COLUMBIA | SIPA
Center for Development Economics and Policy

CDEP-CGEG WORKING PAPER SERIES

CDEP-CGEG WP No. 40

Formal Sector Incentives and Informality

Ama Baafra Abeberese and Ritam Chaurey

August 2017

COLUMBIA | SIPA
Center on Global Economic Governance

Formal Sector Incentives and Informality*

Ama Baafra Abeberese†

Ritam Chaurey‡

July 2017

* We would like to thank participants at the Eastern Economic Association Annual Meeting and the Entrepreneurship and Private Enterprise Development in Emerging Economies Workshop and seminar participants at Wellesley College for helpful comments. All errors are our own.

† Department of Economics, Wellesley College, Wellesley, MA 02481. Email: aabebere@wellesley.edu.

‡ Department of Economics, SUNY Binghamton, Binghamton, NY 13905. Email: rchaurey@binghamton.edu.

Abstract

Informal firms account for over half of output and employment in developing countries. To analyze the barriers to formalization, most of the literature has focused on entry costs in the form of registration fees, rather than on the ongoing costs and benefits of being a formal firm. This paper is the first to study the effect of a simultaneous change in ongoing formal sector costs and benefits on formalization. We analyze an Indian scheme that provided tax exemptions and capital subsidies to formal firms, thereby increasing the benefits and reducing the costs of formal firms. Using a complete enumeration of firms and a difference-in-differences approach, we find that the scheme led to an increase in the registration of existing informal firms, in particular, male-owned firms, urban firms and firms with access to external financing. We find some evidence that the increase in registration was primarily driven by the tax exemption rather than the capital subsidy.

JEL Codes: H20, O12, O17

Keywords: Informality, Taxation, India, Regional Policy

1 Introduction

Informal firms are generally defined as those that are not registered with any government entity. These firms are estimated to account for over half of output and employment in developing countries. In India, the context for this study, informal firms generate half of output and over 90 percent of jobs (Sharma and Chitkara, 2006). Despite the pervasiveness of informal firms in developing countries, there has been relatively little research on this subset of firms. Recent papers (see, for example, Bruhn (2011), Kaplan, Piedra, and Seira (2011), Fajnzylber, Maloney, and Montes-Rojas (2011), de Mel, McKenzie and Woodruff (2013), Galiani, Melendez and Navajas (2015), Campos, Goldstein and McKenzie (2015), de Andrade, Bruhn and McKenzie (2016) and Rocha, Ulyssea, and Rachter (2016)) have attempted to identify the constraints that prevent firms from registering. Theories on these constraints range from the view that firms decide to remain informal because the costs of adhering to government regulations, such as those related to taxes and labor, outweigh the benefits of being formal, such as access to credit, to the view that firms may be unable to afford the initial entry fees associated with registration.

In this paper, we seek to examine whether an increase in the ongoing benefits and a reduction in the ongoing costs of being formal, induced by a tax and capital subsidy policy in India, had an impact on the decision of informal firms to become formal. We also examine whether this impact depends on firm characteristics including the gender and social caste of the firm's owner, the firm's location and the firm's access to external financing. While the existing literature has largely focused on the view that the initial entry costs associated with registration may affect the decision of firms to formalize, there has been little evidence on how the ongoing costs and benefits from formalizing affect this decision. Two notable exceptions are Fajnzylber, Maloney, and Montes-Rojas (2011) and Rocha, Ulyssea, and Rachter (2016), which analyze the impact on formalization of a reduction in the ongoing costs of being a formal firm induced by a Brazilian tax simplification program for small firms. However, to our knowledge, there has been no study that analyzes a simultaneous change in both the ongoing costs and benefits of being a formal firm. The contribution of our paper is, therefore, to fill this void in the literature. In addition, we are able to analyze whether the

primary driver of formalization is a reduction in the ongoing costs or an increase in the ongoing benefits of being a formal firm.

The context for our paper is India's New Industrial Policy for the states of Uttarakhand and Himachal Pradesh through which the central government provided tax exemptions and capital subsidies to formal firms (both new and existing) starting in 2003. One of the benefits of being formal is the ability to more easily access formal sources of credit. Therefore, this policy, which provided formal firms financing for capital investment, represented an increase in the benefits of being formal. A widely cited cost of formalization is the tax burden on formal firms. The policy's tax exemption package, thus, represented a reduction in the cost of being a formal firm. Therefore, given its impact on the ongoing costs and benefits of being formal, one might expect this policy to affect the formalization decision of firms. This context is particularly appropriate for studying the impact of formal sector incentives on the formalization of informal firms. This is primarily because the policy was directed toward larger formal firms and the aim was not to increase registration by informal firms. This characteristic of the policy, therefore, reduces the endogeneity concerns that might arise, for instance, in Fajnzylber, Maloney, and Montes-Rojas (2011) and Rocha, Ulyssea, and Rachter (2016) where the policy studied specifically targeted small firms, which are mostly informal.

We use the Economic Census of India, which is a complete enumeration of all firms, both formal and informal, and a difference-in-differences strategy to analyze the relative change in the registration status of informal firms before and after the policy change in the states affected by the policy compared to the states that were not. We find that this policy led to an increase in the registration of firms. This impact on registration was heterogeneous. Male-owned firms and firms with owners who are not from historically disadvantaged groups[1] were more likely to register in response to this policy. Firms in urban areas were also more likely to register compared to firms in rural areas. We also find that firms with access to external financing were more likely to register following the policy.

[1] These are people from Scheduled Castes and Scheduled Tribes.

To determine whether the increase in the registration of firms was driven by pre-existing informal firms becoming formal or by the entry of new firms, we use data from India's National Sample Survey Organisation's survey of the unorganized manufacturing sector to analyze changes in the number of young firms and find that there was no differential change in the states affected by the policy relative to those that were not. This result suggests that the increase in the registration of firms was primarily driven by the registration of pre-existing informal firms.

To analyze whether the increase in registration was primarily driven by the tax exemption or capital subsidy, we examine whether the likelihood of registration varied with the capital intensity of the firm's industry. We find that firms in more capital-intensive industries were no more likely to register, suggesting that, at least in the short run, the primary driver of registration was the tax exemption.

Overall, our findings support the hypothesis that, even in the absence of any change to the initial entry costs related to registration, changes in the ongoing costs and benefits of being a formal firm can affect the formalization of informal firms. Moreover, we find suggestive evidence that changes in the ongoing costs of being a formal firm may be more effective at inducing formalization than changes in the ongoing benefits. However, not all firms are responsive to these changes. In particular, female-owned firms and firms with owners from historically disadvantaged groups are not as responsive suggesting that further research is needed to understand the constraints that hinder the formalization of this subset of firms.

The next section of the paper provides a brief overview of the existing schools of thought and literature on informal firms. Section 3 describes the scheme in India that we study. Section 4 describes the data. Section 5 presents the empirical strategy and results and Section 6 concludes.

2 The Informal Sector

An informal firm is generally defined as one which is not registered with any government entity and, hence, is not subject to stringent government oversight and regulation. Informal firms are, therefore, generally able to avoid paying taxes and complying with government regulations such as

labor regulations and environmental standards. In developing countries, these informal firms account for a large portion of economic activity. Given the ability of informal firms to avoid paying taxes, their prevalence can have negative implications for a country's tax revenues. Therefore, one incentive for policymakers to reduce the size of the informal sector is the potential to substantially increase the country's tax base. Another concern about the existence of informal firms is that this can result in an inefficient allocation of resources from formal to informal firms (Hsieh and Klenow, 2009). Informal firms can take away market share from formal firms since they are able to operate with lower marginal costs by avoiding the costs associated with taxes and regulations faced by formal firms. Other incentives for reducing the size of the informal sector include the potential to increase compliance with labor and environmental standards.

Given these concerns, policymakers have attempted to encourage the formalization of firms. An effective policy depends critically on the reasons for the existence of these informal firms. According to the legalist view, informal firms would like to formalize but are unable to cover the costs of registration (de Soto, 1989, 2000). The voluntarist view, on the other hand, suggests that firms voluntarily remain informal in order to avoid the costs faced by formal firms such as taxes and compliance with government regulations (Maloney, 2004). For these firms, the costs of being a formal firm outweigh the benefits such as access to credit, contract enforcement, stronger property rights and eligibility for government programs.

These reasons for the existence of informal firms suggest that encouraging formalization could be achieved through a reduction of the initial entry costs of becoming formal associated with registration fees and procedures, a reduction in the ongoing costs associated with taxes and compliance with regulations, and an increase in the benefits of formalization. In line with the view that the main barrier to formalization is the initial entry cost, policymakers are increasingly reducing the initial costs of formalization by streamlining registration procedures and reducing registration fees in an attempt to reduce the size of the informal sector. According to the World Bank's Doing Business report, which tracks reforms in about 189 economies, reforms undertaken by policymakers resulted in a reduction in the number of days for starting a business from an

average of 51 in 2003 to 20 in 2015 (World Bank, 2016).

However, the literature on easier and less costly registration processes has generally found that these have little impact on the registration of informal firms. Bruhn (2011) finds that a reform that simplified business registration procedures in Mexico and reduced the number of days required to register a business from about 30 to less than two led to the opening of new firms but had little impact on the registration of informal firms. Kaplan, Piedra, and Seira (2011) study the same reform in Mexico and similarly find that it led to the creation of new formal firms but are unable to prove if these firms were pre-existing informal firms. de Mel, McKenzie and Woodruff (2013) carry out a randomized evaluation in Sri Lanka and find that providing information about registration processes and reimbursing registration costs did not increase the registration of informal firms. de Andrade, Bruhn and McKenzie (2016) also carry out a randomized evaluation in Brazil and find that providing information on registration and eliminating registration costs had no impact on formalization. Similarly, Galiani, Melendez and Navajas (2015) find that the elimination of the initial fixed cost of registration in Colombia was not effective in increasing formalization. One exception is a study by Campos, Goldstein and McKenzie (2015) who find that, in a setting where registration had no tax implications for the firm, firms in Malawi registered in response to a costless business registration intervention.

In line with the results from the literature, data from the World Bank suggest that initial formalization costs may not be the primary barrier to formalizing. For instance in the World Bank's enterprise surveys, the obstacle most frequently reported by firms is access to financing (La Porta, Rafael, and Andrei Shleifer, 2014). Figure 1 shows that, for both formal and informal firms, access to financing is the most common obstacle to their operations, and is more so for informal than for formal firms. About 44 percent of informal firms cite this as their main obstacle relative to about 19 percent of formal firms. In contrast, only about two percent and six percent of formal and informal firms, respectively, cite business licensing and permits as their main obstacle. It is also notable that, for formal firms, tax rates are one of the main obstacles to doing business. Despite these observations, relative to the initial entry cost of formalization, much less attention appears to have

been given to policies that affect the the ongoing benefits of formalization such as access to credit or the ongoing costs of formalization associated with taxes and regulatory compliance.

In their randomized evaluation in Sri Lanka, de Mel, McKenzie and Woodruff (2013) gave some firms a one-time payment if they registered while some firms were given information about registration processes and offered a reimbursement of the registration costs. As discussed above, providing information and reimbursing registration costs had no effect on formalization. The one-time payment, however, increased formalization. Although these results suggest that the initial entry costs of formalization may not be the main barrier to formalizing, given that the payment was a one-time event it is not clear that these results should be taken as evidence that the ongoing costs or benefits of formalization are determinants of formalizing. In contrast, in our paper, we will be able to directly analyze the impact of a simultaneous change in the ongoing costs and benefits of formalization induced by a tax exemption and capital subsidy scheme in India. We describe this scheme in the next section. Furhermore, we determine whether the change in the ongoing costs or the change in the ongoing benefits of formalization is a more effective incentive for formalization.

Most closely related to our work are Fajnzylber, Maloney, and Montes-Rojas (2011) and Rocha, Ulyssea, and Rachter (2016). The authors analyze programs in Brazil, which reduced and simplified taxes for small firms, and find an increase in formalization. Our paper differs from Fajnzylber, Maloney, and Montes-Rojas (2011) and Rocha, Ulyssea, and Rachter (2016) in two ways. First, we study a simultaneous change in both the ongoing costs and benefits of being a formal firm, whereas these two papers study a change in only the ongoing costs. Furthermore, our context allows us to determine whether reductions in the ongoing costs or increases in the ongoing benefits of being a formal firm are more effective at increasing formalization. Second, the intended aim of the tax exemption and capital subsidy policy we study was to incentivize large formal firms to enter and existing firms to grow in the two states. Therefore, our paper differs from these two papers in that the policy we study led to the registration of informal firms that were not directly targeted by the policy, as was the case for the Brazilian programs which specifically targeted small firms. Thus, we show that a policy that affects the costs and benefits of firms in the formal sector

without necessarily targeting small, informal firms can still incentivize these firms to register.

3 The New Industrial Policy

In June 2003, the central government of India, in order to attract industrial investments and generate employment in the states of Uttarakhand and Himachal Pradesh, introduced an incentive package referred to as the New Industrial Policy for Uttarakhand and Himachal Pradesh that reduced taxes and provided capital subsidies. Both Uttarakhand and Himachal Pradesh are two of the smaller states in India and are located in the north of the country. Both states are very hilly and forests constitute a large proportion of their area. Due to their topography, industrialization was limited in these two states. In 2000, the two states together accounted for less than one percent of the number of factories and industrial output in India. The incentive package was, therefore, introduced to encourage industrialization (Chaurey, forthcoming).

The incentive package included the following.

1. 100 percent excise duty exemption for a period of 10 years from the date of commencement of commercial production
2. 100 percent income tax exemption for an initial period of five years and thereafter between 25 percent and 30 percent for a further period of five years
3. A capital investment subsidy equal to 15 percent of a firm's investment in plant and machinery, subject to a ceiling of 3 million rupees (approximately USD 50,000)

Firms eligible for the package included the following.

1. A new industrial unit in designated areas
2. An existing industrial unit in designated areas if it underwent substantial expansion (increase

of at least 25 percent in the value of its fixed capital investment in plant and machinery for the purpose of expansion of capacity or modernization and diversification)

3. All units in thrust sectors[2] irrespective of their location

To claim the incentives, firms had to be registered. The designated areas included (i) Existing Industrial Estates, (ii) Proposed Industrial Estates, (iii) Industrial Activity in Non-industrial Area and (iv) Expansion of Existing Industrial Estates. These designated areas were expanded such that the policy was applicable in essentially any area where industrial activity was possible in the two states. This, therefore, obliterated the differential treatment given to the thrust sectors. Thus, in effect, the scheme affected the entire states of Uttarakhand and Himachal Pradesh.

The tax exemptions were applied to the taxes collected by the central government. In general, the central government taxes firms in India on their worldwide income arising from all sources at corporate income tax rates between 30 percent for domestic corporations and 40 percent for foreign corporations. Central excise duty rates range from 8 percent to 16 percent. The tax exemptions, therefore, implied a substantial reduction in the tax burden of firms.

Firms registered by March 31, 2010 and March 31, 2012 were eligible for the excise duty and income tax exemptions, respectively. Firms registered by January 7, 2013 were to be eligible for the capital subsidy but this deadline was later extended to March 31, 2017.

4 Data

The data used in our analysis are from the fourth round (1998), preceding the introduction of the tax and capital policy, and fifth round (2005), following the introduction of the policy, of the Economic Census of India. This census is a complete enumeration of all firms in India, including both formal and informal firms. Since only industrial firms, that is, non-agricultural firms, were

[2] The thrust sectors were sectors that were supposed to get preferential treatment irrespective of firm location.

eligible for the policy we analyze, we restrict our sample to these firms. The census provides information on whether or not the firm is registered, the gender of the firm's owner, the number of employees and the source of finance. Unfortunately, the census does not provide information on firm output and inputs (other than labor) and so our analysis is restricted to the variables described above.

The 1948 Factories Act in India requires firms to register if they have 10 or more employees and use electricity in their operations or if they have 20 or more employees and do not use electricity. Firms with numbers of employees below these thresholds are also allowed, though not required, to register. Thus, most firms with 10 or more employees are required by law to be registered (see Amirapu and Gechter (2016)) and were, therefore, likely to be already registered when the scheme came into place. Therefore, the firms most likely to be affected would be smaller firms that were not required by law to be registered. We, therefore, restrict our sample to firms with fewer than 10 workers.

Table 1 reports summary statistics for the variables in 1998, prior to the introduction of the scheme, broken down by registration status. Informal firms tend to be smaller and are more likely to be rural, have an owner from a historically disadvantaged group and be self-financed, that is, have no access to external financing[3].

To assess whether changes in registration are driven by pre-existing informal firms rather than new firms, we use data from the fifty-sixth round (2000), preceding the introduction of the policy, and sixty-second round (2006), following the introduction of the policy, of India's National Sample Survey Organisation's survey of the unorganized manufacturing sector.[4] This survey includes

[3] The survey asks firms about their source of finance, with the possible answers being the government, formal financial institutions such as a bank, non-institutional sources such as a money lender, other sources such as non-governmental organizations, and self financing.

[4] The unorganized manufacturing sector refers to all manufacturing enterprises, both formal and informal, that are not covered by India's Annual Survey of Industries. The Annual Survey of Industries covers firms that have 10 or more employees and use electricity in their operations or

information on whether the firm has been in operation for fewer than three years which allows us to analyze differential changes in the entry of new firms following the policy.

5 Empirical Strategy and Results

5.1 Empirical Strategy

We employ a difference-in-differences setup to analyze the impact of the tax exemption and capital subsidy scheme on registration by firms. We exploit the fact that this scheme was introduced in only two states, Uttarakhand and Himachal Pradesh. We compare changes in the registration of firms following the introduction of the scheme in these two states, which we, henceforth, refer to as the treated areas, to changes in states that were not eligible for the scheme, which we, henceforth, refer to as the control areas.

The main estimating equation is as follows:

$$y_{ijdt} = \beta_0 + \beta_1 post_t * treat_d + \alpha_j + \delta_d + \lambda_t + \epsilon_{ijdt} \tag{1}$$

y_{ijdt} is an outcome for firm i in industry j in district d in year t. $post_t$ takes on a value of one in 2003, when the scheme was introduced, and afterwards and zero otherwise. $treat_d$ takes on a value of one if the firm is located in a treated district and zero otherwise. α_j, δ_d and λ_t are industry, district and year fixed effects, respectively. The main outcome variable of interest is the registration status of a firm. Standard errors are clustered at the district level to allow for correlations across firms in the same district. The coefficient of interest is β_1, which is an estimate of the average differential change in the registration status of firms in the treated areas relative to that of firms in the control areas in the same industry following the introduction of the scheme .

Our aim is to compare outcomes in the treated areas to outcomes in areas that are similar to the treated areas such that any differential change in outcomes can be attributed to the scheme. We define these control areas as all neighboring states (states that share a border with the treated states,

have 20 or more employees and do not use electricity.

Uttarakhand and Himachal Pradesh). These neighboring states are Chandigarh, Punjab, Delhi, Haryana and Uttar Pradesh. Figure 2 is a map of India showing the treated and control states. The control states, like the treated states, are all located in North India. North India and South India are generally considered to be economically and socially distinct from each other. Therefore, we consider the neighboring states as valid control areas since they are more similar economically and socially to the treated states than states in South India and are, therefore, likely to have the same trends in industrial outcomes in the absence of any differential schemes. Note that, as discussed in Section 3, the control states were chosen because of low levels of industrial activity and are, therefore, likely to differ from the treated states in terms of the *levels* of industrial outcomes. However, the assumption underlying our strategy is that the control and treated states would have exhibited similar *trends* in outcomes in the absence of the scheme. While it is impossible to observe this counterfactual, below we turn to the usual indirect test of this assumption, that prior to the scheme the control and treated states had similar trends.

We check whether there were similar pre-trends by making use of the 1990 economic census in addition to the 1998 census. Unfortunately, information on the registration status of firms is not available in the 1990 economic census so we are unable to check for the existence of differential pre-trends in this variable. We check for the existence of differential pre-trends in the other available variables which include the number of workers, the location of the firm and whether the owner is from a historically disadvantaged group. Table 2 reports results from the following regression for these outcomes using the data for 1990 and 1998:

$$y_{ijdt} = \gamma_0 + \gamma_1 D1998_t * treat_d + \alpha_j + \delta_d + \lambda_t + \epsilon_{ijdt} \tag{2}$$

where $D1998_t$ takes on a value of one in 1998 and zero otherwise, and the other variables are as described above. γ_1 is an estimate of the differential change in the outcome variable between 1990 and 1998 in treated areas relative to control areas. The results in Table 2 indicate that, for each of the outcome variables, this estimate is small in magnitude and is statistically insignificant, indicating that there were no differential trends between 1990 and 1998, prior to the introduction of

the scheme. Since we find no evidence of the existence of differential pre-trends in any of these variables, it is unlikely that there were differential pre-trends in registration status as well.

Another potential concern with our difference-in-differences methodology is that, if there were other policy changes that differentially applied to the treated areas, we would be unable to attribute any changes in outcomes to the tax exemption and capital subsidy scheme. To our knowledge, no other policies were implemented in the period we study that differentially applied to the treated states. Policies such as the Electricity Act of 2003 and the Special Economic Zones Act of 2005 were national policies that affected all of India and, thus, should not have had a differential effect on the treated states.

5.2 Effect on Registration

We now analyze the impact of the scheme on the registration of firms using the difference-in-differences strategy. The results from estimating Equation 1 are presented in Table 3. Column 1 reports the results for the sample of industrial firms. The dependent variable is a dummy variable which takes on a value of one if the firm is registered and zero otherwise. The coefficient on the interaction between $post_t$and $treat_d$ is positive and statistically significant, indicating that the scheme led to an increase in the registration of firms in the treated states. The magnitude of the coefficient implies that the the scheme resulted in a substantial increase of about 27 percentage points in the probability of a firm registering.

5.3 Heterogeneous Effects

We explore heterogeneity in the impact on registration with respect to various firm characteristics in a difference-in-difference-in-differences framework. We, therefore, modify Equation 1 by including an interaction between $post_t$, $treat_d$ and a specific firm characteristic.

We first examine heterogeneity by enterprise type. In India, manufacturing enterprises are classified into three groups: Directory Manufacturing Establishment (DME), Non-Directory Manufacturing Establishment (NDME) and Own Account Manufacturing Enterprise (OAME). A DME is a manufacturing enterprise employing six or more workers, at least one of whom is a hired

worker and not a household worker. An NDME is a manufacturing enterprise employing fewer than six workers, at least one of whom is a hired worker. An OAME is a manufacturing enterprise employing no hired workers. DMEs are, therefore, the most similar to formal firms, which tend be larger relative to informal firms and to have hired workers. One might, therefore, expect that DMEs may be the most likely type of firms to register relative to NDMEs and OAMEs since they are closest to the formality threshold. We consider this hypothesis in Column 1 of Table 4. The coefficient on the triple interaction term between $post_t$, $treat_d$ and a dummy variable for DME firms is positive and statistically significant, indicating that such firms were most responsive to the scheme. This result is similar to the result in Bruhn (2013), which studies a Mexican reform that simplified business registration. The author finds that informal firms with owners with characteristics similar to those of owners of formal firms were more likely to register following the reform compared to informal firms with owners with characteristics similar to those of wage workers. Similarly, we find that the scheme in India had a larger effect on DMEs, which are more likely to resemble formal firms, compared to NDMEs and OAMEs, which are more likely to resemble wage workers.

In Column 2 of Table 4, we also find that rural firms were less likely to register than urban firms.

We next explore heterogeneity by the characteristics of the firm owner, specifically gender and whether the owner belongs to a historically disadvantaged group. The coefficient on the triple interaction term between $post_t$, $treat_d$ and a dummy variable for female-owned firms in Column 3 of Table 4 is negative and statistically significant, implying that female-owned firms were less likely than male-owned firms to register in response to the scheme. The estimate in Column 4 of Table 4 also indicates that firms with owners from historically disadvantaged groups were less likely to register. Finally, in Column 5 of Table 4, we assess heterogeneity by access to finance. We find that firms that were self-financed, relative to firms with access to external sources of finance, were less likely to register.

Because the characteristics examined above are likely to be correlated, in Column 6 of Table 4 we report estimates from a regression that includes interactions between $post_t$, $treat_d$ and each of

the characteristics. While the coefficient on the interaction with the dummy variable for DME becomes insignificant, the coefficients on the other interactions remain significant. To summarize, we find that rural firms, female-owned firms, firms with owners from historically disadvantaged groups and firms without access to external financing were less likely to register in response to the scheme.

A potential explanation for the result that firms with access to external financing were more likely to register is as follows. As long as the net benefit of being formal is positive, it is in the interest of a firm to borrow to pay the initial registration fees to become formal in order to have access to the benefits of being formal. Thus, firms that are not credit-constrained, that is, have access to external sources of financing in the absence of internal funds, are in a position to pay the initial registration fees and access the benefits of being formal. They would, therefore, be more likely to respond to a scheme that increases formal sector benefits, which is consistent with our result.

Note that, even controlling for access to finance, female-owned firms, firms with owners from historically disadvantaged groups and rural firms were less responsive to the scheme. This finding suggests that these subsets of firms may face other barriers to formalization. One possibility is that female-owned firms and firms with owners from historically disadvantaged groups may face discrimination or other barriers that prevent them from fully realizing the benefits provided by the policy. Hence, such firms may not register because their expected benefits from formalization may be lower. For instance, Iyer, Khanna, and Varshney (2013) find that people from historically disadvantaged groups are significantly underrepresented in the ownership of enterprises and tend to be smaller and more likely to belong to the informal sector. Further research is needed to understand exactly what the formalization barriers faced by these firms may be, which is, unfortunately, not feasible with the data we use in this paper.

5.4 New Firms or Pre-existing Informal Firms?

Our results could be driven by the creation of new firms rather than the registration of

pre-existing informal firms. Table 5 shows that the share of informal firms fell following the introduction of the scheme, driven by a relatively larger decline in the number of informal firms compared to formal firms. However, this result could be consistent with a scenario in which informal firms exit the market altogether rather than register and become formal. We show that this is not the case. The increase in registration is primarily driven by the registration of pre-existing informal firms.

To show this, we look at the age profile of the firms in the treated and control areas before and after the introduction of the scheme in 2003 using data from India's National Sample Survey Organisation's survey of the unorganized manufacturing sector described in Section 4. Table 6 reports results from a regression where the dependent variable is a dummy variable equal to one if the firm has been in operation for fewer than three years. The coefficient on the interaction between $post_t$ and $treat_d$ is small and statistically insignificant, implying that there was no differential increase in the number of young firms following the introduction of the scheme. This result suggests that the increase in registration is not primarily driven by the creation of new firms but rather by the registration of pre-existing informal firms.

5.5 Tax Exemption versus Capital Subsidy

The subsidy provided by the policy was a capital subsidy equal to 15 percent of a firm's investment in plant and machinery. Therefore, this subsidy would have been more salient for firms with high levels of capital investment. To determine if the increase in registration caused by the policy primarily occurred because of the tax exemption or the capital subsidy, we check if there was a differential effect for firms in capital-intensive industries. Such firms would have benefited from both the tax exemption and the capital subsidy whereas less capital-intensive firms would have benefited mainly from the tax exemption. Table 7 presents results from a regression that includes a triple interaction term between $post_t$, $treat_d$ and a dummy variable for firms in capital-intensive industries. We calculate capital intensity using data from India's National Sample Survey Organisation's survey of the unorganized manufacturing sector. For each 3-digit industry, we

calculate capital intensity as the ratio of the industry's capital, defined as the total value of buildings, plant and machinery to its labor, defined as the total number of employees. We define capital-intensive industries as those with capital intensities above the median capital intensity.[5] The coefficient on the triple interaction term in Table 7 is small and statistically insignificant indicating that capital-intensive firms were no more responsive to the scheme than less capital-intensive firms. This finding suggests that the observed impact on registration occurred primarily as a result of the tax exemption rather than the capital subsidy.

5.6 Effect on Other Outcomes

Next, we look at the impact of the policy on outcomes other than registration status. First, we assess the effect of the policy on employment by firms in Column 1 of Table 8. We find no statistically significant difference between firms in treated states relative to firms in control states before and after the policy. While interpreting this result, it must be kept in mind that we are assessing the short-run (two years after the policy was initiated) effects of the policy on firms. In Columns 2 and 3, we look at the proportion of firms that had access to formal finance and the proportion of firms that self-financed their operations. Interestingly, even in the short run, the policy led to an increase in the proportion of firms that received formal finance and a decrease in the proportion of firms that financed their operations using their own resources. Columns 2 and 3 provide suggestive evidence that even in the short run, the policy led to higher access to finance, and might have been one of the incentives for firms to register.

5.7 Robustness Checks

We now assess the robustness of our estimates. Agricultural firms were not eligible for the tax exemption and capital subsidy scheme and, as such, there should be no differential change in the

[5] Note that we have fewer observations in Table 7 relative to the baseline results because some industries do not have capital information available in the National Sample Survey Organisation's dataset.

registration status of these firms in the treated states. In Column 2 of Table 3, we report results from regressions using data on agricultural firms. Reassuringly, there is no evidence of a change in the registration status of these firms in the treated areas relative to the control areas.

A potential concern is that formal firms in the control areas may have relocated to the treated areas to take advantage of the scheme. If this is the case, then our results may simply be reflecting this movement of pre-existing formal firms rather than the registration of informal firms. Since such relocation is more feasible across bordering districts, we can check the robustness of our results by restricting our sample to interior districts, that is, districts that do not lie along the borders between the treated and control states. Column 2 of Table 9 reports the results from this regression while Column 1 reports the baseline regression for comparison. The estimate in Column 2 is essentially the same as the baseline estimate, demonstrating that our results are not driven by the relocation of firms. Furthermore, if the results were driven by the relocation of firms along the borders between treated and control states, we should see a differential decline in the number of firms located along the borders relative to the interior of the control states. We analyze this by running a regression of the log of the number of firms in a district on an interaction between $post_t$ and a dummy variable equal to one if the district shares a border with a treated state. The coefficient on this interaction term in Column 3 is small and statistically insignificant, confirming that our results are not driven by relocation.

Also, to check that our results are not driven by any particular control state, we check the robustness of our results to excluding one control state at a time. These results are reported in Table 10, with the baseline estimate reported in Column 1 of the table for comparison. The estimates remain virtually unchanged across the different samples.

Another concern is that, if existing formal firms expanded as a result of the reform, there may have been some spillovers to informal firms that allowed them to become more profitable and register. In this case, the increase in the registration of informal firms would be due to spillovers from formal firms rather than a direct effect of the tax exemption and capital subsidy on informal firms. We hypothesize that, if this was indeed the case, the increase in registration would have been

higher among informal firms that work directly with formal firms. To test this, we examine if there was a differential change in registration for firms in industries that tend to have linkages between informal and formal firms.

The National Sample Survey Organisation's survey includes the following variables: whether a firm is a supplier to another firm and whether a firm undertakes any work on a contract basis with another firm. We use these two variables as proxies for linkages between informal and formal firms. We define "supplier-intensive industries" and 'contract-intensive industries" as those with the fractions of firms acting as suppliers to other firms or working on a contract basis, respectively, above the medians. Table 11 presents results from regressions that include triple interaction terms between each of these dummy variables, $post_t$ and $treat_d$.[6] The coefficients on the triple interaction terms are statistically insignificant, indicating that industries with linkages between formal and informal firms were not more likely to experience an increase in registration. These results, therefore, suggest that the increase in registration was unlikely to have been driven by spillovers from formal firms.

Finally, since the treatment is at the state level, it may be ideal to cluster the standard errors at the state level instead of at the district level as has been done in the analysis. However, clustering assumes that the number of clusters is sufficiently large, which is not the case here as we have only seven states. The recent econometrics literature has proposed the wild cluster bootstrap-t procedure as an alternative with less bias when there are few clusters (Cameron, Gelbach, and Miller, 2008). However, this procedure performs well when there are at least ten clusters. Unfortunately, there is currently no ideal approach offered in the literature on how to deal with clustered standard errors when there are very few clusters, as is the case in this paper. Nevertheless, as a robustness check, we use the wild cluster bootstrap-t procedure to calculate standard errors. The results from this procedure are shown in Appendix Table A1. The increase in registration in the treated states is still

[6] Note that we have fewer observations in Table 11 relative to the baseline results because some industries do not have information on suppliers and contracts available in the National Sample Survey Organisation's dataset.

statistically significant using this procedure.

6 Conclusion

Despite the prevalence of informal firms in developing countries and attempts by policymakers to encourage these firms to register, relatively little is known about the incentives to which these firms are most likely to respond. The little literature that exists on creating incentives for informal firms to formalize has mainly focused on incentives related to reducing the initial costs rather than the ongoing costs and benefits of registration. In this paper, we attempt to fill this void in the literature by looking at the responsiveness of registration to changes in the ongoing costs and benefits of being a formal firm. Using a difference-in-differences framework, we are able to show that registration increased in response to a scheme in India that reduced taxes and provided capital subsidies to formal firms.

This impact was, however, heterogeneous, with the increase in registration concentrated among firms that were located in urban areas, had access to external financing, had male owners, and had owners who were not from historically disadvantaged groups. Furthermore, we find suggestive evidence that the increase in registration was driven primary by the reduction in the ongoing costs rather than by the increase in the ongoing benefits of being a formal firm.

While the analysis in this paper advances our knowledge about the responsiveness of firm registration to formal sector incentives, we find that certain subsets of firms are not as responsive to these incentives. Further work is, therefore, needed to determine the mechanisms behind this heterogeneity.

References

[1] Amirapu, Amrit, and Michael Gechter. 2016. "Indian Labor Regulations and the Cost of Corruption: Evidence from the Firm Size Distribution." Unpublished manuscript.

[2] Bruhn, Miriam. 2011. "License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico. *Review of Economics and Statistics*, 93(1): 382-386.

[3] _____. 2013. "A Tale of Two Species: Revisiting the Effect of Registration Reform on Informal Business Owners in Mexico." *Journal of Development Economics*, 103(2013): 275-283.

[4] Cameron, A., J. Gelbach, and D. Miller. 2008. "Bootstrap-based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90(3): 414-427.

[5] Campos, Francisco, Markus Goldstein, and David McKenzie. 2015. "Short-Term Impacts of Formalization Assistance and a Bank Information Session on Business Registration and Access to Finance in Malawi." World Bank Policy Research Working Paper No. 7183.

[6] Chaurey, Ritam. Forthcoming. "Location-Based Tax Incentives: Evidence from India." *Journal of Public Economics.*

[7] de Andrade, Gustavo Henrique, Miriam Bruhn, and David McKenzie. 2016. "A Helping Hand or the Long Arm of the Law? Experimental Evidence on What Governments Can Do to Formalize Firms." *World Bank Economic Review*, 30(1): 24-54.

[8] de Mel, Suresh, David McKenzie, and Christopher Woodruff. 2013. "The Demand for, and Consequences of, Formalization Among Informal Firms in Sri Lanka." *American Economic Journal: Applied Economics*, 5(2): 122-50.

[9] de Soto, H. 1989. *The Other Path: The Economic Answer to Terrorism*. New York: HarperCollins.

[10] ___. 2000. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. New York: Basic Books.

[11] Fajnzylber, Pablo, William F. Maloney, and Gabriel V. Montes-Rojas. 2011 "Does Formality Improve Micro-Firm Performance? Evidence from the Brazilian SIMPLES program." *Journal of Development Economics*, 94(2): 262-276.

[12] Galiani, Sebastian, Marcela Meléndez, and Camila Navajas. 2015. "On the Effect of the Costs of Operating Formally: New Experimental Evidence." NBER Working Paper No. 21292.

[13] Harris, John R., and Michael P. Todaro. "Migration, Unemployment and Development: A Two-Sector Analysis." *American Economic Review*, 60(1): 126-142.

[14] Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4): 1403-1448.

[15] Iyer, Lakshmi, Tarun Khanna, and Ashutosh Varshney. 2013. "Caste and Entrepreneurship in India." *Economic and Political Weekly*, 48(6): 52-60.

[16] Kaplan, David S., Eduardo Piedra, and Enrique Seira. 2011. "Entry Regulation and Business Start-Ups: Evidence from Mexico." *Journal of Public Economics*, 95(11): 1501-1515.

[17] La Porta, Rafael, and Andrei Shleifer. 2014. "Informality and Development." *Journal of Economic Perspectives*, 28(3): 109–126.

[18] Maloney, W.F.. 2004. "Informality Revisited." *World Development*, 32(7): 1159-1178.

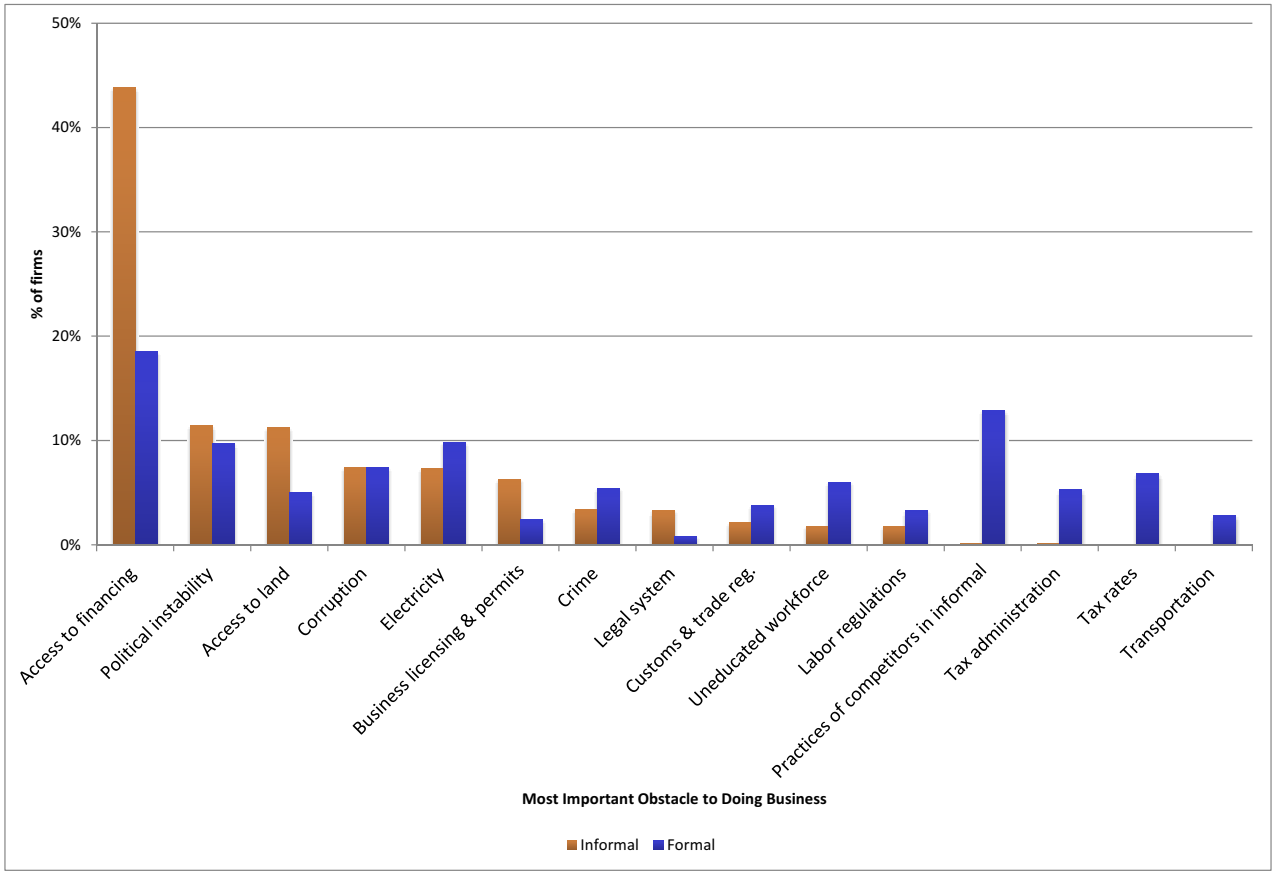
[19] Moser, C. 1978. "Informal Sector or Petty Commodity Production: Dualism or Dependence in Urban Development?" *World Development*, 6(9/10): 1041-1064.

[20] Rocha, Rudi, Gabriel Ulyssea, and Laísa Rachter. 2016. "Do Lower Taxes Reduce Informality? Evidence from Brazil." Mimeo. Pontifícia Universidade Católica do Rio de Janeiro.

[21] Sharma, Rajiv, and Sunita Chitkara. 2006. “Informal Sector in the Indian System of National Accounts.” Expert Group on Informal Sector Statistics (Delhi Group), Paper No. 06.

[22] World Bank. 2016. *Doing Business 2016 : Measuring Regulatory Quality and Efficiency*. Washington, DC: World Bank.

Figure 1: Obstacles to Doing Business



Notes: Data are from La Porta, Rafael, and Andrei Shleifer, 2014.

Figure 2: Treated and Control States

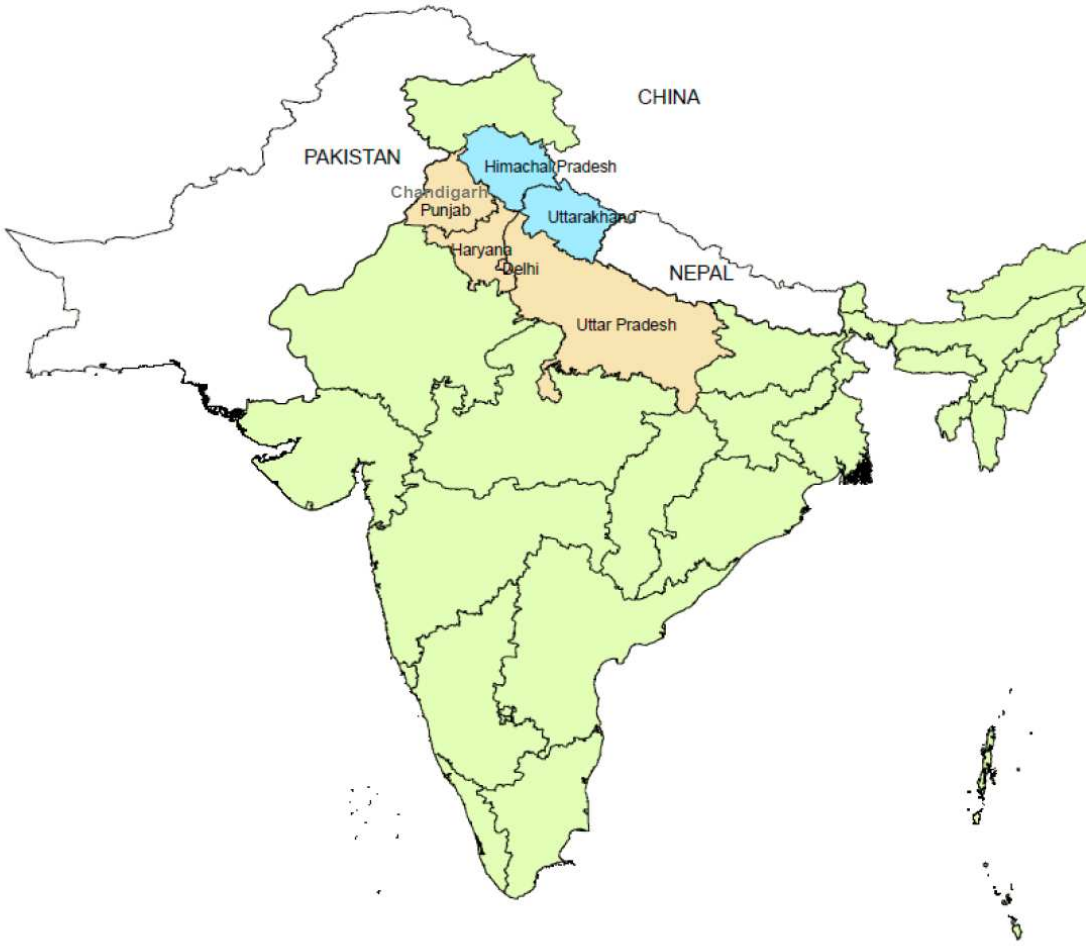


Table 1: Summary Statistics

	(1)	(2)	(3)	(4)
	Informal Firms	Formal Firms	All Firms	Informal - Formal
Number of workers	1.83	3.58	1.85	-1.755***
	(1.35)	(2.49)	(1.38)	
Number of hired workers	0.52	2.04	0.54	-1.515***
	(1.19)	(2.51)	(1.23)	
Rural	0.39	0.27	0.39	0.124***
	(0.49)	(0.44)	(0.49)	
Female owner	0.03	0.03	0.03	-0.00283***
	(0.16)	(0.17)	(0.16)	
Disadvantaged owner	0.41	0.28	0.41	0.135***
	(0.49)	(0.45)	(0.49)	
Self-financed	0.91	0.76	0.91	0.157***
	(0.28)	(0.43)	(0.28)	
No. of Firms	4,528, 318	55,925	4,584,243	

Notes: Standard deviations are in parentheses. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 2: Pre-trends

	(1)	(2)	(3)	(4)	(5)
	at least six workers with one hired	rural	no. of workers	no. of hired workers	historically disadvantaged owner
post1990*treat	0.000327	0.0147	-0.0216	0.00342	-0.0121
	(0.00148)	(0.00980)	(0.0176)	(0.0183)	(0.00800)
No. of Observations	8,670,916	8,670,916	8,670,916	8,670,916	8,670,916

Notes: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Robust standard errors, in parentheses, are clustered at the district level. All regressions include district, 3-digit industry and year fixed effects.

Table 3: Effect on Registration

	(1)	(2)
	registered	registered
post*treat	0.269***	-0.00132
	(0.0337)	(0.0162)
No. of Observations	10,932,520	610,663
Sample	industrial firms	agricultural firms

Notes: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Robust standard errors, in parentheses, are clustered at the district level. All regressions include district, 3-digit industry and year fixed effects.

Table 4: Heterogeneous Effects

	(1)	(2)	(3)	(4)	(5)	(6)
	registered	registered	registered	registered	registered	registered
post*treat	0.271***	0.366***	0.273***	0.295***	0.319***	0.433***
	(0.0338)	(0.0492)	(0.0338)	(0.0284)	(0.0370)	(0.0433)
post*treat*DME	0.0464**					-0.00636
	(0.0220)					(0.0205)
post*treat*rural		-0.115***				-0.119***
		(0.0387)				(0.0370)
post*treat*female			-0.101***			-0.0853***
			(0.0192)			(0.0147)
post*treat*disadvantaged				-0.114***		-0.103***
				(0.0315)		(0.0304)
post*treat*self-financed					-0.0698**	-0.0568**
					(0.0274)	(0.0264)
No. of Observations	10,932,520	10,932,520	10,932,520	10,932,520	10,932,520	10,932,520

Notes: A Directory Manufacturing Establishment (DME) is a manufacturing enterprise employing six or more workers, at least one of whom is a hired worker and not a household worker. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Robust standard errors, in parentheses, are clustered at the district level. All regressions include district, 3-digit industry and year fixed effects.

Table 5: Effect on Number of Firms

	(1)	(2)	(3)
	share of informal firms	log(informal firms)	log(all firms)
post*treat	-0.261***	-0.391***	-0.111**
	(0.0258)	(0.0743)	(0.0496)
No. of Observations	37,772	36,988	37,772

Notes: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Robust standard errors, in parentheses, are clustered at the district level. All regressions include district, 3-digit industry and year fixed effects.

Table 6: Effect on Firm Entry

	(1)
	young firm
post*treat	0.00642
	(0.0442)
No. of Observations	59,595

Notes: A young firm is one that has been in operation for fewer than three years. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Robust standard errors, in parentheses, are clustered at the district level. All regressions include district, 3-digit industry and year fixed effects.

Table 7: Heterogeneity by Capital Intensity

	(1)
	registered
post*treat	0.118***
	(0.0256)
post*treat*capital-intensive	-0.0179
	(0.0367)
No. of Observations	2,383,675

*** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Robust standard errors, in parentheses, are clustered at the district level. All regressions include district, 3-digit industry and year fixed effects.

Table 8: Effect on Employment and Financing

	(1)	(2)	(3)
	log(no. of workers)	access to formal finance	self-financed
post*treat	0.0288	0.0345***	-0.0322**
	(0.0184)	(0.0112)	(0.0152)
No. of Observations	10,932,520	10,932,520	10,932,520

Notes: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Robust standard errors, in parentheses, are clustered at the district level. All regressions include district, 3-digit industry and year fixed effects.

Table 9: Robustness to Relocation of Firms

	(1)	(2)	(3)
	registered	registered	log(all firms)
post*treat	0.269***	0.270***	
	(0.0337)	(0.0475)	
post*bordering			0.0203
			(0.0622)
No. of Observations	10,932,520	9,157,922	32,379
Sample	all districts in treated and control states	interior districts in treated and control states	all districts in control states

Notes: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Robust standard errors, in parentheses, are clustered at the district level. All regressions include district, 3-digit industry and year fixed effects.

Table 10: Robustness to Sample of Control States

	(1)	(2)	(3)	(4)	(5)	(6)
	registered	registered	registered	registered	registered	registered
post*treat	0.269***	0.270***	0.276***	0.271***	0.270***	0.251***
	(0.0337)	(0.0337)	(0.0337)	(0.0341)	(0.0341)	(0.0335)
No. of Observations	10,932,520	10,837,567	9,590,828	9,779,314	9,409,165	4,918,767
Sample excludes:	none	Chandigarh	Delhi	Haryana	Punjab	U. Pradesh

Notes: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Robust standard errors, in parentheses, are clustered at the district level. All regressions include district, 3-digit industry and year fixed effects.

Table 11: Heterogeneity by Linkages with Formal Firms

	(1)	(2)
	registered	registered
post*treat	0.104***	0.0978***
	(0.0282)	(0.0274)
post*treat*supplier-intensive	0.00165	
	(0.0383)	
post*treat*contract-intensive		0.0594
		(0.0472)
No. of Observations	2,383,635	2,383,635

*** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Robust standard errors, in parentheses, are clustered at the district level. All regressions include district, 3-digit industry and year fixed effects.

A Appendix

Table A.1: Effect on Registration (Wild Cluster Bootstrap)

	(1)	(2)
	registered	registered
post*treat	0.269**	-0.00132
	[0.017]	[0.913]
No. of Observations	10,932,520	610,663
Sample	industrial firms	agricultural firms

Notes: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. P-values, in square brackets, are based on standard errors clustered at the state level using the wild cluster bootstrap-t procedure. All regressions include district, 3-digit industry and year fixed effects.