Location-Based Tax Incentives: Evidence From India

Ritam Chaurey*

Department of Economics, State University of New York, Binghamton

August 24, 2014

Abstract

While policies targeting particular geographic regions are widely used by governments, there have been few rigorous evaluations of their causal impacts. In this paper, I study the impact of a location-based tax incentive scheme in India. Using aggregated and firm-level panel data, I find large increases in employment, total output, fixed capital, and the number of firms as a result of the program. These increases are due to both the growth of existing firms as well as the entry of new firms. There is supporting evidence that the new firms entering the treated regions are larger and more productive. I find no evidence for relocation of firms or spillovers in industrial activity between treatment and control areas. Finally, using data from household surveys, I show that wages of workers rise but find no changes in housing rents or migration across the treated and control regions. My results therefore suggest that the policy increased welfare, and I also conclude that the policy was cost-effective. This provides support for "place-based" policies to correct for regional economic disparities, especially in settings with low labor mobility.

*Department of Economics, SUNY Binghamton, P.O. Box 6000, Binghamton, NY 13902-6000. E-mail: *rchaurey@binghamton.edu*. I am grateful to Eric Verhoogen, Amit Khandelwal, and Suresh Naidu for advice and guidance, and to Ama Baafra Abeberese, S Anukriti, Prabhat Barnwal, Christopher Boone, Tirthankar Chakravarty, Giacomo De Giorgi, Rajeev Dehejia, Jonathan Dingel, Rana Hasan, Jonas Hjort, Elise Huillery, Supreet Kaur, Siddharth Kothari, Tao Li, Cristian Pop-Eleches, Raul Sanchez de la Sierra, Miguel Urquiola, and participants at the Development Colloquium, Columbia University for helpful comments. All errors are my own.

I Introduction

Many countries in the world have massive economic disparities across regions. To reduce these regional inequalities, state and local governments often use "location-based" policies that seek to generate employment and productivity in particular regions.[1] These policies include tax exemptions, subsidies, land grants, and other infrastructural benefits to firms in order to incentivize them to locate to disadvantaged regions.[2] The benefits and distortions caused by these policies have long been debated by economists.[3] Whether such spatially targeted policies are able to generate economic gains in a cost-effective manner is largely an empirical question.

The empirical evaluation of location-based policies is complicated because they have both direct effects (on employment and output) and indirect effects (on local prices). Detailed micro data on firms, workers, migration, and local prices is needed in order to quantify their overall effects, and such data is often not available. This may partly explain the lack of empirical work assessing spatially targeted policies, especially in developing countries. I fill this gap in the literature by studying a place-based policy in India. Specifically, I examine the federally financed New Industrial Policy for the states of Uttarakhand and Himachal Pradesh. This policy provided tax exemptions and capital subsidies for new and existing firms starting in 2003, with the primary aim of inducing industrialization and generating employment in the two states.

The causal effect of the 2003 policy is identified using difference-in-differences (DID) and synthetic control methods [Abadie, Diamond, and Hainmueller (2010)]. To estimate the treatment effect on industrial outcomes such as employment, output, wage bill, and fixed capital, I use an aggregated state-industry level dataset and a firm-level panel dataset. This allows me to look at the entry of new firms and the growth of existing firms. I use several different comparison groups, with varying levels of stringency, to ensure robustness in the identification strategy. I show that the treated and control states follow similar trends prior to the policy change, which supports the credibility of the difference-in-differences estimator. First, I use all major states of India as the control group. To ensure that some unobserved

[1] I use the terms location-based policies and place-based policies interchangeably.

[2] Some examples include Empowerment Zones and Enterprise Zones in the United States, Zones Franches Urbaines (ZFU) in France and Regional Selective Assistance in the United Kingdom.

[3] Theoretical arguments can be found in Glaeser (2001), Glaeser and Gottlieb (2008), Moretti (2011), and Kline and Moretti (2013b).

shock to one region of the country is not driving the results, I then compare the treated states only to neighboring states. The strictest empirical specifications compare only those firms that are located in districts lying on the border between treated and control states. These districts are quite similar in terms of observable characteristics and this specification provides a strict check on the other specifications. Finally, synthetic control methods employ a weighted average of all available control states as a counterfactual for the treated states. All results are robust to these different control groups.

The main results are as follows. At the aggregate state-industry level, I find large increases in employment (37 percent), number of factories (27 percent), total output (57 percent), fixed capital (72 percent), and industrial wage bill (39 percent) in treatment states relative to control states. These results provide an estimate for the combined effect of entry of new firms and growth of existing firms as a result of the policy change. The differential impact of the policy on existing firms in the treated areas compared to the control areas is given by the firm-level results. These results show that on average, existing firms in the treated areas increased employment (7.5 - 11 percent), output (8.7 - 18 percent), and made additions to plant and machinery (25 - 28 percent) as compared to firms in control states.

Although the DID regressions show a differential effect of the policy on various outcomes, one concern is that the effects maybe driven by spillovers from the treated to the control regions. If the policy simply causes economic activity to relocate[4] from the control to the treated areas, then the estimated treatment effect might overstate the aggregate effect of the policy. Alternately, there could be positive spillovers due to agglomeration economies on the nearby control areas that might lead us to underestimate the effect of the policy change.

I test for these channels explicitly by taking the following steps. First, I run difference-in-differences specifications comparing the neighboring states and districts to regions located even further away from the actual treated group (Uttarakhand and Himachal Pradesh) based on the assumption that spillovers are likely to be stronger in regions located closer to the treated states. I find no differential outcomes in neighboring states and bordering districts compared to those located further away.[5] Second, I test for differential firm closures across

[4]Place-based policies have often been criticized for simply relocating economic activity across different locations without actually increasing aggregate output [Kline and Moretti (2013a), Glaeser and Gottlieb (2008), Mayer, Mayneris, and Py (2012)].

[5]This is in contrast to Givord, Rathelot, and Sillard (2013), who find evidence for negative spillovers from treated to neighboring control areas in French ZFUs.

treated and control states and find no statistically significant difference. Finally, I also rule out that the overall results are being driven by multi-establishment firms reallocating production across plants to take advantage of the tax exemptions.[6]

One economic justification for providing tax incentives to attract new firms is the possibility of agglomeration economies on the existing firms in the locality [Greenstone, Hornbeck, and Moretti (2010), Kline and Moretti (2013a), Glaeser and Gottlieb (2008)]. If more productive firms enter a location, there might be positive spillovers on existing plants, leading to overall growth.[7] I find that the policy change attracted larger and more productive plants to enter the treated areas. However, I do not find any differential effect on TFP (total factor productivity) for existing plants. One possible explanation for the lack of productivity spillovers could be that the existing firms take time to internalize the agglomeration economies generated by new plants. This paper however, only studies the short-term firm-level responses on productivity and hence might be unable to find evidence for such spillovers.

Having estimated the reduced form effects of the policy on industrial outcomes, I then analyze the impact on the local population. This is often complicated because place-based policies have general equilibrium effects. For instance, if workers are mobile, tax incentives for firms to locate to a particular region might be ineffective in raising real wages of residents. The increase in labor demand by firms and the consequent rise in nominal wages for residents might be partially or even completely offset by increases in housing rents and costs of living as new workers move into the area [Roback (1982), Moretti (2011)]. Hence, in order to get an estimate of the effects of the policy change on the local population, I combine data on wages and expenditures from household surveys with data on rents and migration. I find that nominal wages and monthly per capita expenditure differentially increase but housing rents remain unchanged in treated states after the policy change. Since the policy change might have affected other local prices instead of housing rents, I adjust the nominal wages and per capita expenditure measures using a state-level price deflator. I document that real wages and real expenditure per capita also increase (12 percent and 10 percent respectively) differentially in the treated areas. Theoretically, this is consistent with low mobility

[6]Removing multi-establishment plants from the sample does not change the magnitudes of the treatment effects.

[7]For example, when the elasticity of agglomeration with respect to economic density in the receiving region is higher, reallocating economic activity from one region to another leads to a long run increase in output [Glaeser and Gottlieb (2008)].

of workers across regions as the increase in nominal wages following a labor demand shock is not offset by increases in the cost of living due to entry of new workers. I explicitly test for differential migration and find no statistically significant difference between the treated and control states. The results on low migration in India are also consistent with previous literature.[8]

Finally, I assess the costs and benefits of this location-based tax incentive scheme. Since there is no official government estimate on the effectiveness of the policy, I conduct a back-of-the-envelope quantitative assessment using the estimates from the paper. I find that the policy resulted in approximately 30,000 jobs, between 500-600 new firms, and around 7.8 billion Rupees in industrial wages in 2007-08 (the last year in the data set). I conclude that the policy was cost-effective as the gains in profits for firms and the total wage bill for workers in the treated states outweigh my estimates of the costs (including both actual costs of subsidies and foregone tax revenues).[9] My estimates suggest a gain of 0.5 percent of the GDP (in 2007-08) in the two states relative to what would have happened in the absence of it.

Prior empirical work evaluating place-based policies has primarily focused on developed countries, mostly in the United States and Europe. In the United States, the focus has been on Federal Empowerment Zones (EZ) and State Enterprise Zones (ENTZ); these are neighborhoods receiving tax breaks and job subsidies. The results on the efficacy of these zones in creating jobs have been mixed.[10] Other recent papers on place-based policies have studied programs in European countries. These include the "Regional Selective Assistance" in the United Kingdom [Criscuolo, Martin, Overman, and van Reenen (2012)], the French ZFUs [Mayer, Mayneris, and Py (2012), Givord, Rathelot, and Sillard (2013)] and Italy's Law 488/1992 [Bronzini and de Blasio (2006)]. My paper contributes to this growing literature by rigorously evaluating the incidence and welfare impacts of a location-based policy in a developing country. Wang (2013) is the only other paper in a developing country and

[8]See Munshi and Rosenzweig (2009), Topalova (2010), and Hnatkovska and Lahiri (2013).

[9]The deadweight loss in this setting where workers are not very mobile will be low and most of the benefits will accrue to local workers. [Busso, Gregory, and Kline (2013)].

[10]Neumark and Kolko (2010), Greenbaum and Engberg (2004), and Bondonio and Greenbaum (2007) find no effects of enterprise zones on employment growth. However, Ham, Swenson, Imrohoroglu, and Song (2011) find positive effects for EZs, ENTZs, and Federal Enterprise Communities. Busso, Gregory, and Kline (2013) also find that the EZ program increased employment and wages inside the zones at moderate efficiency costs.

studies the impact of Chinese SEZs on municipality level outcomes such as foreign direct investment, agglomeration and local prices.[11] I extend this relatively new literature studying location-based policies in developing countries by conducting an overall assessment of the policy change using a comprehensive set of outcomes (both detailed industrial- and household-level). I also carry out a cost-benefit analysis of the policy using my estimates and add to the small recent literature on local labor markets that studies the overall costs and benefits of place-based policies [Busso, Gregory, and Kline (2013), Kline and Moretti (2013a)].

This paper also contributes to the literature on firms' location decisions in response to tax differentials. Duranton, Gobillon, and Overman (2011) find a negative impact of local taxes on firm employment but no impact on firm entry in the United Kingdom. Rathelot and Sillard (2008) look at French micro data and find a weak response of firms' location decisions to higher taxes. In contrast, I document a large increase in the entry of firms as a result of the tax exemptions.

The results of this paper, thus, inform policy-makers about the efficacy of tax benefits for industrializing backward regions. In this context, I observe large responses of firms to tax benefits, but very little migration response by individuals. This suggests that it might be easier to provide incentives for firms to move to a particular location than to move people. Especially in settings with low labor mobility, such spatially targeted policies could be a cost-effective way to generate employment, output, and real earnings gains for workers.[12]

The rest of the paper is organized as follows. Section II presents the background for the study and the details of the policy. Section III discusses the empirical strategy, Section IV describes the data, and the results are discussed in Section V. Finally, Section VI conducts a cost-benefit analysis and Section VII concludes.

[11] The analysis in Wang (2013) is restricted as detailed firm-level data is not available to study the dynamics of new and old firms. Furthermore, the paper cannot directly test for migration and only alludes to low worker mobility because nominal wages rise faster than cost of living. A quantitative cost-benefit analysis is also not possible due to data constraints.

[12] Kline and Moretti (2013b) also note that "..in the idealized model..the most efficient demand side subsidy was one that yielded no mobility response at all and simply raised local wages."

II Background and Policy Details

Himachal Pradesh and Uttarakhand are two states in the north of India (see Figure 1). In November 2000, the northwestern districts of Uttar Pradesh were split off to form the state of Uttarakhand. After the formation of Uttarakhand, it was placed in the list of "special category" states[13] that included Jammu and Kashmir, Himachal Pradesh, Arunachal Pradesh, Assam, Sikkim, Manipur, Meghalaya, Mizoram, Nagaland, and Tripura.

Both Uttarakhand and Himachal Pradesh are two of the smaller states in India, together covering roughly 3.5% of India's total area. They are predominantly covered by hilly areas and forests.[14] According to the 2001 Census, the total population of Himachal Pradesh and Uttarakhand was around 6.1 million and 8.5 million respectively (around 1.4% of India's total population). Industrialization was considered a policy challenge in the two states, owing to the topography. For instance, in 2000, the two states together accounted for less than 1% of the number of factories and industrial output in India. Beginning 2003, the Government of India (central government), in order to attract industrial investments and generate employment in the states of Uttarakhand and Himachal Pradesh, decided to provide the following incentive package.

I. New industrial units and existing industrial units on their substantial expansion (increase by at least 25% in the value of fixed capital investment in plant and machinery of an industrial unit for the purpose of expansion of capacity/modernization and diversification) set up in 'designated' industrial estates/growth centers were entitled to:

(a) 100% excise duty exemption for a period of 10 years from the date of commencement of commercial production

(b) 100% income tax exemption for an initial period of five years and thereafter 30% for companies and 25% for others for a further period of five years

(c) all new firms and existing units (upon substantial expansion) in the notified locations would be eligible for capital investment subsidy equaling 15% of their investment in plant and machinery, subject to a ceiling of Rs. 3 million (approximately USD 50,000).

[13] These areas have hilly and difficult terrain, very low level of infrastructural development and significant tribal population. Almost all of them are border states with considerable international borders. These states get preferential treatment in federal assistance.

[14] According to India State of Forest Report, 2011, forest area covered 66.5% of the area of Himachal Pradesh and 64.8% of the area in Uttarakhand.

II. A list of 'thrust sector' industries was compiled that would be eligible for the benefits listed above irrespective of whether they located in an industrial estate or not.

These tax exemptions pertained to the taxes collected by the central government. In general, companies resident in India are taxed on their worldwide income arising from all sources at corporate income tax rates between 30% (for domestic corporations) and 40% (for foreign corporations). Central excise duty rates varied between 8%-16%. These tax exemptions were large enough to incentivize firms to enter the states.

Importantly, a few months after the policy (in June 2003) was initiated, the Government of India issued a notification[15] designating the areas in the two states where industrial units would be eligible to get these tax incentives. The notification included (i) Existing Industrial Estates (ii) Proposed Industrial Estates (iii) Industrial Activity in Non-industrial Area and (iv) Expansion of Existing Industrial Estates. This notification made almost all of the existing industrial activity prior to 2003 and surrounding areas eligible for the benefits. This in practice, blurred the differential treatment accorded to the thrust sector industries, because there was virtually no area (where industrial activity was possible) in the two states where the policy would not be applicable.

The central excise tax exemption was removed on 31st March 2010, and the income tax exemption was removed on the 31st of March 2012. Essentially, any new industrial units set up or existing units undertaking substantial expansion in these states prior to the above dates would continue to be eligible for these benefits.

III Empirical Strategy

In this paper, I empirically test whether the centrally sponsored location-specific tax incentives led to differential increases in industrial outcomes in the treated areas as compared to control areas. The empirical strategy uses the 2003 policy change that provided tax incentives to firms in the two states of Uttarakhand and Himachal Pradesh in a difference-in-differences setup. I use this state-year variation to compare outcomes before and after the policy change (2003) in the treated areas to a set of control units. To the best of my knowledge, no other policy was implemented in these two states beginning 2003 that affected

[15] See Notification No. 50/2003 - Central Excise, Dated: June 10, 2003, available at http://himachal.nic.in/industry/welcomelat.htm

industrial outcomes, and this helps me to identify the treatment effect of the particular tax incentive scheme.

Ideally we would like to compare the treated states to an observationally similar control group. I consider a few different control groups for the analysis. I compare industrial outcomes in the treated states to a set of neighboring states and then to all major states taken together. The most stringent specification compares outcomes of firms in districts located on either side of the borders in the treated and control areas. This is a strict test on the identification, as districts on either side of the border tend to be more similar as compared to geographically distant locations. Finally, I also perform robustness checks using the synthetic control method where the control group is formed using a weighted average of all the non-treated states in order to best match the treated states. For all outcome variables, the synthetic control group comprises states both near and away from the treated states.

Indian firm-level data sets do not provide exact location identifiers (to the level of street address and zip codes) below the district level. Since each district in the two treated states had at least one designated area that was eligible for the incentives, an empirical strategy comparing firms or industries across districts within the treated states would not be possible.[16] Following visits to the respective state industry departments, it became clear that the notification brought almost all existing industrial activity within the ambit of the policy change and also added new areas. Therefore, unlike Mayer, Mayneris, and Py (2012), the closing down of existing firms in ineligible areas to re-open in an eligible industrial area *within the state* is not a concern here.

It is thus reasonable to consider this policy as affecting the entire states of Himachal Pradesh and Uttarakhand. In this paper, it will not be possible to separate out the effects of the tax incentives from the capital subsidy provided.

Before looking at regression specifications, Figures 2 and 3 plot the raw data over time for the variables of interest at the state-industry level. These plots show that the pre-2003 trends in employment, number of factories, total output, and fixed capital were similar across the treated and control states, only diverging after 2003. The pre-treatment trends look parallel and provide visual support to the use of difference-in-differences (DID) strategy in this

[16] I use the terms industrial estate and designated area interchangeably because many non-industrial areas with existing industrial activity prior to 2003 were included as eligible areas for the policy through Notification 50/2003.

context to estimate the causal effect of the policy change.[17]

I run two main types of regressions (DID specifications) to estimate the treatment effect of the policy change on industrial outcomes. First, I run *state×3-digit industry level* regressions of the form:

$$y_{sjt} = \delta_s + \lambda_{jt} + \beta(post_t \times treat_s) + \gamma(X_{st}) + \varepsilon_{sjt} \tag{1}$$

where s,j,t indexes state, 3-digit industry, and time respectively, y_{sjt} represents an outcome variable such as employment, number of factories, total output, fixed capital, or industrial wage bill that varies at the state, industry, and year level, δ_s represents state fixed effects, λ_{jt} represents industry-year fixed effects and X_{st} represents time varying controls.[18] The coefficient β, on the interaction term $post_t \times treat_s$, where

$$post_t = \begin{cases} 1 & \text{if year is 2003 or after} \\ 0 & \text{if year is pre 2003} \end{cases}$$

$$treat_s = \begin{cases} 1 & \text{if state is Uttarakhand or Himachal Pradesh} \\ 0 & \text{otherwise (control states),} \end{cases}$$

is then the causal effect of the policy change. I include pre-treatment state-level variables from the 2001 Census such as population, number of industrial and agricultural workers etc. in the state and interact them with a time dummy for each year as control variables.[19] The regressions with employment, output, wage bill, and fixed capital therefore combine both the extensive (entry and exit of firms) and intensive margin (growth by existing firms) of the policy change. The regression with number of factories as the regressand gives us the extensive margin directly and is a cumulative effect that takes into account both entry and exit of firms.

The next set of regressions are at the *firm-level*:

$$y_{idjt} = \alpha_i + \lambda_{jt} + \beta(post_t \times treat_d) + \gamma(X_{idjt}) + \varepsilon_{idjt} \tag{2}$$

[17] The trends look similar even on adding all states. To be visually clear, these graphs only show the trends in the nearby states.

[18] Note that in all state-industry regressions, the control group is either neighboring states or all major states. District identifiers are not available in this dataset.

[19] These regressions are similar to the area level employment and number of plants regressions in Criscuolo, Martin, Overman, and van Reenen (2012).

where i,d,j,t indexes firm, state or district (depending on choice of control group),[20] 4-digit industry and time respectively, and y_{idjt} represents a firm-level outcome variable such as employment, output, fixed capital, additions to fixed capital, additions to plant and machinery, or wage bill. I also use age and age-squared as controls in the regressions.[21] Note that the inclusion of firm fixed effects removes the effect of new firms entering after 2003. Hence, this regression looks at the impact of the policy change on the outcome measures only for incumbent firms and can be interpreted as the intensive margin of the policy change.

While the DID regressions estimate the differential effect of the policy between the treatment and control areas, it is conceivable that the results are being affected by changes caused by the policy in the control areas. For example, relocation of firms from the control states to the treated states might lead us to wrongfully attribute the observed effects as being caused by the policy change. To check whether firms close down in control states to reopen in the treated states, I take three approaches. First, I look at trends in the number of operational factories in the treated states and the neighboring control states. Then, I run a regression at the state-industry level with the number of closed firms[22] as the dependent variable to look at the differential impact on firm closures across the treated and control states, before and after the policy change.[23] Finally, I run regressions comparing the impact of the policy change in neighboring states to states further away from the treated states. The underlying assumption is that firms closer to the treatment states would be more likely to relocate production into those states in response to the policy. If there is substantial relocation, we would expect to see lower industrial activity in neighboring control states relative to states that are further away from the treated areas. Rather than closing down an existing plant in a control area and reopening in the treated states, a multi-establishment firm might move production between its various plants to take advantage of the tax benefits. To rule this out, I also run regressions that omit multi-establishment firms.

A related concern might be that the policy induces spillovers in the nearby control areas.

[20]For the firm-level data, district identifiers are available and thus the control groups include neighboring states or bordering districts.

[21]Since a particular firm does not change location in the dataset, the firm fixed effect subsumes the state or district fixed effects.

[22]I define the number of closed firms as the difference between the total number of firms and the number of operational firms.

[23]If firms are relocating to the treated states, we would expect to see a larger number of firm closures in the control states as compared to the treated states.

Positive spillovers in industrial activity from the treated states to the neighboring control states would lead us to underestimate the effect of the policy change. Such externalities might be substantial, especially in control districts bordering the treated districts, and may lead to a differential response on firms in districts nearer to the treated states relative to those further away. To check whether the firm-level results are being influenced by spillovers, I run a regression specification comparing the firms along the border in the treated states to those in districts further away from the border in the control states (essentially omitting the bordering control districts from the regression). I also run a specification to see the effect of the policy change on firms in bordering control districts compared to firms in districts further away in the control states.

Finally, to look at the effects of the policy on wages, rents, and migration, I run regressions of the form:

$$y_{kdst} = \delta_d + \lambda_t + \beta(post_t \times treat_s) + \gamma(X_{kdst}) + \varepsilon_{kdst} \quad (3)$$

where k,d,s,t indexes household or individual, district, state and time respectively. y_{kdst} represents wages or migration status in the individual-level regressions and rents or monthly per capita expenditure in the household-level regressions. For the individual-level regressions, I control for age, sex, marital status, education status, and the industry of work. The regressions with housing rents use attributes of the house as controls, such as roof type, dwelling type, floor, number of rooms, and area.

IV Data

I combine data from multiple sources to evaluate the impact of the policy change. For industrial outcomes, I use two datasets: (i) the Annual Survey of Industries (ASI) state×3-digit industry panel (from 1999-2000 to 2007-08) and (ii) ASI firm-level panel (1999-2000 to 2007-08). To study the effects on individual and household outcomes, I use (a) Employment-Unemployment rounds of the National Sample Survey (NSS) for the years 1999-2000, 2004-05, 2005-06 and 2007-08 and (b) Housing Conditions rounds of the NSS for the years 2002 and 2008.

The Annual Survey of Industries (ASI), conducted by the Ministry of Statistics and Program Implementation (MoSPI), is the main source of industrial statistics in India. The

ASI covers the entire Factory Sector comprising industrial units (called factories) registered under the Sections 2(m)(i) and 2(m)(ii) of the Factories Act, 1948. This includes all firms employing 10 or more workers using power and 20 or more workers without the use of power. Geographically, it covers the entire country except the states of Arunachal Pradesh, Mizoram, and Sikkim, and Union Territory of Lakshadweep for the surveys. The ASI dataset is well-suited to answer this question as it covers formal sector firms that are affected by tax changes.

For the state-industry level regressions, I use the ASI state×3-digit industry panel. Each observation is at the state-industry-year level. Industries are classified at the 3-digit National Industrial Classification (NIC) codes. This data set includes 65 industries (3-digit NIC), 9 years (1999-2000 to 2007-08) and all major states. Table 1 shows descriptive statistics for the variables of interest at the state×3-digit industry level. As Table 1 shows, the two treated states of Himachal Pradesh and Uttarakhand had smaller industrial employment, number of factories, total output, fixed capital, and wage bill as compared to neighboring states or the rest of India before 2003. For example, average employment size in a 3-digit industry before 2003 in the treated states was 590 as compared to the figures for neighboring states (3154) or all the major states together (4952). Post-2003, the average size of industrial employment goes up throughout India, but the increase is highest in the treated states. Similar increases can be seen for number of factories, total output, and fixed capital after 2003 in the treated states as compared to other states. Mean total output and fixed capital at the state industry level rises almost three-folds in the treated states after 2003, much larger than the increase in the other states.

For the firm-level regressions, I use the ASI firm-level panel for the years 1999-2000 to 2007-08. The ASI frame is divided into census (surveyed every year) and sample (sampled every few years) sectors. However, the definition of the two sectors has changed from time to time. Five industrially backward states[24] are always covered in the census sector. For the rest of India, the definition of the census sector has changed from 200 or more employees (1998-2000) to 100 or more employees (2001 onwards). To take into account the changes in the sampling frame, I run firm-level regressions using the sampling weights provided by ASI. I restrict the sample to the major states and union territories of India as covered by

[24]Manipur, Meghalaya, Nagaland, Tripura and Andaman and Nicobar Islands.

the ASI.[25]

"Firm" in this context means a factory, the unit of observation in the data set. Table 2 shows summary statistics for the different outcome variables at the firm-level broken up by treated states, neighboring states, and major states for periods before and after the policy change. I use the sampling weights from the data set to construct the summary statistics for the estimated population.[26] Average employment within the firm increases post 2003, irrespective of which group we look at. Median employment after 2003 however, increases by almost 56% for firms in the treated state whereas the increase is negligible for firms in the rest of the country. Mean output and fixed capital almost double for firms in the treated group after 2003. This increase is much larger as compared to any other group.

To study the welfare effects of the policy change, I use migration, wages, and house rents data from the National Sample Survey (NSS). NSS is a nationally representative household survey in India, also conducted by the MoSPI. Specifically, I use rounds 55 (1999-2000), 61 (2004-05), 62 (2005-06), and 64 (2007-08) of the employment-unemployment surveys of the NSS for the wages data. The survey provides information on wages and employment for each household member over the last seven days before the interview. To study migration, I use NSS Rounds 55 (1999-2000) and Round 64 (2007-08)[27] - one round each before and after the policy change. The survey elicits information about the last usual place of residence for the household members. I define an external (internal) migrant as one whose last usual place of residence was another state or country (same state but another district). To look at the effect of the policy on rents, I use two NSS rounds of the Housing Conditions schedule for the years 2002 (round 58) and 2008 (round 65). These rounds include questions on housing rents and the attributes of the house such as total floor area, kitchen type, floor type, number of rooms, type of roof, and type of dwelling. Finally, I construct a state-level GDP deflator using the state GDP at constant and current prices from the Reserve Bank of India (RBI) Handbook of Statistics on the Indian Economy to deflate nominal values.[28] Summary statistics for these NSS data are shown in Table 3.

[25] I do not include Jammu & Kashmir or the states in the North-east namely Assam, Manipur, Meghalaya, Nagaland, and Tripura.

[26] See Harrison, Martin, and Nataraj (2011) and Bollard and Sharma (2013).

[27] Only these two rounds have information on migration for the relevant time frame.

[28] See Appendix Table A2.

V Results

I begin by reporting the results for the difference-in-differences regressions at the state-industry level for different outcome variables (Subsection V.1) and then look at the firm-level results (Subsection V.2). The synthetic control results are reported in subsection V.3. Subsection V.4 discusses results on productivity and results on wages, rents, and migration are discussed in subsection V.5.

V.1 State-industry results

The state-industry DID regression results are reported in tables 4 through 8. For each of these tables, Panel A uses neighboring states as the control group and includes Haryana, Punjab, Delhi, Chandigarh, and Uttar Pradesh. The states of Haryana, Punjab and Uttar Pradesh border the treated states, whereas Chandigarh and Delhi are the commercial hubs near the treated states. In Panel B, all the major states of India are included as the control group. This includes the neighboring states and the states of Rajasthan, Bihar, Andhra Pradesh, Chhattisgarh, Maharashtra, Madhya Pradesh, Orissa, Goa, Kerala, Karnataka, Tamil Nadu, Jharkhand, Gujarat, and West Bengal. All the results in this subsection can be interpreted as the cumulative effect of the growth of existing firms and the entry of new firms at the state-industry level.

For tables 4 to 8, Column 1 includes state, year, and 3-digit industry fixed effects. State fixed effects control for time invariant state characteristics like the area and topography of the state. The year fixed effects control for macroeconomic shocks affecting all states and the industry fixed effects control for time invariant industry characteristics. Column 2 is a more flexible specification as it includes industry-year fixed effects which control for time varying industry characteristics. This is important because some industries like pharmaceuticals and IT (information technology) have grown in India over the last decade, and the industry-year fixed effects controls for these changes.[29] Column 3 adds time varying controls at the state level to the specification in Column 1. I include pre-2003 state-level variables from the 2001 Census such as population, agricultural, and industrial workers etc. and interact them with

[29]A related concern might be that the industrial composition in the treated states is different than the control states and to control for this, I run regressions controlling for state×industry fixed effects. These results are shown in Table B6 in the Appendix.

a time dummy for each year as control variables. Column 4 includes the specification of Column 2 with time varying controls.[30] Standard errors are clustered at the state-year level. For each regression, I also report Cameron, Gelbach, and Miller (2011) multi-way clustered standard errors at the state and year level. Since the policy change only affected two states and number of state clusters are small, inference using standard cluster-robust techniques may lead to over rejection. As a robustness check, I report the wild cluster bootstrap-t p-values [Cameron, Gelbach, and Miller (2008)] for the main regressions in Appendix Table B7.[31] The results remain significant on using the bootstrapped p-values.

The dependent variable in Table 4 is the log of employment at the 3-digit industry level. In columns 1 to 4, the coefficient of interest on the interaction *post*treat* is positive and significant at the 1% level. Mean employment at the industry level differentially increases in the treated states relative to the control states by around 37% - 42% (Panel A). In Panel B, I run the same specifications with all major states as the comparison group. The coefficient of interest in Panel B is between 43% and 45%. The magnitudes of the treatment effect are similar and do not depend on the choice of the control group. This translates into approximately 30,000 additional jobs in the treated states in 2007-08 (the last year in the dataset) compared to what would have happened in the absence of the policy.

Table 5 looks at the same specifications for log of total number of factories as the dependent variable. In this table, the coefficient on *post*treat* can be interpreted as the extensive margin of the policy change as it takes into account new entry by firms as a result of the policy change. Columns 1 through 4, Panel A, show that the effect of the policy change on the average number of factories in an industry (in treated states relative to control states) is between 27% and 29%. The corresponding estimates from Panel B show a 31% increase in the average number of factories. Table 5 confirms that the policy change led to a large differential increase in the number of new firms coming in to the treated states relative to control states. These estimates translate into a total additional increase of around 550 firms in the treated states in 2007-08.

Table 6 reports the results for log of total output at the state-industry level. The effect of the policy change on total output (in treated states relative to control states) is even larger than the effect on total employment, ranging between 58% and 64%. Results for log of fixed

[30] I also run regressions controlling for yearly state-level GDP. All results are robust to adding this control variable.

[31] Also see Busso, Gregory, and Kline (2013).

capital are shown in Table 7. The results show an increase of 87% in fixed capital and part of this can be attributed to the "substantial expansion" clause where existing firms needed to increase their investment of fixed capital by at least 25% to receive the tax exemptions. Furthermore, capital investment subsidies were also provided in these two states after 2003 to both new and existing firms, contributing to the massive increase in fixed capital at the industry level. Finally in Table 8, I show that the industrial wage bill differentially increases by 40% in the treated states as compared to the control states.

The estimated coefficients over time along with standard errors at the 95% confidence level are plotted in Figure 4. These coefficients are obtained from a regression of the outcome variable (log employment, log factories, log output, and log fixed capital) on the interaction between *treat* (indicator variable for treated states) and time dummies after controlling for state, year, and industry fixed effects. These graphs visually show that before 2003, there were no trends in the outcomes and the effects only show up after 2003. These graphs also provide visual evidence for the treatment effect of the policy change and show that the effect increases every year after 2003.

There may be some concern that firms close down in the neighboring control states to reopen in the treated states to take advantage of the tax incentives. To check for this, I take the following steps. First, I plot the number of operational factories in the treated states and the neighboring control states. If the policy change in 2003 caused factories to close down in the neighboring control states and reopen in the treated states, there should be a decline in the number of operational factories in the neighboring states. Figure 5 plots the trends in operational factories and there is no evidence that more factories closed down in control states compared to treated states. To check for differential closure of firms across treated and control states, I run difference-in-difference specifications with the number of firm closures as the dependent variable in Table 9. I find no differential response in terms of firm closures across the treated and control states. In Table 10, I run regressions similar to placebo checks. I remove the treated states from the sample and run regressions assuming that the neighboring states got treated by the policy change. The underlying assumption is that the relocation of factories, workers, and capital is easier from nearby places as compared to regions further away from the treated states. Hence, the policy change should have differentially impacted neighboring states as opposed to states further away from the treated states. The results in Table 10 compare outcomes in the neighboring states to other major states before and after the policy change. There is no statistically significant systematic difference in the outcomes between the neighbors and all other states.[32] Overall, there is no evidence of differential closure of firms or relocation of industrial activity across the treated and control states.

However, firms in control states might re-direct capital and produce output across their different plants to take advantage of these tax incentives without closing down. To rule out this channel completely, I would need to check whether multi-establishment firms are driving the results. I discuss this issue along with the firm-level results next.

V.2 Firm-level results

Firm-level regressions are reported in Table 11. I restrict the sample to open firms. Different rows show the results for the various outcomes of interest. All columns, 1 through 4, include firm, year, and 4-digit industry fixed effects. Columns 2 and 4 also control for 4-digit industry-year fixed effects. Firm fixed effects control for any time invariant unobserved heterogeneity at the firm-level and 4-digit industry-year fixed effects take into account time varying effects across industries. I also control for age and age-squared in all regressions. Columns 1 and 2 use firms in the neighboring states as the control group. In columns 3 and 4, I restrict the sample to bordering districts.[33] Districts on the border along the treated and control states tend to be observationally similar, differing only because of differential benefits provided to firms. In these specifications, I compare outcomes for firms across bordering districts (in treated and neighboring control states) before and after the policy change. These regressions are a strict test on the identification strategy and provide credible support to my results from using firms in neighboring states as the control group. As mentioned earlier, these regressions only show the effect of the policy change on existing firms. This is because the effect of new firms entering after 2003 is removed by the inclusion of firm fixed effects.

In Table 11, the coefficients on the interaction term *post*treat* can be interpreted as the intensive margin of the policy change as it shows the effect of the policy change on incumbent firms (firms present both before and after the policy change). Columns 1 and 2 show that the mean employment for existing firms in the treated states differentially increases by 7-10% as

[32]Employment is marginally significant at the 10% level but has the opposite sign to what we would expect.
[33]List of bordering districts is shown in Appendix Table A1.

compared to those in control states. The results in columns 3 and 4 with firms in bordering districts as the control group also shows a differential increase in mean employment in firms by around 11%. Total output and wage bill also differentially increase for existing firms in the treated states compared to those in control states. The differential increase in total output is between 8.7% and 23.7% depending on the choice of the control group. Wage bill also increases by 8% to 13%.

I also run regressions with fixed capital, additions to fixed capital, and additions to plant and machinery as outcome variables. Fixed capital includes depreciation whereas additions to fixed capital and additions to plant and machinery are measures of actual additions before depreciation. The measure of stock of fixed capital is more likely to suffer from measurement error than the numbers for actual additions[34] made during the year. This is also clear from the regression results. For example, fixed capital shows an increase of around 7-8% but in most cases is statistically indistinguishable from zero. Actual additions to fixed capital however, increased by around 28% and additions to plant and machinery increased by around 26% for existing firms in treated areas compared to control areas. The coefficients from the bordering districts regressions also have similar magnitudes. These results provide suggestive evidence that existing firms took advantage of the "substantial expansion" clause and increased investment on fixed capital to receive the tax benefits. In this respect, these results confirm that the policy was successful in incentivizing firms to invest more in plant and machinery.

It is conceivable that the firm-level results are being driven by multi-establishment firms reallocating production across their various plants to take advantage of the incentives. In Table 12, I directly test for this by removing multi-establishment firms from the sample. I find similar coefficients using this sample and hence it is unlikely that the firm-level results are being driven by these establishments. A separate concern in the bordering districts regressions is the possibility that results are downward biased because of positive spillovers from treated to control areas. To check for this, I run a regression specification with the firms along the border in the treated states compared with firms further away from the bordering districts in the control states. These results are shown in Table 13. For existing firms in the treated districts as compared to control districts, employment went up by 15% and output by

[34]Actual additions are similar to measures of investment.

16% after 2003.[35] In Table 14, I compare firms in districts in the neighboring control states that border the treated states to firms in districts further away from the treated states. I find no differential effects in firm-level outcomes in this regression. The results from Tables 13 and 14 taken together, suggest that spillovers do not play a substantial role in the firm-level results and especially lends support for the bordering districts regression specification.

To explore differences between new and existing firms, I plot kernel density graphs (figures 6 through 9) comparing firms across treated and control regions. The graphs clearly indicate that the new firms entering the treated states are larger and more productive than both the existing firms (in the treated states) and the new firms entering the neighboring control states. Since informal firms tend to be smaller in size than formal firms, this is suggestive evidence that the effects of the policy change are not completely being driven by informal firms becoming formal. However, it is not possible to say where the new firms come from and where they might have set up in the absence of the policy.

V.3 Synthetic Control Methods

The New Industrial Policy for the states of Uttarakhand and Himachal Pradesh, provides an ideal setup to apply the synthetic control method [Abadie and Gardeazabal (2003), Abadie, Diamond, and Hainmueller (2010)]. I use the method to find a "synthetic control" group for the treated states, by using a weighted average of the available "donor pool" (units that are unaffected by the policy), that best matches the value of the predictors of the outcomes of interest (directly employed, number of factories, total output, and fixed capital) before 2003. This synthetic control group then approximates the trajectory of the outcomes for the treated units in the counterfactual event that the policy had not been in place. Since the choice of the control units is data driven, the method is extremely transparent. Following this, I conduct placebo tests where the synthetic control method is applied to all the control units in the sample, with the treated unit in the donor pool. The inference is then exact, in the sense that I look at the ratio of the post-intervention (2003) to pre-intervention mean squared prediction error (MSPE) from the treated units' outcome and compare it to the same ratio for all the placebo runs. This is similar to a permutation test and essentially examines whether the treatment effect is much larger in the treated states as compared to a randomly

[35] These magnitudes are a bit larger than those obtained in the firm-level regressions in Table 11, suggesting the possibility of a downward bias in the treatment effects.

chosen placebo. If the effect is of a similar magnitude in the treated units as compared to a randomly-chosen placebo, then it is not possible to conclude that the treatment effect is valid.

First, I aggregate the data from the two treated states into one treated unit,[36] and use the total aggregated numbers over all industries in the state in a year from the ASI state-industry level panel data. Since I do not include Jammu & Kashmir or the states in the north-east of India in the 'donor pool', there are 23 available control states and union territories to choose from.

The outcome variables of interest are the number of employed persons, number of factories, total output, and fixed capital, for all industries together in a state in a given year. As predictors of the outcome variables at the state level, I include number of males, females, literacy rate, number of workers, number of main and marginal workers, cultivators, agricultural laborers, and percentage of Scheduled Caste and Scheduled Tribes. These variables are available from the 2001 Census. I also include the lagged values of the outcome variables in the pre-intervention period (2000-03).

Figure 10, Panel A graphically shows the treatment effect of the 2003 policy change on the number employed in all industries in the states. The solid line shows the trajectory for the treated states, which is closely matched by the dashed line (synthetic control) in the pre-intervention phase, only to diverge substantially after 2003. Table B1a in the appendix shows the weights on the states in the donor pool, that provide a good synthetic control group. Only Chandigarh, Chhattisgarh, Dadra and Nagar Haveli, Jharkhand, and Rajasthan have positive weights, with the rest of the states getting zero weights. Both Chandigarh and Dadra and Nagar Haveli are union territories and the results are unchanged on removing the union territories from the donor pool. Using logs of the outcome variables also keeps the results unchanged. Table B1b shows the values for various predictors of the outcome variable for both the treated and the synthetic control groups. Figure 10, Panel B is used to conduct the inference test. It visually shows the ratio of post-intervention to pre-intervention MSPE for the treated states and all the placebo runs for states in the donor pool. Clearly, this ratio is the largest for the treated states. Hence, if the treatment was randomly assigned to any unit in the sample, the probability of obtaining a ratio as large as in the treated states

[36]The synthetic method works well when there is one treated unit and a number of control units to choose from.

would be $1/24 = 0.042$. This implies that the differential effect in employment between the treated and control states is significant at the 5% level.[37] Also note that the total treatment effect in 2007-08 is around 25,000 jobs and is close to the estimates from the regression in Table 4.

Figure 11, Panel A shows the synthetic control method graph for the outcome variable - number of factories in all industries at the state level. Table B2a shows that positive weights have been applied to Andaman & Nicobar Islands, Chandigarh, Chhattisgarh, Jharkhand, and Rajasthan. Results remain unchanged on removing the union territories (Andaman & Nicobar Islands and Chandigarh). Table B2b shows the mean of the predictors used for number of factories. I conduct a similar inference test in Figure 11, Panel B and plot the ratio of post-intervention to pre-intervention MSPE. Again, the ratio is largest for the treated states and hence the probability of obtaining the largest ratio is $1/24 = 0.042$. As in the previous case, the total treatment effect in 2007-08 (around 600 firms) is close to the estimates from the regressions in Table 5.

Finally, Figures 12 and 13 show the synthetic control method graphs for total output and fixed capital respectively. For both these outcome variables, the post/pre 2003 MSPE is the largest for the treated states and hence the p-value is $1/24 = 0.042$. Hence, the treatment effect is significant at the 5% level for all outcomes of interest.

V.4 Productivity results

A major economic rationale for providing tax incentives to firms to locate to a particular region is the possibility of agglomeration economies. New firms entering a region might lead to positive productivity spillovers on existing firms. To test for agglomeration economies, I look at the differential effect of the policy change on TFP (total factor productivity) measures in the treated states compared to the control states. I use two different measures of productivity. First, I construct industry and firm-level TFP measures using the methodology of Levinsohn and Petrin (2003).[38] I also construct labor productivity measures defined as output per man-day and value added per man-day. Columns 1 through 3 in Table 15 show the results for the state-industry level DID regressions comparing aggregate productivity in the treated states to the control states. I find large increases in aggregate productivity across

[37] Note that this is a one-tailed test and the p-value depends on the number of control states.

[38] Details in the Appendix.

various measures of TFP. This differential increase in productivity is the cumulative effect of both new and existing firms in the treated states. A natural question to ask is whether the entry of new firms led to increases in productivity for the existing firms. In columns 4 and 5, I run firm-level regressions to look at the effect of the policy change on the productivity of existing firms, and find no effects. This suggests that the most of the aggregate productivity gains in the treated states are being driven by the entry of new firms. Although surprising, it must be kept in mind that this paper looks only at the short to medium term effect of the policy change and generally existing firms take some time to internalize agglomeration economies.[39] These differences in productivity levels between new and old firms are also shown in Figure 9.

V.5 Wages, Rents, and Migration

After looking at various industrial outcomes, it is important to investigate changes in the local economy. First, I test whether the policy resulted in earnings gains for the residents of the treated states. It is conceivable that migrant workers move in from other states to take advantage of the jobs available after the policy change. This might partially or completely offset the nominal wage gains due to the increase in local prices and rents in the treated states. I explicitly test for real earnings gain and differential migration as well. Table 16, column 1 reports the results of a difference-in-differences regression specification comparing total wages in treated states to neighboring states. This regression controls for district fixed effects and industry-year fixed effects along with individual controls such as age, sex, and education status. I find an 11% differential increase in nominal wages[40] for all workers in the treated states compared with the neighboring states. In column 2, I run the same specification restricted to districts along the border and find a 13% increase in wages. The magnitude of this effect is similar across the two specifications. Columns 3 and 4 look at wages of workers involved in non-agricultural activities and these wages go up by 14.5%. In columns 5 and 6, I look at the wages of agricultural workers and find no differential increase across the treated and control groups. This provides additional support

[39]For example, Greenstone, Hornbeck, and Moretti (2010) find that the total factor productivity (TFP) of incumbent plants grows five years after the opening of a large plant in their county. Also, Wang (2013) shows that there is positive TFP growth more than six years after the opening of an SEZ in China.

[40]The wages data comes from the NSS dataset that asks each household member their wages over the seven days preceding the interview.

for the results as the policy was similar to a labor demand shock for the industrial sector and should not have affected the agricultural sector. In columns 7 and 8, I compare housing rents across the treated and control groups controlling for district and year fixed effects along with housing attributes such as floor area, dwelling type, roof type, number of rooms etc. I find no statistically significant differential effect of the policy change on housing rents. This suggests that nominal wages might have gone up without corresponding increases in local prices.

However, housing rents might not be an apt measure of overall price levels in India and we would need a state-level consumer price index (CPI) to deflate the nominal values. A state-level CPI is not readily available for different states going back to the early 2000s. I construct an alternate price index at the state level (1999-2000 as the base year) using state GDP at constant and current prices from the RBI Handbook of Statistics. The price index for the neighboring states is shown in Appendix Table A2. I deflate wages and monthly per capita expenditure and run difference-in-differences specification comparing these outcomes in treated and control areas. In Table 17, columns 1 and 2 show that real wages increase by 12% -15%. The magnitude is similar to the increase in nominal wages in Table 17, suggesting that the policy did not differentially affect the price levels. Furthermore, in columns 3 and 4, I also find a differential increase in monthly per capita expenditure by around 10%. Finally, for column 5, I aggregate the total wages earned in the entire states (the state-level wage bill). I compare the total wage bill in the treated states to the neighboring control states and find a 52.8% increase in total wage bill.[41] This is comparable to the 40% increase in industrial wage bill estimated in Table 8.

These results suggest that the policy change did not induce differential migration into the treated regions from control states. I explicitly test for migration in treated areas compared with control areas in Table 18. The definition of migrants follows the questions in the NSS surveys that elicit information on the last usual place of residence of the respondent. Using this measure, I define an external migrant as a person whose last usual place of residence was another state; an internal migrant from within the state but another district. I also run specifications with economic migrants who report their reason for migration being work related. I find no statistically significant effect on external migrants and economic migrants. I find a negative differential effect on internal migrants. This might be because each district

[41]This is the differential increase in the total wage bill over seven days.

within the treated states had an industrial estate (and more jobs) leading to less within-state migration. Taken together, Table 18 suggests that there was no differential migration in response to the policy change in treated states compared with control states. This is consistent with previous literature on India documenting low migration.[42]

VI Cost-Benefit Analysis

Although the results in the paper suggest that the policy was successful in creating employment, output, and real earnings gains in the treated states, it might have come at large costs. I use the treatment effect coefficients to conduct a back-of-the-envelope cost-benefit analysis. I broadly follow Busso, Gregory, and Kline (2013) for this analysis, but in addition I include firm profits (as a benefit of the policy).[43] The benefits of the policy accrue to firm owners, workers,[44] and landowners in the treated states, whereas the costs to the government include the foregone tax revenues and the actual cost of the capital subsidy. For the ease of comparison, I provide all numbers in terms of 2007-08 (the last year in my analysis).

The benefits in the treated states can be broken down into three components: (i) increase in profits of firms in the treated areas, (ii) real wage earnings increases in the treated areas, and (iii) rental rate gains for landowners in the treated areas. For profits, I first estimate the treatment effect coefficient using a difference-in-differences specification as in all the earlier specifications. This is shown in Appendix Table B5. I use the estimated treatment effect coefficient on *post*treat* to calculate the magnitude of the total treatment effect. Note that the total treatment effect is the difference between the actual total and the counterfactual total in 2007-08. The counterfactual total is the amount that would have accrued in the absence of the policy and equals Actual total/$(1+\beta)$. Similarly, for the real wage bill gains in the treated states, I use the estimates from the total real wage bill regressions from Table 17, column 5. I multiply the weekly total wage bill by 52 to get the yearly total wage bill. As shown in Table 17, the rental rates do not change differentially and hence I assume the

[42]For example, Munshi and Rosenzweig (2009) find that in rural areas permanent migration rates of men out of their origin villages were as low as 8.7 percent in 1999.

[43]Busso, Gregory, and Kline (2013) measure the benefits of the EZ program as the total earnings increase of zone workers, earnings increase for non-resident commuters, improvements in local amenities, and value of rent reductions outside the zone due to decreases in population. They model firms as price-takers with constant returns to scale technology. Hence, profits of firms in their analysis is zero.

[44]Note that in my case, migration is zero and there are no non-resident commuters.

rental rate gains to landowners are negligible. I show the numbers in Table 19a. The total gains from the policy change are around 102 billion Rupees, of which 72 billion Rupees (USD 1.2 billion) accrues to firm owners[45] (as profits) and 30 billion Rupees (USD 480 million) accrues to workers (as wage bill).

To calculate the costs, I take into account (i) foregone corporate income tax revenue, (ii) foregone central excise tax revenue, and (iii) actual costs of the capital subsidy. I estimate the foregone tax revenue by calculating the revenue that the government would have collected in the absence of the policy. I use the estimated coefficient on *post*treat* on corporate income (Appendix table B5) and calculate the counterfactual.[46] I use a 35% corporate income tax rate to measure the total foregone revenue from the corporate income tax exemption. I use the same method to calculate the foregone revenue from the central excise tax exemption. The central excise tax is levied on the total value of output. I use the actual central excise tax receipts collected by the Government of India as a percentage of the value of output as the effective excise tax rate (7%).[47] These numbers are shown in Table 19b. The foregone revenue from corporate income taxes is 34.7 billion Rupees and that from central excise taxes is 29.3 billion Rupees.

The actual cost of the capital subsidies is also not readily available. In this analysis, I calculate an upper bound for the actual cost. The policy provided capital subsidies to new and old firms equaling 15% of their investment in plant and machinery up to a maximum amount of Rs. 3,000,000. I assume that every firm gets the maximum amount to provide an upper bound. In 2007-08, there were 2634 firms in the treated states.[48] Assuming each firm received Rs. 3,000,000, the total amount spent by the government would be Rs. 7.9 billion. Hence, the total loss for the government was approximately Rs. 72 billion (USD 1.2 billion). Comparing the costs and benefits, gives us a figure of Rs. 30 billion (USD 480 million) in benefits from the policy change. This is roughly 0.5% of the combined GDP of the two treated states.[49] Although these numbers provide suggestive evidence on the cost-

[45]Around 21% of the firms in the treated states were public limited companies. Individual proprietorships (13%), partnerships (26%) and private limited companies (38%) accounted for the bulk of the remaining firms. No other detailed shareholder information is available in the dataset.

[46]Counterfactual = Actual total/$(1+\beta)$.

[47]This figure comes from the Comptroller and Auditor General of India, Report No. CA 20 of 2009-10 - Union Government (Indirect Taxes).

[48]Figure from the ASI state$\times$industry panel data.

[49]Combined state GDP for Himachal Pradesh and Uttarakhand in 2007-08 was Rs. 598.6 billion (USD 9.6 billion).

effectiveness of the policy, I cannot conclude whether the policy was Pareto-improving or if the tax-incentive scheme was the most efficient transfer to the treated regions.

VII Conclusion

Many argue that a spatially targeted industrial policy is a waste of taxpayers' money as it simply reallocates economic activity across regions and does not lead to overall growth. Policy makers throughout the world however, use such location-based policies to help develop economically lagging regions. Whether location-based tax incentives are effective and help in industrialization and employment generation at the local level is largely an empirical question. In the last few years, there has been a growing empirical literature on place-based policies, mainly as more micro-data has become available. However, these policies have been understudied in developing countries where regional economic disparities can be large and labor mobility might be low.

In this paper, I critically examine a location-based tax incentive scheme that provided tax exemptions and capital subsidies to new and existing firms in two states in India, beginning 2003. I find that the policy change resulted in large increases in employment, output and capital - both due to entry of new firms and growth by existing firms. I document that the new firms entering the treated areas are larger and more productive but find no evidence for relocation of economic activity across the treated and control areas. The policy led to earnings increases for residents of the treated states without any corresponding increases in housing rents or local prices. I also show that these real earnings gain might be a result of low migration into the treated states. Finally, I use my estimates to conduct a simple cost-benefit analysis that suggests that the policy was cost-effective.

An important caveat is that these results are at best medium-term effects of the policy change on various economic outcomes. It will be interesting to look at the long run impacts of this policy after the removal of the incentives [see Kline and Moretti (2013a)]. Whether or not such policies have a lasting impact (for example, agglomeration economies)[50] or only attract fly-by-night operators that shut shop and relocate to the next area with such benefits is an important issue but beyond the current scope of this paper. With more data available in the following years, this seems to be a promising avenue for future research.

[50] See Greenstone, Hornbeck, and Moretti (2010).

References

Abadie, A., A. Diamond, and J. Hainmueller (2010): "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 105(490), 493–505.

Abadie, A., and J. Gardeazabal (2003): "The Economic Costs of Conflict: A Case Study of the Basque Country," *The American Economic Review*, 93(1), pp. 113–132.

Bollard, Albert, P. K., and G. Sharma (2013): "India's Mysterious Manufacturing Miracle," *Review of Economic Dynamics*.

Bondonio, D., and R. T. Greenbaum (2007): "Do local tax incentives affect economic growth? What mean impacts miss in the analysis of enterprise zone policies," *Regional Science and Urban Economics*, 37(1), 121–136.

Bronzini, R., and G. de Blasio (2006): "Evaluating the impact of investment incentives: The case of Italy's Law 488/1992," *Journal of Urban Economics*, 60(2), 327–349.

Busso, M., J. Gregory, and P. Kline (2013): "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 103(2), 897–947.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 90(3), 414–427.

——— (2011): "Robust Inference With Multiway Clustering," *Journal of Business & Economic Statistics*, 29(2), 238–249.

Criscuolo, C., R. Martin, H. Overman, and J. van Reenen (2012): "The Causal Effects of an Industrial Policy," *NBER Working Paper 17842*.

Duranton, G., L. Gobillon, and H. G. Overman (2011): "Assessing the Effects of Local Taxation using Microgeographic Data," *Economic Journal*, 121(555), 1017–1046.

Givord, P., R. Rathelot, and P. Sillard (2013): "Place-based tax exemptions and displacement effects: An evaluation of the Zones Franches Urbaines program," *Regional Science and Urban Economics*, 43(1), 151–163.

Glaeser, E. L. (2001): "The Economics of Location-Based Tax Incentives," *Harvard Institute of Economic Research Working Papers 1932*.

Glaeser, E. L., and J. D. Gottlieb (2008): "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity*, 39(1), 155–253.

Greenbaum, R. T., and J. B. Engberg (2004): "The impact of state enterprise zones on urban manufacturing establishments.," *Journal of Policy Analysis and Management*, 23, 315–339.

Greenstone, M., R. Hornbeck, and E. Moretti (2010): "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 118(3), 536–598.

Ham, J., C. Swenson, A. Imrohoroglu, and H. Song (2011): "Government Programs Can Improve Local Labor Markets: Evidence from State Enterprise Zones, Federal Empowerment Zones and Federal Enterprise Communities," *Journal of Public Economics*, 95, 779–797.

Harrison, A. E., L. A. Martin, and S. Nataraj (2011): "Learning versus stealing: How Important are market-share reallocations to India's productivity growth?," *World Bank Economic Review*, forthcoming.

Hnatkovska, V., and A. Lahiri (2013): "Structural Transformation and the Rural-Urban Divide," *Working Paper, University of British Columbia*.

Kline, P., and E. Moretti (2013a): "Local Economic Development, Agglomeration Economies and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *forthcoming Quarterly Journal of Economics*.

——— (2013b): "People, Places and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *Working Paper*.

Levinsohn, J., and A. Petrin (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70(2), 317–341.

Mayer, T., F. Mayneris, and L. Py (2012): "The Impact of Urban Enterprise Zones on Establishments' Location Decisions: Evidence from French ZFU," *mimeo Sciences-Po*.

Moretti, E. (2011): "Local Labor Markets," *Handbook of Labor Economics*, Chapter 14, Volume 4, 1237–1313.

Munshi, K., and M. Rosenzweig (2009): "Why is Mobility in India so Low? Social Insurance, Inequality, and Growth," NBER Working Papers 14850, National Bureau of Economic Research, Inc.

Neumark, D., and J. Kolko (2010): "Do enterprise zones create jobs? Evidence from California's enterprise zone program," *Journal of Urban Economics*, 68(1), 1–19.

Rathelot, R., and P. Sillard (2008): "The Importance of Local Corporate Taxes in Business Location Decisions: Evidence From French Micro Data," *Economic Journal*, 118(527), 499–514.

Roback, J. (1982): "Wages, Rents, and the Quality of Life," *The Journal of Political Economy*, 90(6), 1257–1278.

Topalova, P. (2010): "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India," *American Economic Journal: Applied Economics*, 2(4), 1–41.

Wang, J. (2013): "The economic impact of Special Economic Zones: Evidence from Chinese municipalities," *Journal of Development Economics*, 101, 133 – 147.

Table 1: Summary statistics at the state×3-digit industry level

	State	Time period	Observations	Mean	Std Dev
Number employed	All states	Pre-2003	3604	3962.88	12496.62
		Post-2003	6179	4162.30	13082.27
	Treated	Pre-2003	231	589.68	1214.31
		Post-2003	424	834.16	1545.70
	Neighboring states	Pre-2003	709	3153.96	6118.20
		Post-2003	1203	3554.21	6825.79
	All major states	Pre-2003	2746	4952.17	14023.84
		Post-2003	4630	5237.32	14814.82
Number of factories	All states	Pre-2003	3604	107.64	252.96
		Post-2003	6179	112.77	270.42
	Treated	Pre-2003	231	15.98	27.82
		Post-2003	424	21.15	32.17
	Neighboring states	Pre-2003	709	103.90	163.59
		Post-2003	1203	110.14	188.27
	All major states	Pre-2003	2746	134.31	282.95
		Post-2003	4630	141.98	304.79
Total output	All states	Pre-2003	3604	83793.55	240913.20
(in '00,000 Rs.)		Post-2003	6179	162689.00	515128.30
	Treated	Pre-2003	231	15068.07	26703.81
		Post-2003	424	41973.70	94545.83
	Neighboring states	Pre-2003	709	73225.33	128034.00
		Post-2003	1203	131358.60	256813.00
	All major states	Pre-2003	2746	103814.20	271097.10
		Post-2003	4630	202135.20	584000.10
Fixed capital	All states	Pre-2003	3604	35414.08	133253.90
(in '00,000 Rs.)		Post-2003	6179	51037.45	180756.90
	Treated	Pre-2003	231	7358.37	20079.97
		Post-2003	424	18860.29	61088.17
	Neighboring states	Pre-2003	709	24231.10	57330.15
		Post-2003	1203	32682.40	79836.12
	All major states	Pre-2003	2746	44306.31	150713.20
		Post-2003	4630	63988.66	205354.20
Wage bill	All states	Pre-2003	3604	5373.16	14082.43
(in '00,000 Rs.)		Post-2003	6179	7623.88	19639.85
	Treated	Pre-2003	231	1167.24	3509.69
		Post-2003	424	1936.53	4662.96
	Neighboring states	Pre-2003	709	4166.21	6912.09
		Post-2003	1203	6265.78	10883.73
	All major states	Pre-2003	2746	6775.80	15810.68
		Post-2003	4630	9684.05	22205.34

Notes: Treated states: Uttarakhand and Himachal Pradesh; Neighboring states: Haryana, Punjab, Delhi, Chandigarh, Uttar Pradesh; All major states: Neighboring states plus Rajasthan, Bihar, Andhra Pradesh, Chhattisgarh, Maharashtra, Madhya Pradesh, Orissa, Goa, Kerala, Karnataka, Tamil Nadu, Jharkhand, Gujarat, and West Bengal. Observations here are state×3-digit-year observations.

Table 2: Summary statistics at the firm-level

	State	Time period	Observations	Mean	Std Dev
Number employed	All states	Pre-2003	93276	41.50	241.12
		Post-2003	195581	42.64	198.55
	Treated	Pre-2003	1915	39.63	151.29
		Post-2003	5222	43.43	127.28
	Neighboring states	Pre-2003	18192	32.74	120.00
		Post-2003	39710	36.03	125.26
	All major states	Pre-2003	84879	41.55	245.93
		Post-2003	176605	42.62	202.64
Total output	All states	Pre-2003	88126	927.58	13589.21
(in ‘00,000 Rs.)		Post-2003	187796	1726.92	32981.24
	Treated	Pre-2003	1743	1097.01	4461.96
		Post-2003	4733	2300.66	8030.17
	Neighboring states	Pre-2003	17234	800.43	6734.20
		Post-2003	38348	1361.89	1361.89
	All major states	Pre-2003	80680	919.38	13874.97
		Post-2003	170818	1700.22	33725.87
Fixed capital	All states	Pre-2003	96718	357.52	8204.424
(in ‘00,000 Rs.)		Post-2003	206686	494.65	11057.01
	Treated	Pre-2003	1961	471.02	3227.43
		Post-2003	5461	928.29	9743.92
	Neighboring states	Pre-2003	18741	240.86	3613.18
		Post-2003	42523	309.13	3426.12
	All major states	Pre-2003	87607	359.06	8383.13
		Post-2003	185860	493.56	493.56
Wage bill	All states	Pre-2003	98444	22.79	280.09
(in ‘00,000 Rs.)		Post-2003	211175	26.98	336.92
	Treated	Pre-2003	1972	28.25	264.65
		Post-2003	5537	31.39	292.88
	Neighboring states	Pre-2003	18847	16.67	110.60
		Post-2003	42922	20.74	131.72
	All major states	Pre-2003	89306	23.05	285.46
		Post-2003	190186	27.31	344.27

Notes: Treated states: Uttarakhand and Himachal Pradesh; Neighboring states: Haryana, Punjab, Delhi, Chandigarh, Uttar Pradesh; All major states: Neighboring states plus Rajasthan, Bihar, Andhra Pradesh, Chhattisgarh, Maharashtra, Madhya Pradesh, Orissa, Goa, Kerala, Karnataka, Tamil Nadu, Jharkhand, Gujarat, and West Bengal. Observations here are firm-year observations.

Table 3: Summary statistics for migration, wages, and housing rents

	States	Time Period	Observations	Mean	Std. Dev
External migrant	Treated states	Pre-2003	15194	0.05	0.22
		Post-2003	16575	0.09	0.28
	Neighboring states	Pre-2003	118045	0.04	0.19
		Post-2003	103797	0.05	0.22
Internal migrant	Treated states	Pre-2003	15194	0.09	0.28
		Post-2003	16575	0.06	0.24
	Neighboring states	Pre-2003	118045	0.09	0.28
		Post-2003	103797	0.09	0.29
Economic migrant	Treated states	Pre-2003	15194	0.10	0.30
		Post-2003	16575	0.11	0.31
	Neighboring states	Pre-2003	118045	0.05	0.22
		Post-2003	103797	0.06	0.23
Total wages (in Rs.)	Treated states	Pre-2003	1714	879.36	861.36
(all workers)		Post-2003	6327	1080.16	1117.18
	Neighboring states	Pre-2003	11772	691.05	945.31
		Post-2003	31724	829.97	1075.31
Housing rents	Treated states	Pre-2003	392	422.92	377.45
(in Rs.)		Post-2003	511	749.74	804.42
	Neighboring states	Pre-2003	2384	656.28	954.79
		Post-2003	3517	1176.71	1502.73

Notes: Treated states: Uttarakhand and Himachal Pradesh; Neighboring states: Haryana, Punjab, Delhi, Chandigarh, Uttar Pradesh; External migrant is defined as a person whose last usual place of residence is outside the state; Internal migrant is one whose last usual place of residence is the same state but a different district; Economic migrant is one who migrated for a work related reason. Total wages are defined as the wages earned over the seven days preceding the interview.

Table 4: Log employment

	(1)	(2)	(3)	(4)
Panel A				
*post*treat*	0.371***	0.379***	0.418***	0.372***
	(0.0740)	(0.0827)	(0.0884)	(0.0999)
	[0.129]	[0.143]	[0.141]	[0.154]
Observations	2,567	2,567	2,567	2,567
R-squared	0.688	0.709	0.691	0.712
Panel B				
*post*treat*	0.427***	0.443***	0.439***	0.448***
	(0.0877)	(0.0899)	(0.0788)	(0.0822)
	[0.124]	[0.138]	[0.131]	[0.141]
Observations	8,028	8,028	8,028	8,028
R-squared	0.625	0.634	0.626	0.634
state FE	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes
3 digit industry FE	Yes	Yes	Yes	Yes
3 digit industry-year FE	No	Yes	No	Yes
time varying controls	No	No	Yes	Yes

Notes: Dependent variable is log employment in the 3 digit industry in a particular state. The coefficient on the interaction term post*treat shows the treatment effect. Time varying controls include pre-treatment state-level variables interacted with a time dummy for each year. Panel A shows the regressions with the neighboring states as the control group whereas Panel B uses all major states as the control group. Standard errors in parentheses are clustered at the state-year level. Standard errors in square brackets are clustered at the state and year level using CGM multi-way clustering. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 5: Log of total factories

	(1)	(2)	(3)	(4)
Panel A				
*post*treat*	0.268***	0.267***	0.291***	0.270***
	(0.0569)	(0.0616)	(0.0792)	(0.0815)
	[0.107]	[0.107]	[0.140]	[0.133]
Observations	2,567	2,567	2,567	2,567
R-squared	0.725	0.736	0.727	0.738
Panel B				
*post*treat*	0.310***	0.311***	0.315***	0.315***
	(0.0680)	(0.0693)	(0.0596)	(0.0610)
	[0.0943]	[0.105]	[0.0955]	[0.104]
Observations	8,031	8,031	8,031	8,031
R-squared	0.715	0.719	0.716	0.719
state FE	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes
3 digit industry FE	Yes	Yes	Yes	Yes
3 digit industry-year FE	No	Yes	No	Yes
time varying controls	No	No	Yes	Yes

Notes: Dependent variable is the log of total number of factories in a 3 digit industry in a particular state. The coefficient on the interaction term post*treat shows the treatment effect. Time varying controls include pre-treatment state-level variables interacted with a time dummy for each year. Panel A shows the regressions with the neighboring states as the control group whereas Panel B uses all major states as the control group. Standard errors in parentheses are clustered at the state-year level. Standard errors in square brackets are clustered at the state and year level using CGM multi-way clustering. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 6: Log of total output

	(1)	(2)	(3)	(4)
Panel A				
*post*treat*	0.579***	0.584***	0.619***	0.566***
	(0.0969)	(0.108)	(0.116)	(0.132)
	[0.158]	[0.180]	[0.199]	[0.199]
Observations	2,567	2,567	2,567	2,567
R-squared	0.655	0.679	0.658	0.682
Panel B				
*post*treat*	0.561***	0.577***	0.623***	0.639***
	(0.106)	(0.107)	(0.101)	(0.104)
	[0.146]	[0.159]	[0.160]	[0.173]
Observations	8,031	8,031	8,031	8,031
R-squared	0.611	0.622	0.611	0.622
state FE	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes
3 digit industry FE	Yes	Yes	Yes	Yes
3 digit industry-year FE	No	Yes	No	Yes
time varying controls	No	No	Yes	Yes

Notes: Dependent variable is the log of total output in a 3 digit industry in a particular state. The coefficient on the interaction term post*treat shows the treatment effect. Time varying controls include pre-treatment state-level variables interacted with a time dummy for each year. Panel A shows the regressions with the neighboring states as the control group whereas Panel B uses all major states as the control group. Standard errors in parentheses are clustered at the state-year level. Standard errors in square brackets are clustered at the state and year level using CGM multi-way clustering. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 7: Log of fixed capital

	(1)	(2)	(3)	(4)
Panel A				
*post*treat*	0.718***	0.714***	0.878***	0.866***
	(0.132)	(0.147)	(0.209)	(0.225)
	[0.236]	[0.264]	[0.322]	[0.347]
Observations	2,567	2,567	2,567	2,567
R-squared	0.668	0.691	0.672	0.695
Panel B				
*post*treat*	0.711***	0.728***	0.776***	0.787***
	(0.158)	(0.160)	(0.145)	(0.149)
	[0.230]	[0.243]	[0.244]	[0.258]
Observations	8,030	8,030	8,030	8,030
R-squared	0.627	0.635	0.628	0.636
state FE	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes
3 digit industry FE	Yes	Yes	Yes	Yes
3 digit industry-year FE	No	Yes	No	Yes
time varying controls	No	No	Yes	Yes

Notes: Dependent variable is the log of fixed capital in a 3 digit industry in a particular state. The coefficient on the interaction term post*treat shows the treatment effect. Time varying controls include pre-treatment state-level variables interacted with a time dummy for each year. Panel A shows the regressions with the neighboring states as the control group whereas Panel B uses all major states as the control group. Standard errors in parentheses are clustered at the state-year level. Standard errors in square brackets are clustered at the state and year level using CGM multi-way clustering. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 8: Log of wage bill

	(1)	(2)	(3)	(4)
Panel A				
*post*treat*	0.392***	0.400***	0.322***	0.275***
	(0.0826)	(0.0957)	(0.0788)	(0.0935)
	[0.148]	[0.173]	[0.134]	[0.140]
Observations	2,567	2,567	2,567	2,567
R-squared	0.660	0.682	0.663	0.685
Panel B				
*post*treat*	0.412***	0.427***	0.437***	0.449***
	(0.0938)	(0.0983)	(0.0891)	(0.0948)
	[0.138]	[0.157]	[0.145]	[0.165]
Observations	8,030	8,030	8,030	8,030
R-squared	0.604	0.611	0.604	0.612
state FE	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes
3 digit industry FE	Yes	Yes	Yes	Yes
3 digit industry-year FE	No	Yes	No	Yes
time varying controls	No	No	Yes	Yes

Notes: Dependent variable is the log of wage bill in a 3 digit industry in a particular state. The coefficient on the interaction term post*treat shows the treatment effect. Time varying controls include pre-treatment state-level variables interacted with a time dummy for each year. Panel A shows the regressions with the neighboring states as the control group whereas Panel B uses all major states as the control group. Standard errors in parentheses are clustered at the state-year level. Standard errors in square brackets are clustered at the state and year level using CGM multi-way clustering. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 9: Number of firm closures

	(1) closed	(2) closed	(3) closed	(4) closed
*post*treat*	0.0487	0.311	0.469	-0.377
	(0.396)	(0.304)	(0.556)	(0.319)
Observations	2,567	2,567	8,031	8,031
R-squared	0.312	0.321	0.275	0.277
state FE	Yes	Yes	Yes	Yes
3 digit industry-year FE	Yes	Yes	Yes	Yes
time varying controls	No	Yes	No	Yes
Control Group	Neighboring states	Neighboring states	Major states	Major states

Notes: The coefficient on the interaction term post*treat shows the treatment effect. Time varying controls include pre-treatment state-level variables interacted with a time dummy for each year. Standard errors are clustered at the state-year level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 10: Testing for relocation of industrial activity

	(1)	(2)	(3)	(4)
	Log (employed)	Log (total factories)	Log (total output)	Log (fixed capital)
*post*neighbors*	0.0501*	0.00974	-0.0371	-0.0273
	(0.0267)	(0.0168)	(0.0338)	(0.0374)
Observations	7,375	7,375	7,375	7,375
R-squared	0.629	0.716	0.621	0.635
state FE	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes
3 digit industry FE	Yes	Yes	Yes	Yes

Notes: The coefficient on the interaction term post*neighbors shows the effect of the policy change on neighboring states as compared to all other major states. The two treated states are omitted in this regression. Standard errors are clustered at the state-year level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 11: Firm-level regressions

	(1)	(2)	(3)	(4)
Log (employed)				
*post*treat*	0.0747*	0.103**	0.0740	0.110*
	(0.0386)	(0.0400)	(0.0577)	(0.0544)
Observations	63,629	63,629	13,185	13,185
R-squared	0.939	0.942	0.946	0.953
Log (total output)				
*post*treat*	0.0866**	0.114***	0.177**	0.237***
	(0.0430)	(0.0414)	(0.0711)	(0.0722)
Observations	60,664	60,664	12,315	12,315
R-squared	0.965	0.967	0.970	0.976
Log (wage bill)				
*post*treat*	0.0807**	0.113**	0.108**	0.129**
	(0.0399)	(0.0423)	(0.0466)	(0.0565)
Observations	63,650	63,650	13,193	13,193
R-squared	0.949	0.952	0.954	0.960
Log (fixed capital)				
*post*treat*	0.0552	0.0709*	0.0804	0.0812
	(0.0368)	(0.0387)	(0.0501)	(0.0495)
Observations	67,033	67,033	14,125	14,125
R-squared	0.973	0.974	0.978	0.981
Log (additions to fixed capital)				
*post*treat*	0.278**	0.275**	0.277	0.406**
	(0.120)	(0.108)	(0.170)	(0.154)
Observations	52,906	52,906	10,503	10,503
R-squared	0.856	0.864	0.875	0.899
Log (additions to plant and machinery)				
*post*treat*	0.255**	0.256**	0.199*	0.246
	(0.110)	(0.101)	(0.113)	(0.216)
Observations	41,674	41,674	8,222	8,222
R-squared	0.858	0.869	0.883	0.914
Control group	Neighboring states	Neighboring states	Border districts	Border districts
firm FE	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes
4-digit industry FE	Yes	Yes	Yes	Yes
4-digit industry year FE	No	Yes	No	Yes
Age Controls	Yes	Yes	Yes	Yes

Notes: The coefficient on the interaction term post*treat shows the treatment effect. Time varying controls include pre-treatment state-level variables interacted with a time dummy for each year. Standard errors in columns 1 to 2 are clustered at the state-year level, and at the district level for columns 3 and 4. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 12: Removing multi-establishment firms

	(1)	(2)	(3)	(4)
	Log (employed)	Log (output)	Log (fixed capital)	Log (additions to plant and machinery)
*post*treat*	0.0919**	0.144***	0.0656*	0.232**
	(0.0437)	(0.0476)	(0.0383)	(0.103)
Observations	55,599	52,682	58,796	34,695
R-squared	0.939	0.969	0.973	0.871
firm FE	Yes	Yes	Yes	Yes
4-digit industry-year FE	Yes	Yes	Yes	Yes
Age Controls	Yes	Yes	Yes	Yes

Notes: The coefficient on the interaction term post*treat shows the treatment effect. Control group includes the neighboring states. Standard errors are clustered at the district level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 13: Firm-level regression - testing for spillovers

	(1)	(2)	(3)
	Log (employed)	Log (total output)	Log (fixed capital)
*post*treat*	0.151***	0.162**	0.0388
	(0.0535)	(0.0698)	(0.0628)
Observations	17,456	16,316	18,139
R-squared	0.958	0.975	0.978
firm FE	Yes	Yes	Yes
4-digit industry-year FE	Yes	Yes	Yes
Age Controls	Yes	Yes	Yes

Notes: The coefficient on the interaction term post*treat shows the treatment effect. Control group includes districts that are away from the borders in the neighboring states. Standard errors are clustered at the district level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 14: Testing for spillovers - comparing bordering districts to districts further away

	(1)	(2)	(3)
	Log(employed)	Log(total output)	Log(fixed capital)
*post*neighoring-district*	0.0575	-0.128	-0.0749
	(0.0579)	(0.0799)	(0.0808)
Observations	17,451	16,679	18,708
R-squared	0.955	0.979	0.978
firm FE	Yes	Yes	Yes
4-digit industry-year FE	Yes	Yes	Yes
Age Controls	Yes	Yes	Yes

Notes: The coefficient on the interaction term post*neighoring-district shows the treatment effect where the treated group is the bordering districts in the neighboring states. Control group includes districts that are away from the borders in the neighboring states. Standard errors are clustered at the district level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 15: Productivity regressions

	(1)	(2)	(3)	(4)	(5)
	Log (TFP) (Levinsohn-Petrin)	Log (labor productivity 1) (Value added / man-days)	Log (labor productivity 2) (Total output/man-days)	Log (labor productivity) (Total output/man-days)	Log (TFP) (Levinsohn-Petrin)
*post*treat*	0.288***	0.388***	0.191***	-0.0334	-0.00412
	(0.0726)	(0.0640)	(0.0381)	(0.0348)	(0.0767)
Observations	2,472	2,482	2,511	58,086	46,213
R-squared	0.634	0.332	0.486	0.932	0.859
state FE	Yes	Yes	Yes	-	-
year FE	Yes	Yes	Yes	Yes	Yes
3 digit industry FE	Yes	Yes	Yes	-	-
firm FE	No	No	No	Yes	Yes
4-digit industry year FE	No	No	No	Yes	Yes
Age Controls	No	No	No	Yes	Yes

Notes: The coefficient on the interaction term post*treat shows the treatment effect. Columns 1 through 3 show state-industry regressions with the neighboring states as the control group. Columns 4 and 5 show firm-level regressions. Firm fixed effects subsume state fixed effects. Standard errors are clustered at the state-year level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 16: Nominal wages and Rents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log (wages) [all]	Log (wages) [all]	Log (wages) [non-agricultural]	Log (wages) [non-agricultural]	Log (wages) [agricultural]	Log (wages) [agricultural]	Log (rent) [all]	Log (rent) [all]
*post*treat*	0.111**	0.132*	0.147***	0.145*	0.0215	-0.0273	0.396	-0.0583
	(0.0531)	(0.0640)	(0.0561)	(0.0714)	(0.135)	(0.189)	(0.266)	(0.313)
Observations	51,455	10,189	40,964	8,387	10,491	1,802	3,500	603
R-squared	0.667	0.615	0.642	0.660	0.433	0.239	0.540	0.451
district FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No
controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
control group	Neighboring states	Bordering districts	Neighboring states	Bordering districts	Neighboring states	Bordering districts	Neighboring states	Bordering districts

Notes: The coefficient on the interaction term post*treat shows the treatment effect. Controls for columns 1 through 6 include age, sex, educational status, marital status and relationship to household head. Controls for columns 7 and 8 include attributes of the house such as number of rooms, kitchen type, dwelling type, roof type and floor type. Standard errors are clustered at the district level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 17: Real wages and real monthly per capita expenditure

	(1) Log (real wages)	(2) Log (real wages)	(3) Log (real MPCE)	(4) Log (real MPCE)	(5) Log (real wage bill)
*post*treat*	0.122**	0.146**	0.0967**	0.107	0.528***
	(0.0495)	(0.0616)	(0.0408)	(0.0739)	(0.168)
Observations	51,455	10,189	88,731	17,157	531
R-squared	0.668	0.605	0.358	0.243	0.287
district FE	Yes	Yes	Yes	Yes	Yes
industry-year FE	Yes	Yes	No	No	No
Controls	Yes	Yes	Yes	Yes	No
control group	Neighboring states	Bordering districts	Neighboring states	Bordering districts	Neighboring states

Notes: The coefficient on the interaction term post*treat shows the treatment effect. Controls for columns 1 and 2 include age, sex, educational status, marital status and relationship to household head. Controls for columns 3 and 4 include household type, social group, rural-urban and religion. Standard errors are clustered at the district level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 18: Testing for differential migration

	(1) external migrant	(2) external migrant	(3) internal migrant	(4) internal migrant	(5) economic migrant	(6) economic migrant
*post*treat*	0.0172	0.0396	-0.0252**	-0.0236	0.00486	0.0349
	(0.0272)	(0.0412)	(0.0120)	(0.0173)	(0.0244)	(0.0311)
Observations	253,611	45,301	253,611	45,301	253,611	45,301
R-squared	0.087	0.043	0.031	0.017	0.076	0.026
district FE	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes
Control group	Neighboring states	Bordering districts	Neighboring states	Bordering districts	Neighboring states	Bordering districts

Notes: The coefficient on the interaction term post*treat shows the treatment effect. External migrant is one whose last usual place of residence was another state or country. Internal migrant's last usual place of residence was the same state but another district. An economic migrant migrated for work related reasons. Standard errors are clustered at the district level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table 19a: Total Benefits

	Actual Total in 2007-08 (billion Rupees)	Treatment effect coefficient	Total impact in 2007-08 (in billion Rupees)
profits	148.85	0.93	71.80
total wage bill	87.36	0.528	30.19

Table 19b: Total Costs

	Actual Total in 2007-08 (in billion Rupees)	Treatment effect coefficient	Counterfactual (in billion Rupees)	Tax rate (percent)	Loss in revenue (in billion Rupees)
corporate income	176.31	0.78	99.05	0.35	34.67
total output	660.75	0.579	418.46	0.07	29.29

Figure 1: Map of India

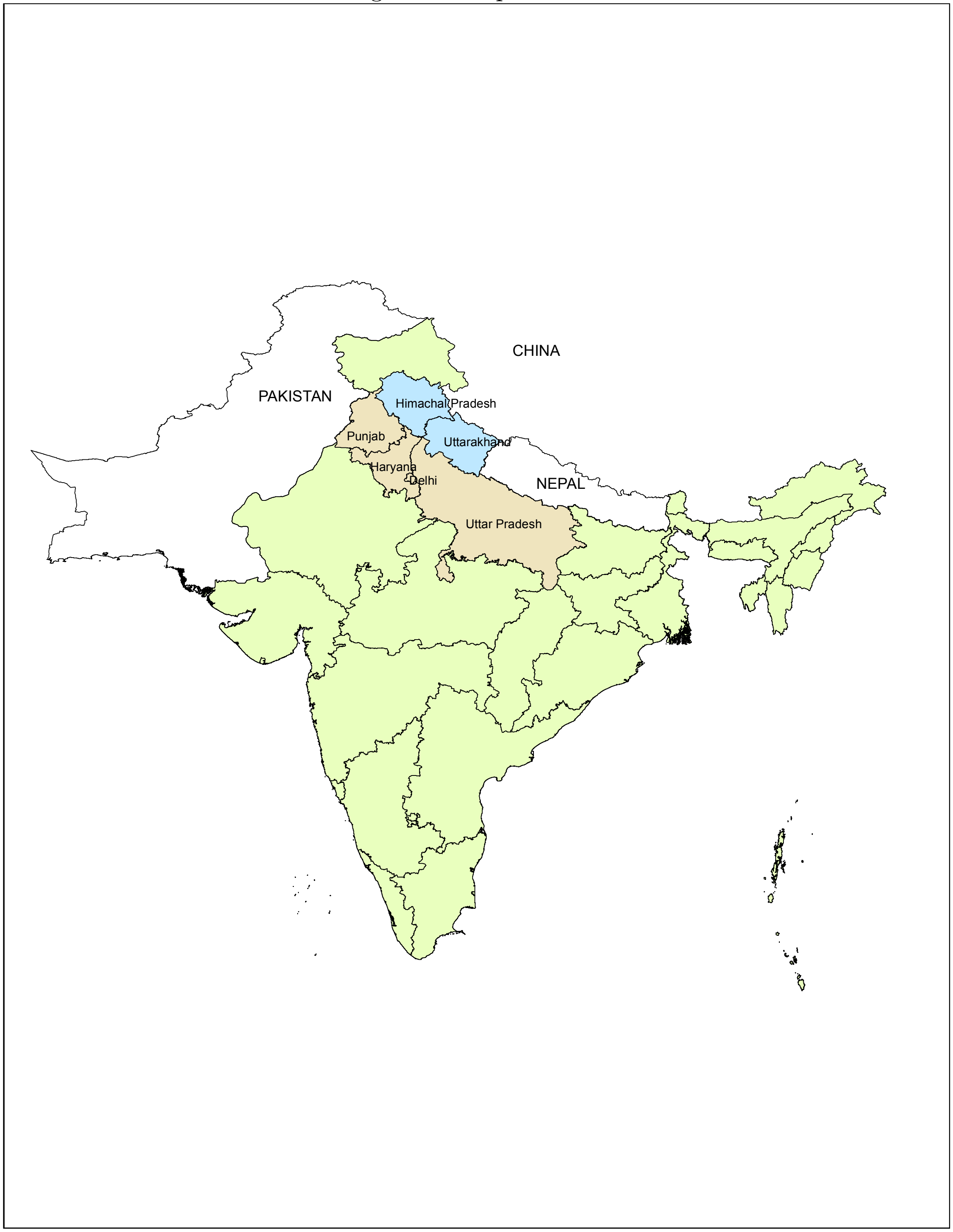
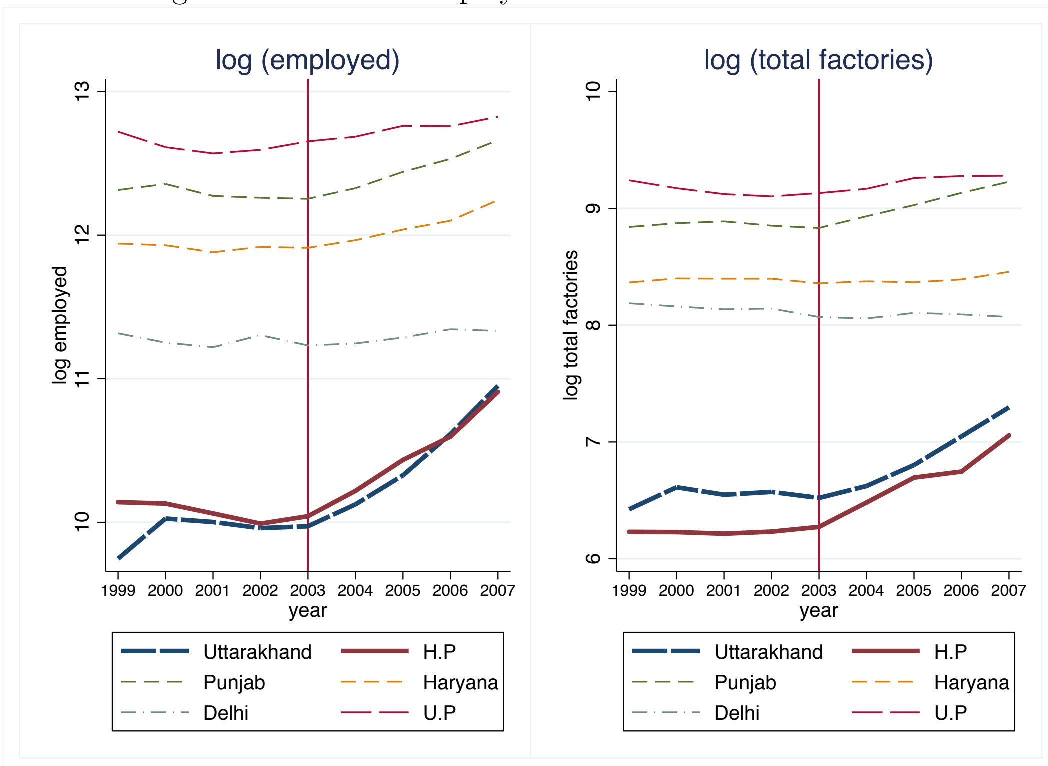
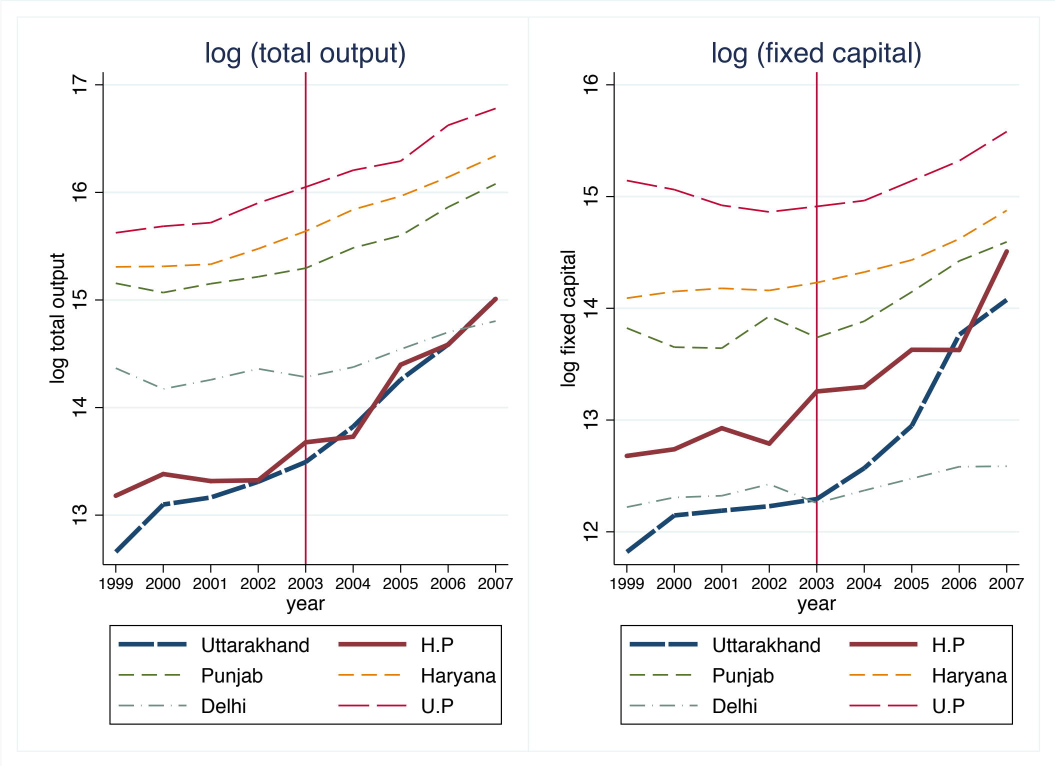


Figure 2: Trends in employment and number of factories



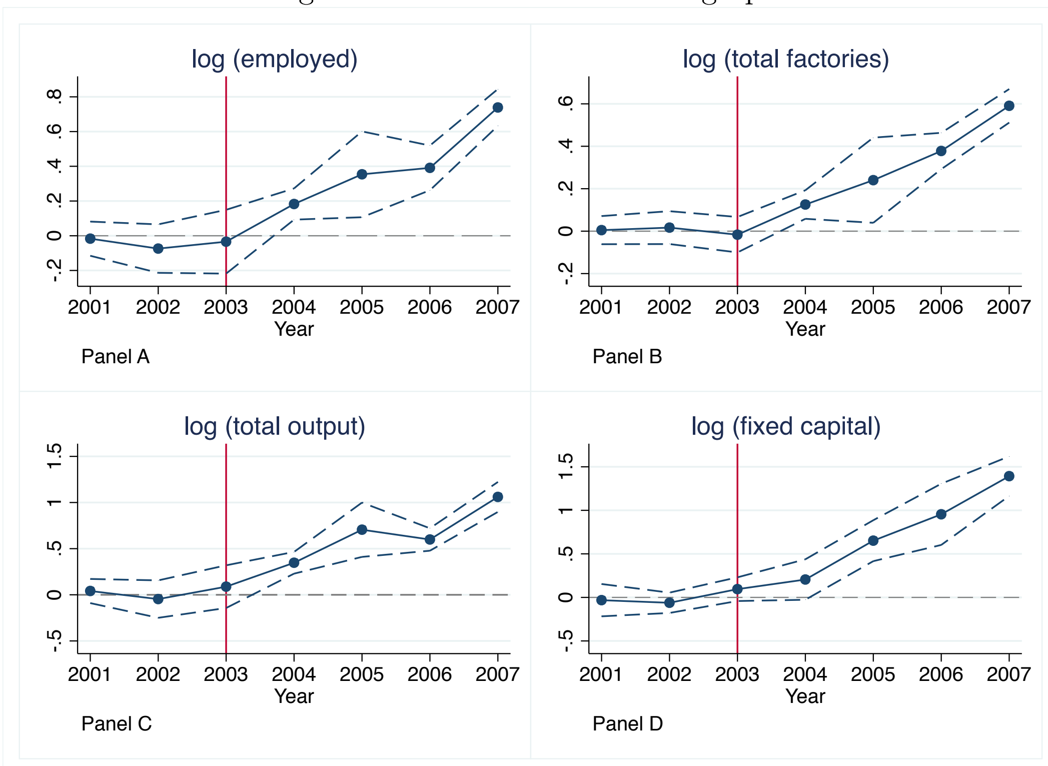
Notes: ASI state×industry data from 1999-2000 to 2007-08.

Figure 3: Trends in total output and fixed capital



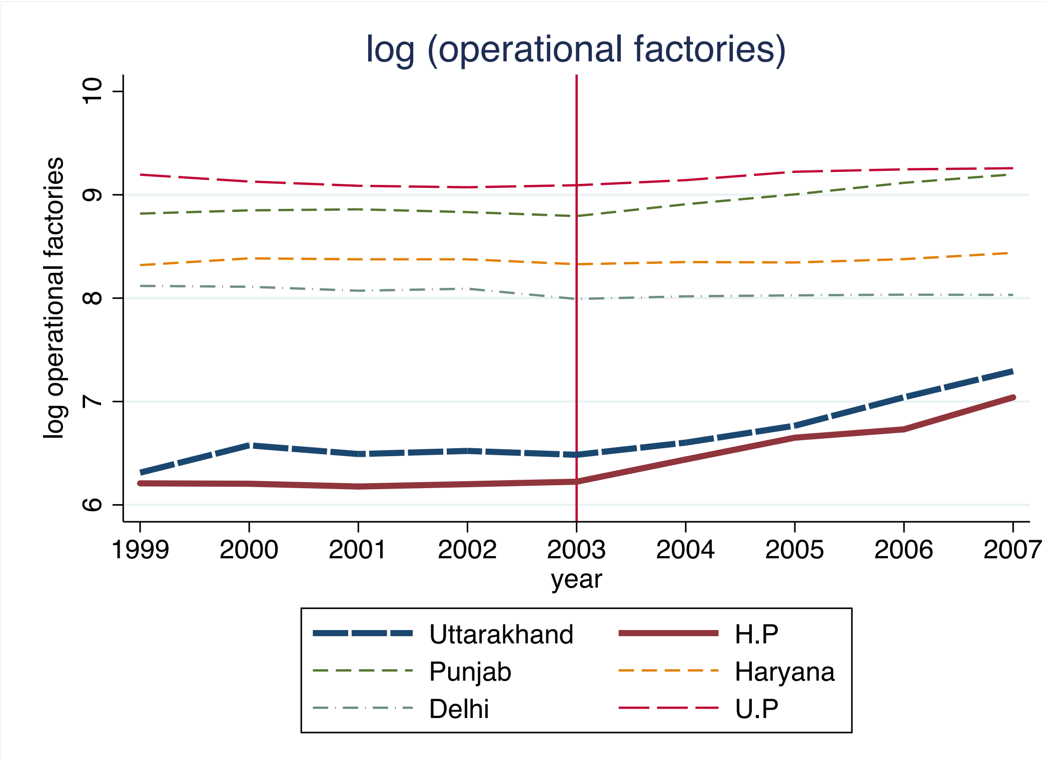
Notes: ASI state×industry data from 1999-2000 to 2007-08.

Figure 4: Estimated coefficient graphs



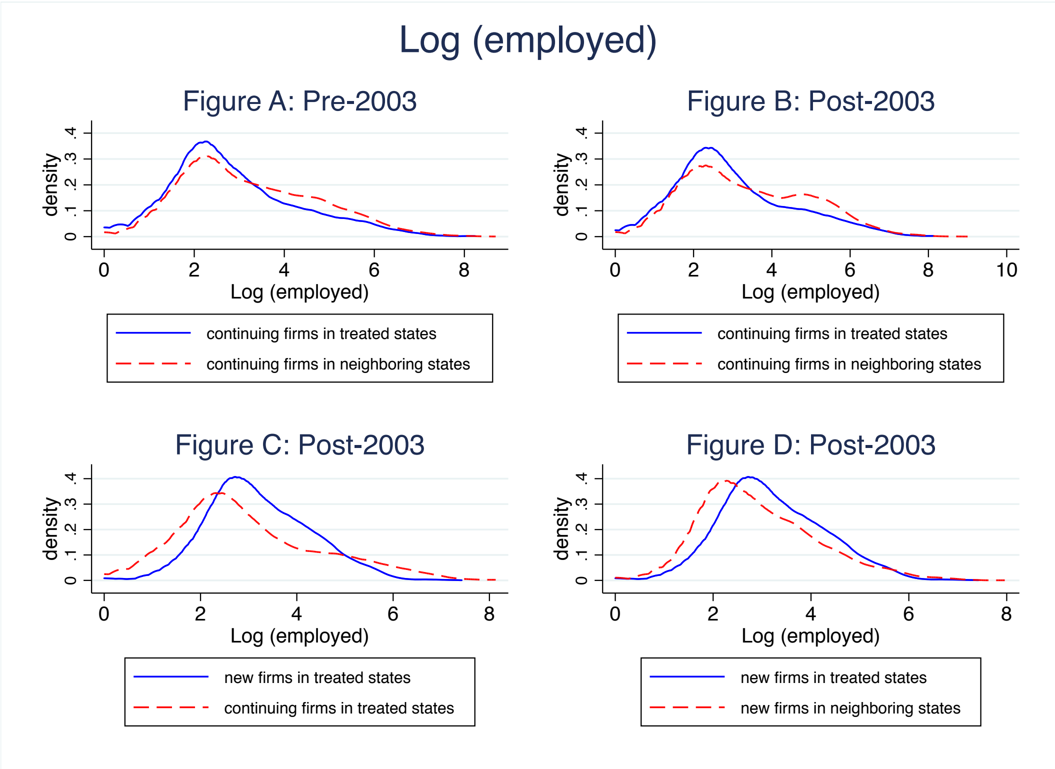
Notes: These graphs plot the coefficients obtained from a regression of the outcome variable (mentioned on top of the graph) on the interaction between the treated states dummy and year dummies. The regressions control for state, year, and 3-digit industry fixed effects. The Y-axis shows the estimated coefficients and the X-axis shows the various years. Standard errors are clustered at the state-year level.

Figure 5: Trends in operational factories



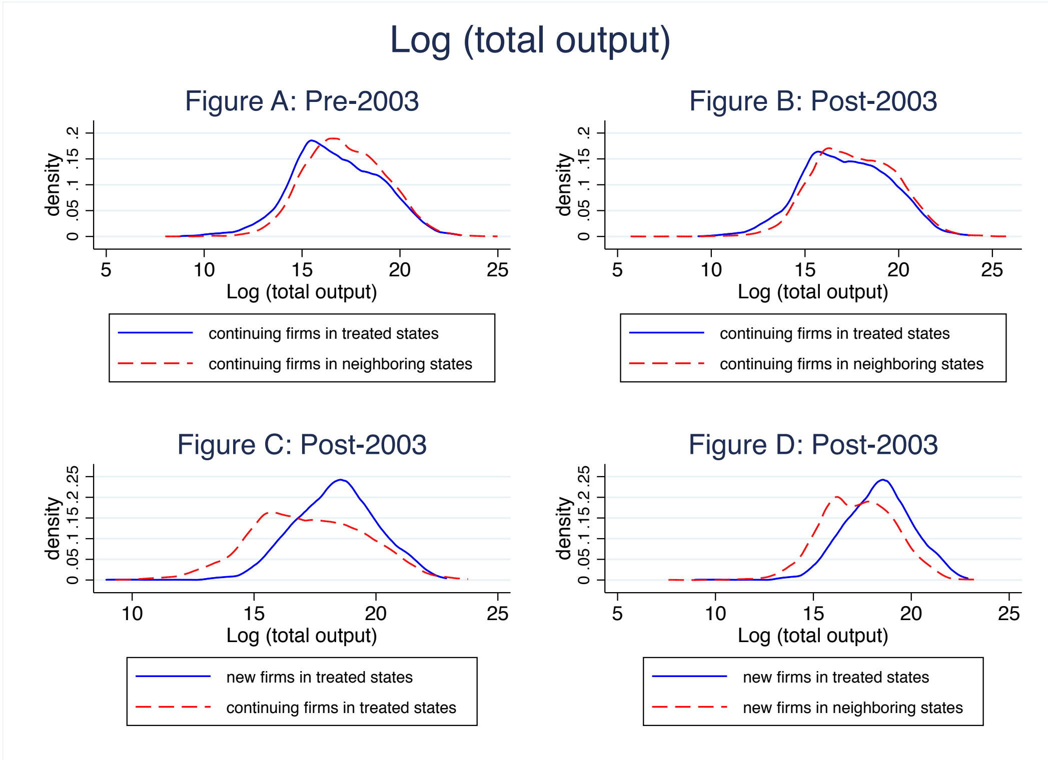
Notes: ASI state×industry data from 1999-2000 to 2007-08.

Figure 6: Kernel density of Log (employed)



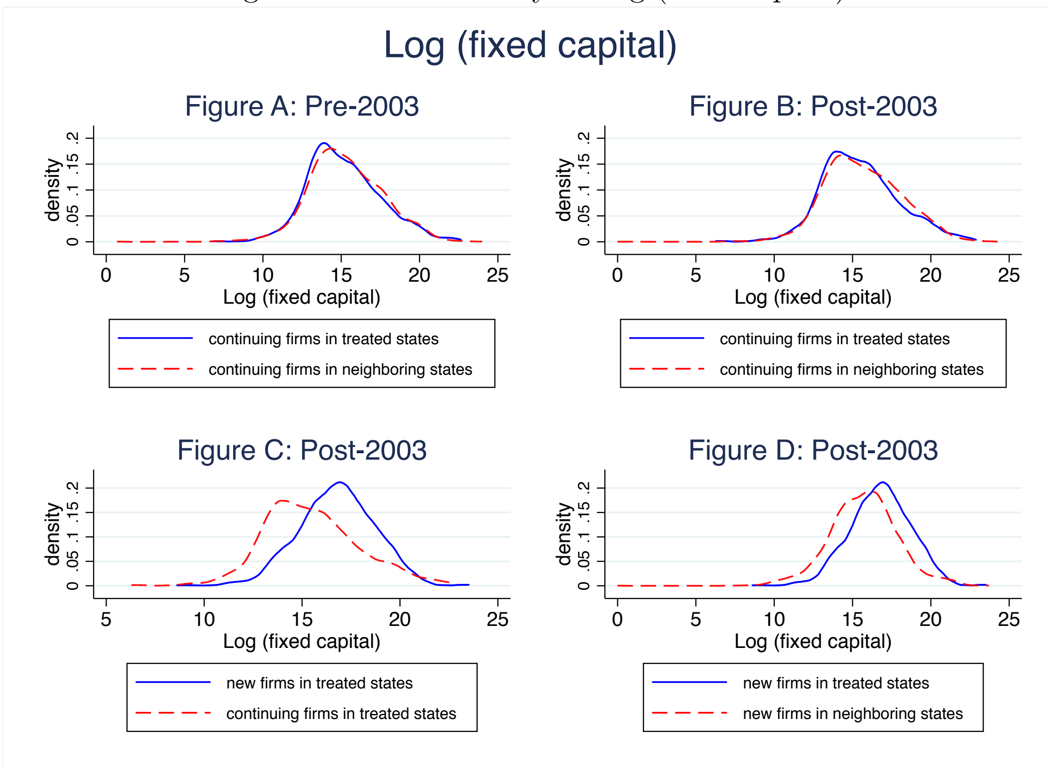
Notes: ASI firm-level data from 1999-2000 to 2007-08.

Figure 7: Kernel density of Log (total output)



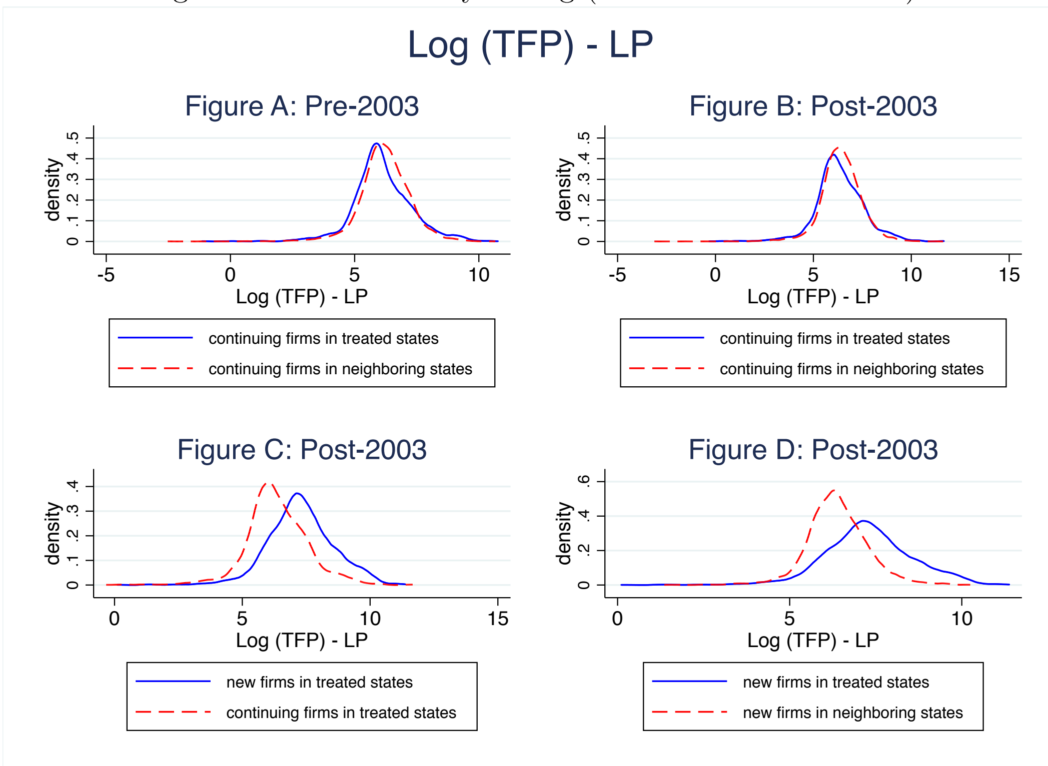
Notes: ASI firm-level data from 1999-2000 to 2007-08.

Figure 8: Kernel density of Log (fixed capital)



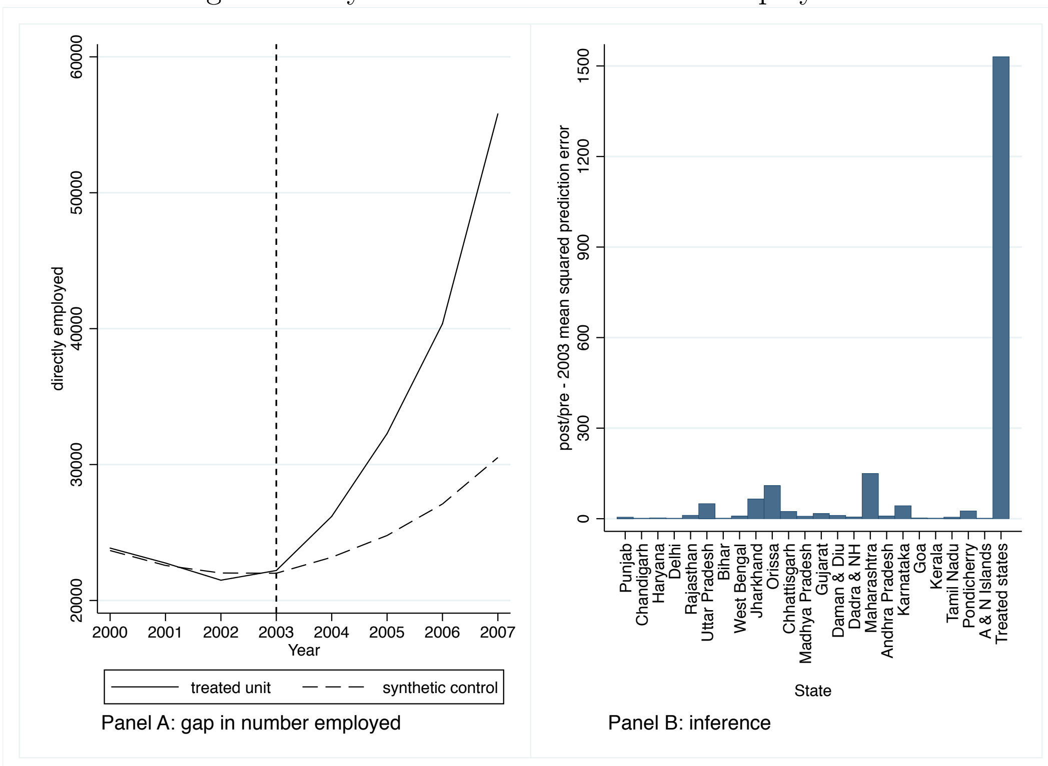
Notes: ASI firm-level data from 1999-2000 to 2007-08.

Figure 9: Kernel density of Log (TFP-Levinsohn-Petrin)



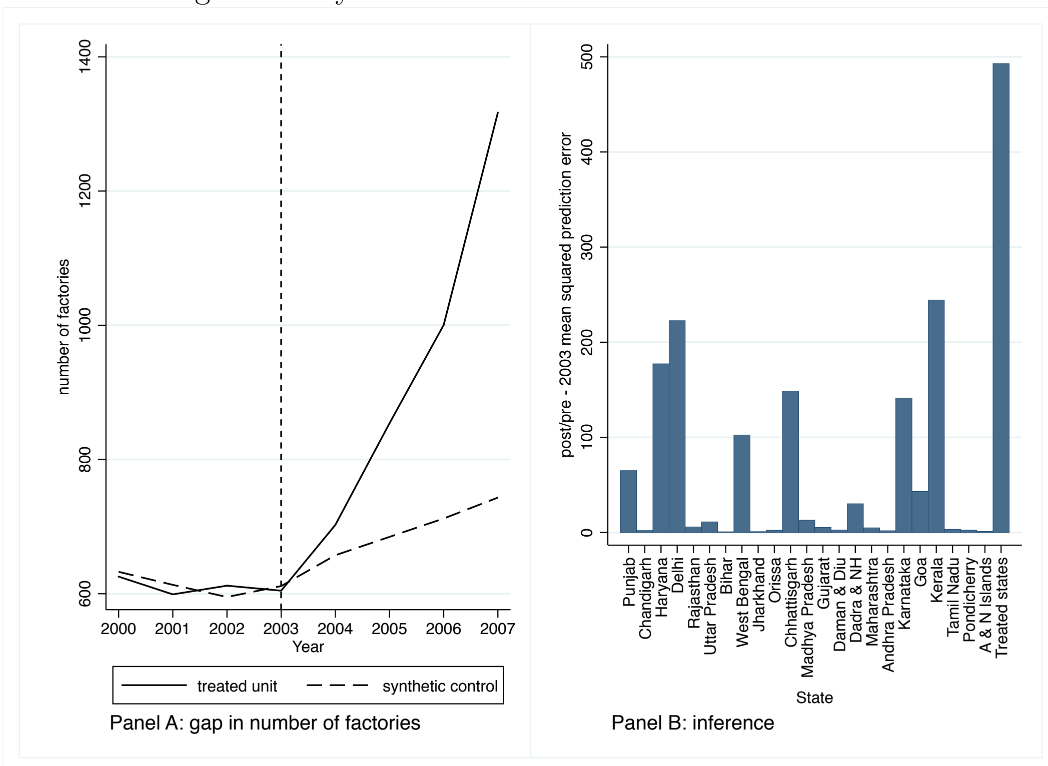
Notes: ASI firm-level data from 1999-2000 to 2007-08.

Figure 10: Synthetic Control Method - employment



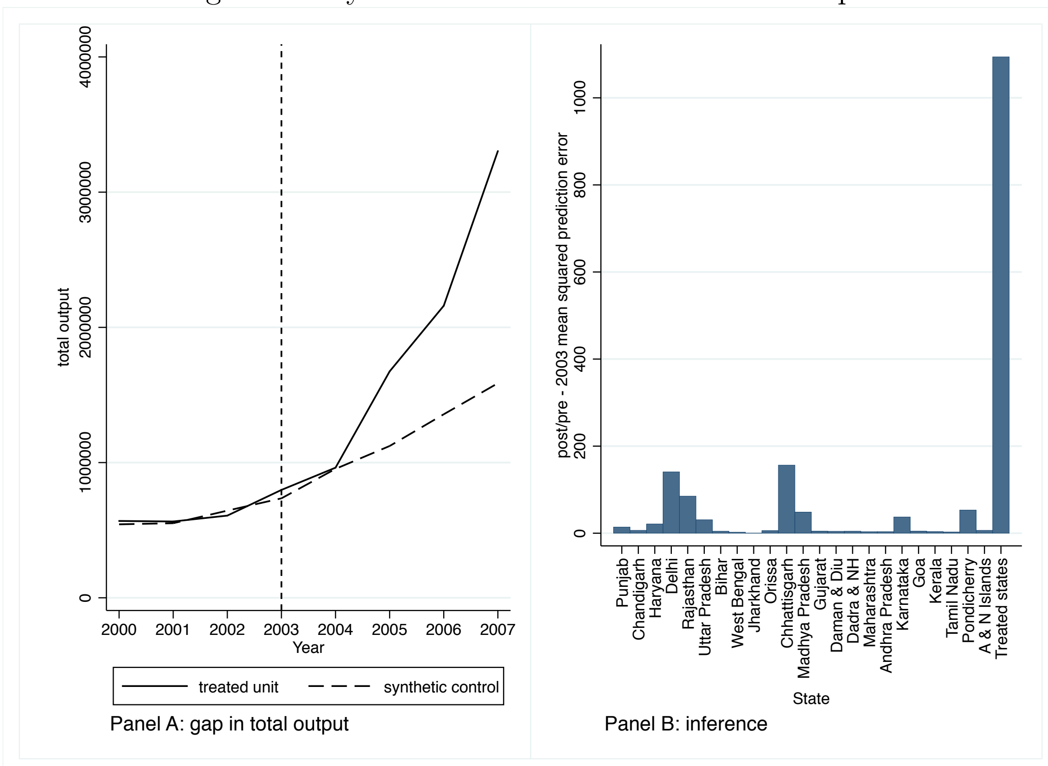
Notes: Panel A shows the gaps in employment between the treated and the synthetic control unit while Panel B shows a histogram for inference.

Figure 11: Synthetic Control Method - total factories



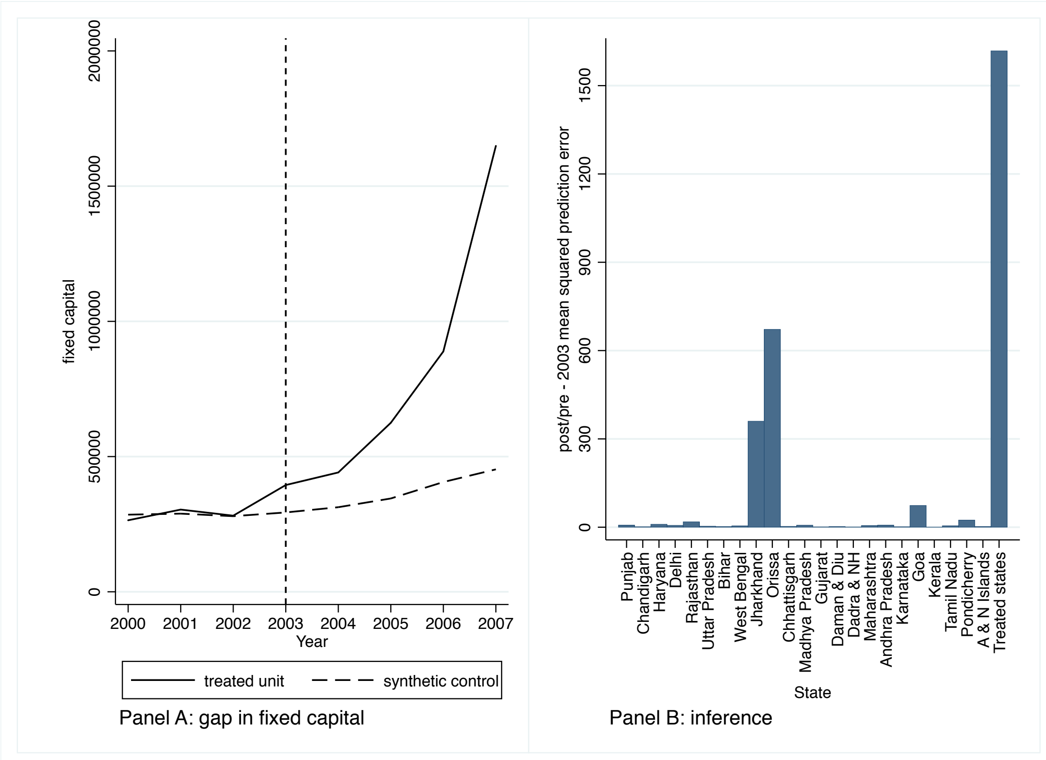
Notes: Panel A shows the gaps in total factories between the treated and the synthetic control unit while Panel B shows a histogram for inference.

Figure 12: Synthetic Control Method - total output



Notes: Panel A shows the gaps in total output between the treated and the synthetic control unit while Panel B shows a histogram for inference. Total output is measured in '00,000 Rs.

Figure 13: Synthetic Control Method - fixed capital



Notes: Panel A shows the gaps in fixed capital between the treated and the synthetic control unit while Panel B shows a histogram for inference. Fixed capital is measured in ‘00,000 Rs.

Appendix

Levinsohn-Petrin methodology

Assume that a firm i in industry j at time t has a Cobb-Douglas production function

$$y_{ijt} = \alpha + \beta_l(l_{jt}) + \beta_p(p_{jt}) + \beta_m(m_{jt}) + \beta_k(k_{jt}) + \omega_{ijt} + \varepsilon_{ijt}$$

where y is output, l is labor, p is power and electricity expenditure, and m is expenditure on raw materials (all variables in logarithms). The simultaneity problem arises because firms observe their own productivity ω_{ijt}, before choosing their inputs of power, labor and other raw materials. However, this is not observable to the econometrician. Levinsohn and Petrin (2003) use raw material expenditure (m_{ijt}) as a proxy for the unobserved productivity shock. They show that if these raw material inputs are monotonic in the firm's productivity at all levels of capital, then it can be inverted to express productivity in terms of capital and raw materials.

$$\omega_{ijt} = \omega_{jt}(m_{ijt}, k_{ijt})$$

This function can then be inserted into the equation above. Then the estimation takes place in two stages. In the first stage, a flexible functional form of capital and raw materials is included and the coefficients on l and p are estimated using semi-parametric techniques. The second stage uses GMM techniques to recover the coefficients on k and m.[51]

I use this method to estimate production function parameters separately for each 2 digit industry. Then, I use these estimates to construct firm-level productivity measures.

[51]For details, see Levinsohn and Petrin (2003).

Table A1: List of bordering districts

Himachal Pradesh	Uttarakhand	Uttar Pradesh	Haryana	Punjab
Sirmaur	Udham Singh Nagar	Pilibhit	Yamunanagar	Pathankot
Solan	Nainital	Bareilly	Ambala	Hoshiarpur
Bilaspur	Pauri	Rampur	Panchkula	Rupnagar
Una	Haridwar	Moradabad	-	SAS Nagar
Kangra	Dehradun	Bijnor	-	Gurdaspur
Chamba	-	Muzzafarnagar	-	-
-	-	Saharanpur	-	-

Table A2: State-level price index for neighboring states

States	1999-00	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08
Himachal Pradesh	1.00	1.04	1.09	1.15	1.16	1.19	1.23	1.26	1.30
Uttarakhand	1.00	1.02	1.06	1.13	1.16	1.18	1.25	1.35	1.42
Haryana	1.00	1.05	1.09	1.14	1.18	1.23	1.28	1.36	1.49
Punjab	1.00	1.07	1.11	1.13	1.15	1.18	1.25	1.33	1.47
Uttar Pradesh	1.00	1.01	1.03	1.09	1.13	1.17	1.24	1.32	1.40
Chandigarh	1.00	1.03	1.09	1.14	1.19	1.28	1.39	1.46	1.51
Delhi	1.00	1.04	1.09	1.11	1.17	1.21	1.26	1.32	1.38

Notes: These deflators have been calculated using the state GDP at current and constant prices from the RBI Handbook of Statistics on the Indian Economy. Base year: 1999-2000.

Table B1a: State weights in the synthetic ‘treated’ group

State	Weight
A & N Islands	0
Andhra Pradesh	0
Bihar	0
Chandigarh	0.777
Chhattisgarh	0.086
Dadra & NH	0.034
Daman & Diu	0
Delhi	0
Goa	0
Gujarat	0
Haryana	0
Jharkhand	0.018
Karnataka	0
Kerala	0
Madhya Pradesh	0
Maharashtra	0
Orissa	0
Pondicherry	0
Punjab	0
Rajasthan	0.085
Tamil Nadu	0
Uttar Pradesh	0
West Bengal	0

Table B1b: Directly employed predictor means

	Treated	Synthetic
Male	3706932	4049442
Female	3576693	3737691
Literacy rate	74.05	77.2577
Number of workers	3063249	3302985
Main workers	2143115	2463767
Marginal workers	920134	839217.7
Cultivators	1762493	1560680
Agricultural laborers	176927	532634
Household industrial workers	62483.5	85517.69
Percent SC	21.3	16.3341
Percent ST	3.5	6.394
Directly employed (2000-03)	22700	22756.61

Table B2a: State weights in the synthetic ‘treated’ group

State	Weight
A & N Islands	0.008
Andhra Pradesh	0
Bihar	0
Chandigarh	0.729
Chhattisgarh	0.139
Dadra & NH	0
Daman & Diu	0
Delhi	0
Goa	0
Gujarat	0
Haryana	0
Jharkhand	0.109
Karnataka	0
Kerala	0
Madhya Pradesh	0
Maharashtra	0
Orissa	0
Pondicherry	0
Punjab	0
Rajasthan	0.015
Tamil Nadu	0
Uttar Pradesh	0
West Bengal	0

Table B2b: Number of factories predictor means

	Treated	Synthetic
Male	3706932	3781787
Female	3576693	3558227
Literacy rate	74.05	76.0972
Number of workers	3063249	3053144
Main workers	2143115	2185583
Marginal workers	920134	867560.7
Cultivators	1762493	1222037
Agricultural laborers	176927	778797.7
Household industrial workers	62483.5	87648.31
Percent SC	21.3	15.9141
Percent ST	3.5	7.5423
Number of factories(2000-03)	612.1667	613.6527

Table B3a: State weights in the synthetic 'treated' group

State	Weight
A & N Islands	0
Andhra Pradesh	0
Bihar	0
Chandigarh	0.806
Chhattisgarh	0
Dadra & NH	0
Daman & Diu	0
Delhi	0
Goa	0
Gujarat	0
Haryana	0
Jharkhand	0.121
Karnataka	0
Kerala	0
Madhya Pradesh	0
Maharashtra	0
Orissa	0
Pondicherry	0
Punjab	0.015
Rajasthan	0.058
Tamil Nadu	0
Uttar Pradesh	0
West Bengal	0

Table B3b: Total output predictor means

	Treated	Synthetic
Male	3706932	3989818
Female	3576693	3639341
Literacy rate	74.05	77.0457
Number of workers	3063249	3012951
Main workers	2143115	2174101
Marginal workers	920134	838849.6
Cultivators	1762493	1265456
Agricultural laborers	176927	514184.3
Household industrial workers	62483.5	99604.07
Percent SC	21.3	16.9639
Percent ST	3.5	3.9131
Total output (2000-03)	580121	579834

Table B4a: State weights in the synthetic 'treated' group

State	Weight
A & N Islands	0
Andhra Pradesh	0
Bihar	0
Chandigarh	0.791
Chhattisgarh	0
Dadra & NH	0.045
Daman & Diu	0
Delhi	0
Goa	0
Gujarat	0
Haryana	0
Jharkhand	0.071
Karnataka	0
Kerala	0
Madhya Pradesh	0
Maharashtra	0
Orissa	0
Pondicherry	0
Punjab	0
Rajasthan	0.094
Tamil Nadu	0
Uttar Pradesh	0
West Bengal	0

Table B4b: Fixed capital predictor means

	Treated	Synthetic
Male	3706932	4157782
Female	3576693	3789372
Literacy rate	74.05	76.8581
Number of workers	3063249	3226216
Main workers	2143115	2361348
Marginal workers	920134	864868.4
Cultivators	1762493	1514791
Agricultural laborers	176927	440779.2
Household industrial workers	62483.5	97437
Percent SC	21.3	16.3826
Percent ST	3.5	5.8507
Fixed capital (2000-03)	283296.8	284601.6

Table B5: Corporate profits and corporate income

	(1) Log (profit)	(2) Log (profit)	(3) Log (income)	(4) Log (income)
*post*treat*	0.930***	1.045***	0.775***	0.804***
	(0.189)	(0.244)	(0.148)	(0.163)
Observations	2,021	2,021	2,404	2,404
R-squared	0.559	0.610	0.582	0.619
state FE	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes
3 digit industry FE	Yes	Yes	Yes	Yes
3 digit industry-year FE	No	Yes	No	Yes

Notes: The coefficient on the interaction term post*treat shows the treatment effect. Control group includes all neighboring states. Standard errors in parentheses are clustered at the state-year level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table B6: Regressions with state×industry fixed effects

VARIABLES	(1) Log (employed)	(2) Log (total factories)	(3) Log (total output)	(4) Log (fixed capital)	(5) Log (wage bill)
*post*treat*	0.427***	0.295***	0.537***	0.659***	0.430***
	(0.0761)	(0.0598)	(0.107)	(0.131)	(0.0786)
Observations	2,567	2,567	2,567	2,567	2,567
R-squared	0.953	0.969	0.939	0.942	0.949
state-industry FE	Yes	Yes	Yes	Yes	Yes
3 digit industry-year FE	Yes	Yes	Yes	Yes	Yes

Notes: The coefficient on the interaction term post*treat shows the treatment effect. Control group includes all neighboring states. Standard errors in parentheses are clustered at the state-year level. ***Significant at 1%, **Significant at 5%, *Significant at 10%.

Table B7: Inference using wild cluster bootstrap-t procedure

	Log(employment)	Log(employment)	Log(total factories)	Log(total factories)	Log(total output)	Log(total output)	Log(fixed capital)	Log(fixed capital)	Log(wage-bill)	Log(wage-bill)
Panel A: neighboring states										
*post*treat*	0.371	0.379	0.268	0.267	0.579	0.584	0.718	0.714	0.392	0.400
p-value	0	0.022	0.009	0.004	0	0.015	0.004	0.021	.0660	0.051
Observations	2567	2567	2567	2567	2567	2567	2567	2567	2567	2567
Clusters	7	7	7	7	7	7	7	7	7	7
Panel B: major states										
*post*treat*	0.427	0.443	0.310	0.311	0.561	0.577	0.711	0.728	0.412	0.427
p-value	0.039	0.031	0.008	0.008	0.036	0.067	0.182	0.211	0.053	0.070
Observations	8030	8030	8030	8030	8030	8030	8030	8030	8030	8030
Clusters	21	21	21	21	21	21	21	21	21	21
state FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3 digit industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3 digit industry-year FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: The coefficient on the interaction term post*treat shows the treatment effect. Panel A shows the regressions with the neighboring states as the control group whereas Panel B uses all major states as the control group. The p-values are calculated using the Cameron, Gelbach, and Miller (2008) wild cluster bootstrap-t procedure (999 replications).