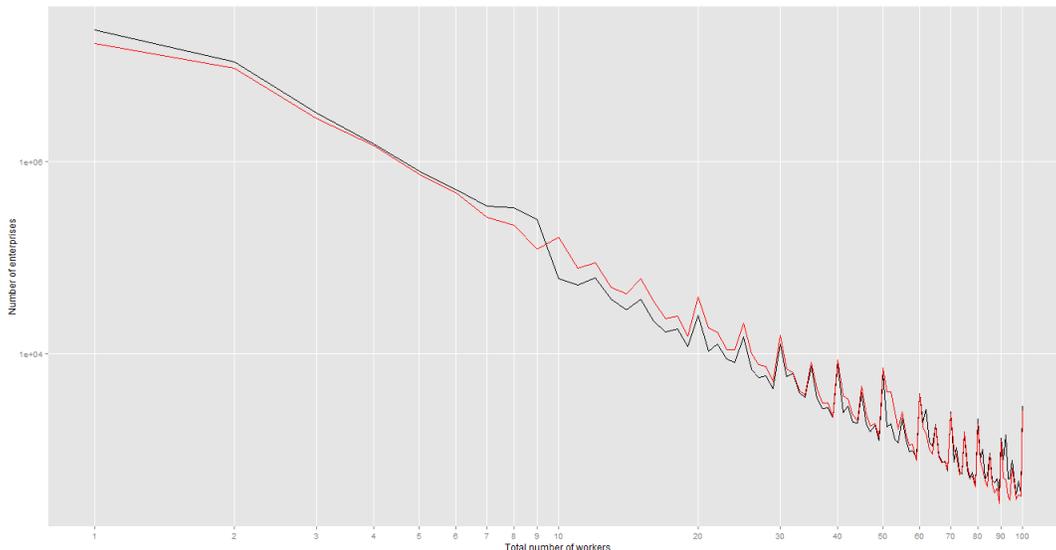
As we have noted, all of the analysis above uses data only from the 2005 Economic Census. However, the discontinuity observed in the 2005 data does not appear in the 1990 or 1998 ECs. Explaining the puzzling inter-temporal variation is a main goal of our future work. Perhaps the simplest explanation is that the data quality of the EC improved between 1998 and 2005. Indeed, the presence of the downshift can easily be erased by introducing measurement error. This explanation has received anecdotal support from our meetings with the Directors of the state Directorates of Economics and Statistics, who largely appeared to be more confident in the results of the 2005 EC than in the results of previous rounds. One reason for this could be that the 2005 EC was the first in which ICR (intelligent character recognition) was used to read in the raw data.

One might also be concerned in the other direction: that somehow the distortion observed in 2005 is anomalous and is perhaps due to something like a change in the EC's enumeration practices. Based on our interviews with officials and enumerators in charge of collecting these data, however, we believe this is unlikely. While we have discovered that enumeration practices have changed slightly over the years, we have not discovered changes that could have produced the specific patterns we observe. For example, unlike the previous waves, the 2005 Economic Census included an "address slip" that was meant to be filled out for establishments with 10 or more workers. It is conceivable that enumerators, in an effort to avoid the extra work of filling out the address slip, preferred to misrepresent the number of workers for establishments with more than 10 workers. However, while this could reasonably explain why there are fewer 10, 11 and 12 person establishments, we find it hard to understand how this kind of phenomenon could explain why there are also fewer 40, 50, 100 and 200 person establishments (see section 4.3 for more on this). Furthermore, in post-enumeration checks done in West Bengal, Bihar and Tamil Nadu, this kind of misrepresentation was not found to be in occurrence.

Nor does the culprit seem to be changes in the regulations themselves. Indeed, none of the regulations which we assume are responsible for the discontinuity in 2005 were greatly changed *in the right direction* between 1990 and 2005. There were reforms on the margin (reflected in the Dougherty (2009) measures), and these do seem to be correlated with our measures of τ , but most of the changes have gone in the direction of loosening regulations and thus cannot explain why τ would have increased over time. Enforcement has also changed to some degree, but we have not yet found evidence that it can explain the variation in our data (see the Interstate analysis above).

Another possible (but ultimately unlikely) explanation is that changes in the competitive environment, particularly related to the increased exposure to international markets and

Figure 7: Variation in the distribution of establishment size across time: 1998 (red) and 2005 (black)



competition, are responsible. The period in question (1998 to 2005) saw heavy reductions in protective barriers from foreign competition - particularly through the elimination of non-tariff barriers. However, in our preliminary analysis (not yet reported) we do not find a strong link between trade liberalization and τ , and such link as exists goes in the opposite direction.

A final possibility is that changes in the availability of contract labor cause the discontinuity to show up in 2005. Indeed, there was a large speed-up in the use of contract labor over the period,⁴¹ however state-level changes in the fraction of contract labor in registered manufacturing are not robustly correlated with τ . The absence of such a correlation is not necessarily evidence that such a link does not exist, but it does not give support to the hypothesis either.

We have been able to share this intertemporal paradox with a number of experts regarding these issues in India but have not yet been able to find a watertight explanation. This continues to be a priority in our ongoing work.

7 Conclusion

Our goals in this paper are 1) to document the effect of size-based labor regulations on the misallocation of resources between firms via the employment decisions of business enterprises,

⁴¹In the registered manufacturing sector, the share of contract labor in total labor increased from 15% to 26%. Prior to this period, its growth was markedly slower.

2) to estimate the net costs of the set of regulations that become binding when establishments choose to employ 10 or more workers, and 3) to shed some light on the source of these costs by demonstrating that corruption in the form of harassment bribery may play a large role in making Indian regulations costly. To the best of our knowledge, this paper is the first to provide cost estimates of regulations in India (particularly non-IDA regulations), the first to analyze the effects of regulations without using necessarily subjective evaluations of state-level amendments to labor laws, and the first to provide evidence for the link between regulatory costs and corruption.

To accomplish these tasks, we use the 2005 Economic Census of India, an uncommonly used dataset which is uniquely suited to our task because it includes the entire universe of non-farm enterprises. In our investigation, we find a significant level shift down in the natural log of the probability mass of establishments with 10 or more workers. Adapting a method from Garicano, LeLarge and Van Reenen (2013), we interpret this as evidence of substantial per-worker costs of operating above the 10 worker threshold. At the all-India level, we find that operating at or above the 10-worker threshold is associated with an 37% increase in the unit cost of labor as modeled. Furthermore, we observe a great deal of variation in our estimated costs by state, industry and ownership type. We estimate the highest costs for privately owned firms (the lowest costs for government-owned firms) and firms in construction (lowest costs for those in public administration and defense).

We also find that Indian states with the highest costs also have the highest levels of corruption and poor governance (as measured through a variety of different indices), and that firms in industries with high bureaucratic dependence are exposed to particularly high costs if they are also in highly corrupt states. Our analysis suggests that the size of regulatory costs may have as much to do with how regulations are implemented and who implements them, as with the content of the specific labor and industrial laws themselves. Finally, we demonstrate an inverse correlation between the size of our estimated costs in a state and the growth rate of employment in the *registered* (but not the unregistered) manufacturing sector, suggesting that higher costs discourage formal manufacturing activity and encourage informality.

Appendix 1: Full Results by State and Industry

[To Be Added]

Appendix 2: Anecdotal Evidence Regarding Harassment Bribery from “ipaidabribe.com”

“I am a small factory owner in Kirti Nagar Industrial Area. We follow almost all rules laid down by government for the welfare of workers. Now, even if we follow everything there is always somethings where we lack and which needs improvement. We have a factory inspector by the name of Mr.R.B.Singh (M: 9818829355). He comes to all the factories in our area, inspects them, find mistakes and then harass and blackmails us. According to him he can get our factories sealed. To avoid this, to save our time and to save the unnecessary paperwork we pay him every year. I have paid him twice in two years i.e. 10000 & 15000 and this is common with all factories. Please take a strict action against him so that he learns a lesson. I am sure he is not alone. All his colleagues are equally corrupt.”

(Reported on August 11, 2014 from New Delhi, Delhi | Report #131791)

“During the routine labor verification process by the labor department at our office, we were advised by the consultant to pay the labor inspector a bribe to ensure that they don’t keep calling us for needless paperwork.”

(Reported on June 28, 2011 from Chennai, Tamil Nadu | Report #35064)

“The Labour Department requires a dozen odd registers to be maintained some of them which are totally outdated and pointless. E.g: Salary register, Attendance register, Leave register etc.

Our IT office has an electronic system that logs all entries/exits and leave taken. We have the records and offered to provide it to them in a printout.

Salaries are paid electronically via bank transfer.

The officer declined and said it must be maintained in a manual register!

Finally an arrangement was made where we maintain a few records manually and the rest he would overlook.

Cost of arrangement Rs 1500 twice a year even if the officer shows up only once a year for the inspection!

He is supposed to inspect twice so expects to be paid even for the time he did not show up!” (Reported on October 13, 2010 from Chennai, Tamil Nadu | Report #44950)

“Well i had gone to renew my labour license and after all the running around in the bank and the department, the signing authority asked me to pay Rs.500 for signing. When asked

why 500, i was told since there are 5 employess for Rs.100 each.” (Reported on December 31, 2010 from Hyderabad, Andhra Pradesh | Report #43509)

“... in my third visit i met one of office peon in Labour office he guided me for the bribe he also investigated and *advised me for bribe according to the number of Employees deployed on contract basis* and for this valueble suggestion he charged me Rs. 100. Again with full confidence i went to the ALCs desk and straight away i offered him the packet which was contains the amount of Bribe Rs. 3000/- ... He issued me the license after office hours...” (Reported on March 30, 2011 from Mumbai, Maharashtra | Report #39133)

Appendix 2: Modeling Harassment Bribery

References

- [1] Adhvaryu, A., Chari, A., & Sharma, S. (2013). Firing Costs and Flexibility. *The Review of Economics and Statistics* (forthcoming).
- [2] Aghion, Philippe, Robin Burgess, Stephen Redding, and Fabrizio Zilibotti (2008), "The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India", *American Economic Review*, American Economic Association, 98(4).
- [3] Ahsan, Ahmad & Pagés, Carmen (2009). "Are all labor regulations equal? Evidence from Indian manufacturing," *Journal of Comparative Economics*, Elsevier, vol. 37(1), pages 62-75, March.
- [4] Ahsan & Mitra (2014), "Trade Liberalization and Labor's Slice of the Pie: Evidence from Indian Firms," *Journal of Development Economics*, 108: 1-16.
- [5] Bardhan, Pranab (2014). "The Labour Reform Myth." *The Indian Express* [New Delhi] 23 August. <http://indianexpress.com/article/opinion/editorials/the-labour-reform-myth/>
- [6] Basu, Kaushik (2011). "Why, for a Class of Bribes, the Act of Giving a Bribe should be Treated as Legal", Working Paper.

- [7] Besley, T., & Burgess, R. (2004). Can labor regulation hinder economic performance? Evidence from India. *The Quarterly Journal of Economics*, 119(1), 91–134.
- [8] Besley, Timothy & McLaren, John, 1993. "Taxes and Bribery: The Role of Wage Incentives," *Economic Journal*, Royal Economic Society, vol. 103(416), pages 119-41, January.
- [9] Bhattacharjea, A. (2006). Labour market regulation and industrial performance in India: A critical review of the empirical evidence. *Indian Journal of Labour Economics*, 49(2).
- [10] Bhattacharjea, A. (2009). How do Indian firms respond to employment protection legislation? *Economic and Political Weekly*. Vol XLIV No. 22.
- [11] Bollard, Klenow & G. Sharma (2013). India's mysterious manufacturing miracle. *Review of Economic Dynamics*, 16:159-85.
- [12] Cain, J. Salcedo, Rana Hasan & Devashish Mitra, 2010. "Trade Liberalization and Poverty Reduction: New Evidence from Indian States," Working Papers 3333, School of International and Public Affairs, Columbia University.
- [13] Chatterjee, Urmila & Ravi Kanbur (2013). Regulation and Non-Compliance: Magnitudes and Patterns for India's Factories Act. Working Paper.
- [14] Chaurey, Ritam (2013). Labor Regulations and Contract Labor Use: Evidence from Indian Firms. Working Paper.
- [15] de Waldemar, F. S. (2012), New Products and Corruption: Evidence from Indian Firms. *The Developing Economies*, 50: 268–284. doi: 10.1111/j.1746-1049.2012.00171.x
- [16] Dougherty, S. (2009). Labour Regulation and Employment Dynamics at the State Level in India. *Review of Market Integration* 1: 295.
- [17] Dougherty, S., Robles, V., & Krishna, K. (2011). Employment Protection Legislation and Plant-Level Productivity in India. *OECD Economics Department Working Papers*, No. 917.
- [18] Eaton, J., Kortum, S., & Kramarz, F. (2011). An anatomy of international trade: Evidence from French firms. *Econometrica*, 79(5).
- [19] Fallon and Lucas (1993).
- [20] FICCI (2004), "FICCI Survey on Priority Issues for National Manufacturing Competitiveness Council", Federation of Indian Chambers of Commerce and Industry, August.

- [21] Garicano, L., C. LeLarge, and J. Van Reenen (2013). Firm size distortions and the productivity distribution: Evidence from France. Working Paper.
- [22] Goldar, B., & Aggarwal, S. (2012). Informalization Of Industrial Labor In India: Effects Of Labor Market Rigidities And Import Competition. *The Developing Economies*.
- [23] Gourio, F., & Roys, N. (2012). Discontinuity in Labor Laws, Firm Size Distribution and Reallocation. Working paper.
- [24] Hasan, R. & K. Jandoc (2012). Labor Regulations and the Firm Size Distribution in Indian Manufacturing. Working Paper.
- [25] Hasan, R., D. Mitra & K. V. Ramaswamy (2007). Trade Reforms, Labor Regulations, and Labor-Demand Elasticities: Empirical Evidence from India. *The Review of Economics and Statistics*. Vol. 89, No. 3, Pages 466-481
- [26] Hernandez-Perez, R., F. Angulo-Brown, and D. Tun (2006, January). Company size distribution for developing countries. *Physica A: Statistical Mechanics and its Applications* 359, 607–618.
- [27] Hsieh, C., & Klenow, P. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4).
- [28] Hsieh, Chang-Tai, and Benjamin Olken (2014). The Missing "Missing Middle". NBER Working Paper No. 19966.
- [29] Khan, Adnan, Asim Khwaja, and Benjamin Olken (2014). Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors. Working Paper.
- [30] Kochhar, Kalpana, Utsav Kumar, Raghuram Rajan, Arvind Subramanian, and Ioannis Tokatlidis (2006), "India's Pattern of Development: What Happened, What Follows?", *Journal of Monetary Economics*, Vol. 53, No. 5, pp. 981-1019.
- [31] Krueger, A. O. (2007). The Missing Middle. In N. C. Hope, A. Kochhar, R. Noll and T. N. Srinivasan (Eds.), *Economic Reform in India: Challenges, Prospects, and Lessons* (pp. 299-318).
- [32] Lucas, R. E., Jr. (1978). On the size distribution of business firms. *The Bell Journal of Economics* 9(2).
- [33] Markovitch, N. M., & Krieger, U. R. (2000). Nonparametric estimation of long-tailed density functions and its application to the analysis of World Wide Web traffic. *Performance Evaluation*, 42(2-3), 205–222.

- [34] Ministry of Finance, Government of India (2013). Seizing the Demographic Dividend. In *Economic Survey 2012-13*.
- [35] Mookherjee, Dilip & I. P. L. Png (1995). Corruptible Law Enforcers: How Should They Be Compensated? *The Economic Journal* 105.
- [36] Nagaraj, R. (2004), “Fall in Organised Manufacturing Employment: A Brief Note”, *Economic and Political Weekly*, 24 July.
- [37] Novosad, Paul and Asher, Sam (2014). Politics and Local Economic Growth: Evidence from India. Working Paper.
- [38] Ramaswamy, K. V. (2013), “Understanding the ‘missing middle’ in Indian Manufacturing: The Role of Size-Dependent Labour Regulations and Fiscal Incentives”, Working Paper
- [39] Sharma, Alakh (2006). Flexibility, Employment and Labour Market Reforms. *Economic and Political Weekly*. Vol 41, No. 21.
- [40] TeamLease Services (2006). India Labour Report 2006: A Ranking of Indian States by their Labour Ecosystem.
- [41] World Bank CII (2004), Improving the Investment Climate in India, The World Bank, Washington DC.