Columbia University

Department of Economics
Discussion Paper Series

Location Based Tax Incentives: Evidence from India

Ritam Chaurey

Discussion Paper No.: 1213-12

Department of Economics
Columbia University
New York, NY 10027

January 2013

Location Based Tax Incentives: Evidence From India

Ritam Chaurey*

Department of Economics, Columbia University

December 2012

Abstract: This paper studies the impact of the Government of India financed "New Industrial Policy for Uttarakhand and Himachal Pradesh", whereby beginning 2003, new industrial units and existing units upon expansion were given 100% income tax and excise tax exemption. Using a difference-in-differences approach, I find a large increase in employment, number of factories, total output and fixed capital at the 3-digit industry level. Using firm level data, I find that the average employment, output, fixed capital, and additions to plant and machinery increased for existing firms as a result of this policy. Hence I show that the policy change affected both the intensive and extensive margins. I also look at heterogeneity by firm size and find that the policy led to employment and output increases for smaller firms. Finally, I use synthetic control methods (Abadie, Diamond, and Hainmueller (2010)) as a robustness check for the treatment effects of the policy change.

JEL Classification Numbers: R58, H25, O25.

*Address: Department of Economics, Columbia University, 1022 International Affairs Building, 420 West 118th Street, New York City, NY 10027, USA. Contact e-mail: *rc2556@columbia.edu*. I am grateful to Amit Khandelwal, Suresh Naidu and Eric Verhoogen 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, Supreet Kaur, Siddharth Kothari, Tao Li, Cristian Pop-Eleches, Raul Sanchez de la Sierra and participants at the Development Colloquium, Columbia University for helpful comments. All errors are my own.

1 Introduction

Economists have debated the efficacy of industrial policy to develop backward regions for a long time, with some viewing these policies as inefficient and a waste of taxpayers' money and others arguing for its effectiveness in solving market failures. In practice, however, state and local governments throughout the world have used "place-based" policies by offering tax incentives, subsidies, land grants, infrastructure and other benefits to firms in an effort to industrialize and generate employment and productivity at the local level.

Glaeser (2001) discusses the logic behind location-based tax incentives. These incentives attract firms that hire local labor and produce output, thereby generating both producer and consumer surplus. Entry of new firms are also beneficial to the existing local economy if they generate positive spillovers and agglomeration economies. Both these channels increase welfare at the local level. Local governments also use these tax incentives to maximize tax revenues in the long run, by giving tax breaks to firms up-front and impose taxes later on. Sometimes different firms are given differential tax benefits depending on their preference for the location (with firms with higher preference for the location receiving lesser benefits). Moretti (2011) uses spatial equilibrium analysis to explain the incidence and welfare implications of centrally financed subsidies to help disadvantaged areas. Whether such policies benefit the disadvantaged area depends on the elasticity of labor supply and housing supply.

Rigorous econometric studies of the causal effects of industrial policies have been difficult in the past because of the absence of clear identification strategies and data limitations. With the availability of more microdata, there has been a spate of recent papers looking at location based policies. However, the debate on the usefulness of such policies is still ongoing.

Most of the prior literature on place-based policies has focused on Empowerment Zones (EZ) in the US; neighborhoods receiving tax breaks and job subsidies. Neumark and Kolko (2010) use geocoded data on establishments to conclude that EZs in California have not led to employment growth and hence have been ineffective. However, Busso and Kline (2012) use confidential microdata and use rejected and future applicants to the EZ program as controls

to find that the program increased employment and wages inside the zones at moderate efficiency costs. Criscuolo, Martin, Overman, and van Reenen (2012) use an instrumental variable technique to find employment effects of the 'Regional Selective Assistance', in the UK. Most of the other recent literature on place based policies looks at the US or Europe. For example, Mayer, Mayneris, and Py (2012) look at the impact of French Enterprise Zones (ZFUs) on the location decisions of firms and find a positive impact albeit with displacement effects from the untreated areas to the treated areas. There is also some related literature looking at the impact of Special Economic Zones (SEZ) in China (Wang (2010)) and industrial clusters in India (Fernandes and Sharma (2012)).

This paper looks at a similar policy in a developing country context in India. Specifically, I look at the central government sponsored New Industrial Policy for the states of Uttarakhand and Himachal Pradesh. As part of the policy change, post-2003, all new industrial units or existing units on their substantial expansion, set up in industrial estates within the state, received fiscal benefits such as 100% excise duty exemption for 10 years and 100% income tax exemption for 5 years, from the date of commercial production. Other benefits included a capital investment subsidy on plant and machinery subject to an upper limit of Rs. 30 lakhs (approximately USD 60,000).

I look at this policy change using a differences-in-differences methodology and find a 37% increase in employment, 27% increase in the number of factories, 57% increase in total output and a 87% increase in fixed capital at the 3-digit industry level, as a result of the policy change. I also find a 10.3% increase in the number of workers employed, 11.4% increase in total output and a 7% increase in fixed capital (including depreciation) in incumbent firms. Existing firms also increase their additions to fixed capital (before depreciation) and additions to plant and machinery by 27.5% and 25.6% respectively. Further, I can also rule out relocation of firms from control states to treated states. There is also some evidence of heterogeneity in these effects based on firm size, and I find most of the employment and output effects concentrated in small firms (based on different cutoffs) and almost no effect on

larger firms. Finally, I use synthetic control methods (Abadie, Diamond, and Hainmueller (2010)) as a robustness check for the treatment effects of the policy change.

To my knowledge, this is the first paper to rigorously evaluate the impact of a central government sponsored policy to industrialize a backward area in a developing country, by providing tax incentives for new and existing firms. Since this policy change leads to growth in both the intensive and extensive margins without any evidence of relocation of firms from untreated to treated areas, this study informs policy makers about the efficacy of tax benefits to industrialize backward regions. In this regard, the results of this paper suggest a larger role for "industrial policy" to correct for market failures and regional economic disparities especially in developing countries.

This paper is organized as follows. The next section presents the background for the study and the details of the policy. Then, Section 3 discusses the empirical strategy, Section 4 describes the data and the results are discussed in Section 5. Section 6 discusses the robustness of these results and finally Section 7, provides some conclusions.

2 Background and Policy Details

Himachal Pradesh and Uttarakhand are two states in the north of India (see Figure 1). In November 2000, the Himalayan and adjoining northwestern districts of Uttar Pradesh were consolidated to form the state of Uttarakhand. After the formation of Uttarakhand, it was placed in the list of "special category" states [1] 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 18.6% of India's total area. They are predominantly covered by hilly areas

[1] 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.

and forests[2]. Industrialization was considered a policy challenge in the two states, owing to the topography.

Beginning 2003, the Government of India (central government), to attract investments in the industrial sector 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 not less than 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 would be 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 industries 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.30 lakh (approximately USD 60,000).

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.

The thrust sector included floriculture, medicinal herbs and aromatic herbs processing, honey, horticulture and agro based industries (such as sauces, ketchup, etc., fruit juices, fruit pulp, preserved fruits and vegetables, jams, jellies, vegetable juices, puree, pickles etc., processing of fresh fruits and vegetables including packaging), food processing industry[3], sugar and its by-products, silk and silk products, wool and wool products, woven fabrics (excisable garments), sports goods, paper and paper products, pharmaceutical products, information and communication technology industry and computer hardware call centers,

[2]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

[3]See Appendix for Negative list - a list of industries that did not receive these benefits

bottling of mineral water, eco-tourism (hotels, resorts, spa, entertainment/amusement parks and ropeways), industrial gases, handicrafts and non-timber forest based industries.

Importantly, a few months after the policy (in June 2003) was initiated, the Government of India issued a notification[4] 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 along with adding new industrial estates. This practically blurred the differential treatment accorded to the thrust sector industries, because there was 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 upon substantial expansion in these states prior to the above dates would continue to be eligible for these benefits as per the above rules.

3 Empirical Strategy

In this paper I intend to empirically test whether the centrally sponsored location specific tax incentives led to an increase in employment, output, fixed capital and the number of factories in the treated states. The empirical strategy uses the 2003 policy change that provided tax incentives to 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 states to a set of control units - either neighboring states, all major states or bordering districts. Since this is a centrally sponsored scheme, this paper is not concerned with different states competing against each other by providing differential

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

benefits to attract firms.

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 affected by the policy change, an empirical strategy comparing firms/industries across districts within the treated states would not be possible.[5] 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. Hence, I believe that this is the best empirical strategy to answer this question in the given context. However, 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 the trends in employment, number of factories, total output and fixed capital were similar across the treated and control states, only diverging post treatment. The pre-treatment trends look parallel and provide visual support to the use of difference-in-differences strategy in this context to estimate the causal effect of the policy change.

More specifically, I run two main types of regressions (DID specifications) to estimate the causal effect of the policy change. First, I run *state-3-digit industry level* regressions of the form:

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

[5]we 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

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 or fixed capital 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. The coefficient β, on the interaction term $post_t * treat_s$ 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 workers, number of illiterates in the state, that I interact with a time dummy for each year as control variables. These regressions are similar to the area level employment and number of plants regressions in Criscuolo, Martin, Overman, and van Reenen (2012). The regressions with employment, output and fixed capital therefore look at 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 as it is an aggregate that takes in to account both entry and exit of firms.

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

$$y_{isjt} = \alpha_i + \delta_s + \lambda_{jt} + \beta(post_t * treat_s) + \gamma(X_{isjt}) + \epsilon_{isjt} \tag{2}$$

where i,s,j,t indexes firm, state, 4-digit industry and time respectively and y_{isjt} represents an outcome variable such as employment, output, fixed capital, additions to fixed capital and additions to plant and machinery at the firm level. I also use age and age-squared as controls in the regressions. Since the policy (post) dummy equals one for all years post 2003

and zero otherwise, the inclusion of firm fixed effects removes the effect of entrants post 2003. Hence, this regression looks at the impact of the policy change on the outcome measures of incumbent firms. Hence, this regression can be interpreted as the "intensive margin" of the policy change.

Displacement 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 bordering states to reopen in the treated state, I take two approaches. First, I look at trends in the number of operational factories in the treated states and the neighboring control states (Figure 5). Then, I run a regression with the number of closed firms[6] at the state-industry level as the dependent variable to look at the differential impact on firm closing across the treated and control states, before and after the policy change. Another related concern might be that the policy induces spillovers in the neighboring states. Since the magnitude of the treatment effects are similar whether we look at only neighboring states or all major states together as the control group, it is reasonable to conclude that the results are not being driven by spillovers. Such externalities might be substantial especially for the bordering districts regressions. To check whether the results are being driven by spillovers, I first 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 (Appendix Table 9). I also perform robustness checks using the synthetic control method where the synthetic control group is formed using a weighted average of all the states to best match the treated states. For all outcome variables (shown below), the synthetic control group is composed of states that are not neighboring the treated states. All these results from various specifications suggest that the results in the paper are not driven by spillovers.

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

4 Data

I use two data sets: (i) Annual Survey of Industries (ASI) state-3-digit industry panel (from 1999-2000 to 2007-08) and (ii) ASI firm level panel (1998-99 to 2007-08) to look at the effect of the policy change.

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 with the aid of power and 20 or more workers without the aid of power. ASI also covers bidi and cigar units, registered under the Bidi and Cigar Workers (Conditions of Employment) Act, 1966. 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 frame is divided into census (surveyed every year) and sample (sampled every few years) sectors. However, the definition of the two sectors have changed from time to time. Five industrially backward states (Manipur, Meghalaya, Nagaland, Tripura and Andaman and Nicobar Islands) 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 in to account the changes in the sampling frame, I run weighted regressions at the firm level with the sampling weights provided by ASI.

I restrict the sample to the major states and union territories of India as covered by the ASI, and do not include Jammu & Kashmir or the states in the north-east namely Assam, Manipur, Meghalaya, Nagaland and Tripura.

For the state-industry level regressions I use the ASI state-3-digit industry panel. Each observation is at the state-industry-year level. Table 1 shows some 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 start out with smaller industrial employment,

number of factories, total output and fixed capital 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."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[7]. 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 compared to any other group.

5 Results

I begin by looking at the results for the difference-in-differences regressions at the state industry level for different outcome variables (Tables 3 to 6) and then look at employment at the firm level (Table 7) and finally discuss the firm size heterogeneity results (Table 8a and 8b).

[7]See Harrison, Martin, and Nataraj (2011) and Bollard, Klenow, and Sharma (2011)

5.1 State-industry results

Having shown the trends in the outcome variables with the raw data in Figures 2 and 3, I now report the regression results. Table 3 reports employment regressions at the state-industry level. 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.

Column 1 includes state, year and 3-digit industry fixed effects. The state fixed effects control for time invariant state characteristics like the area and topography of the state, the year fixed effects control for any 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 have grown in India over the last decade, and the industry-year fixed effects controls for these changes. Column 3 includes 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, number of workers and number of illiterates and interact them with a time dummy for each year as control variables. Column 4 includes the specification of Column 2 with time varying controls. 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.

The dependent variable in Table 3 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 increases by around 37

- 42% in 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%.

Table 4 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 in to 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 is between 27% and 29%. The corresponding estimates from Panel B show a 31% increase in the average number of factories. Table 4 confirms that the policy change led to a large increase in the number of new firms coming in to the treated states.

Table 5 shows the results for log of total output at the state-industry level. The effect of the policy change on total output is even larger than the effect on total employment ranging between 58% and 64%. Results for log of fixed capital are shown in Table 6. The results show an increase of 87% for fixed capital and part of this can be attributed to the "substantial expansion" clause (although I cannot explicitly test this) where existing firms would need to increase their investment of fixed capital by at least 25% to avail the tax incentives. Further, capital investment subsidies were also provided in these two states post 2003 to both new and existing firms, leading to this massive increase in fixed capital at the industry level.

Figure 4 plots the estimated coefficients over time along with standard errors at the 95% confidence interval level. 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* (dummy 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.

There can be some concern that firms are closing down in the neighboring states to reopen in the treated states to take advantage of the tax incentives. 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 states and reopen in the treated states, there should be a drop in the number of operational factories in the neighboring states. Figure 5 plots the trends in operational factories and there is no evidence that factories in the neighboring states closed down to relocate in the treated states. I also run a regression specification similar to column 4 in the tables with state and industry-year fixed effects and time varying controls with log of number of closed firms as the dependent variable. The coefficient on *post*treat* is -0.0139 with a standard error of 0.245. Hence I see no differential impact of the policy change on number of closed firms across neighboring states. This suggests that there is no evidence for relocation of firms from neighboring states in to treated states.

5.2 Firm level results

Firm level regressions are reported in Table 7. I restrict the sample to only open firms. 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, with the idea being that the districts on the side of Uttarakhand and Himachal Pradesh receive the tax incentives as compared to the districts on the other side of the border in the neighboring states, which then act as a suitable control group. I essentially compare outcomes for firms across bordering districts before and after the policy change. These regressions provide a strict test on the identification strategy and provide credible support to my results from using firms in neighboring states as the control group.

In Table 7, 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). This is because all the specifications control for firm fixed effects, and this removes the effect of new entrants. Columns 1 and 2 show that the mean employment at the firm level increases by 7-10%. The results in columns 3 and 4 with firms in bordering districts as the control group also shows an increase in mean employment by around 11%. Output at the firm level increases by 11.4% (Column 2). The magnitude of the effect on output is larger (23.7%) in the bordering districts regression but are roughly in the same range.

I also run regressions with fixed capital, additions to fixed capital and additions to plant and machinery as the outcome variables. Fixed capital includes depreciation and additions to fixed capital, plant and machinery are measures of actual additions before depreciation. Fixed capital shows an increase of around 7-8%. Actual additions to fixed capital increases by around 28% and additions to plant and machinery increases by around 26%. The results from the bordering districts regressions are very similar to these magnitudes. These results provide suggestive evidence that existing firms took advantage of the "substantial expansion" clause and increased investment on fixed capital by more than 25% to take get the tax benefits. In this respect, these results confirm that the policy was successful in incentivizing firms to invest more in plant and machinery.

Restricting the sample to districts along the border of treated and control states may bias the results because of spillovers. 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 9 in the Appendix. For existing firms in the treated districts as compared to control districts, employment went up by 15% and output by 16% after 2003. This lends support for the bordering districts regression specification in Table 7.

5.3 Firm size heterogeneity

Criscuolo, Martin, Overman, and van Reenen (2012), find that the effect of the "Regional Selective Assistance" in the UK on employment is confined to smaller firms with almost no effect on larger firms. To check for such treatment effect heterogeneity by size, I run regressions to look at the treatment effects of the policy on firms with different cutoffs in Tables 8a and 8b.

Table 8a looks at 3 size cutoffs - below 50 workers, between 50 and 100 workers and above 100 workers. Table 8b reports a different set of cutoffs - below 40 workers, between 40 and 80 workers and more than 80 workers. An interesting pattern emerges from the table. Mean employment and output in Table 8a goes up for smaller firms (less than 50 workers) by 9% and 10% respectively, with negligible effect for the large firms (more than 100 workers). This is the case even in Table 8b where employment and output increase by 8.6% and 11.3% for firms with less than 40 workers. There is no effect (negative and insignificant effect) for firms with more than 80 workers. These results are also broadly similar to Bronzini and Iachini (2011) and Lach (2002), who find that R & D incentives have an effect on smaller firms but not on larger firms in Italy and Israel respectively. However, when we look at additions to fixed capital or additions to plant and machinery, large firms show a large and significant increase whereas smaller firms show a statistically insignificant increase.

Most employment and output growth is concentrated with the smaller firms. A plausible rationale for this heterogeneous effect might be that the smaller firms are credit constrained and the tax benefits relax the constraint, thereby allowing them to expand output and hire more labor. These results also conform with Banerjee and Duflo (2012), where they look at a directed lending program in India and find that many firms were credit constrained and actually used the credit to expand production rather than as a substitute for other forms of credit.

There is also some suggestive evidence that large firms increase their investments in plant and machinery to take advantage of the tax incentives (substantial expansion clause) but

do not increase output or employment. For example in Table 8a, firms with more than 100 workers increase their additions to plant and machinery and fixed capital by 39.8% and 32.4% respectively but show negligible effect for output and employment. The same pattern holds true in Table 8b for firms with over 80 workers. However, the results are not strong enough to conclude that larger firms are gaming the system.

6 Robustness

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), so that it 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. I, then perform "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 placebos. This is similar to permutation tests, and essentially tests whether the effect of the treatment is much larger in the treated states as compared to a randomly chosen placebo. If the effect is of a similar magnitude between the treated and a randomly chosen placebo, then it is not possible to conclude with any confidence that the treatment effect is real.

First, I aggregate the data from the two treated states into one treated unit, as the

synthetic method works when there is one treated unit and a bunch of control units to choose from. In this analysis, I use the ASI state-industry level panel data and use the total aggregated numbers over all industries in the state in a year. 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 variable of interest is directly employed, 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 6, Panel A graphically shows the treatment effect of the 2003 policy change on 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 post 2003. Appendix Table 10a shows the weights on the states in the donor pool, to provide a good synthetic control. Only Chandigarh, Chhattisgarh, Dadra and Nagar Haveli, Jharkhand and Rajasthan have positive weights, with the rest of the states with 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. Appendix Table 10b shows the values for various predictors of the outcome variable for both the treated and the synthetic control groups. Figure 6, Panel B is used to conduct the inference test. It 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. Evidently, the 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 that of the treated states

would be $1/24 = 0.042$. Hence, it is significant at the 5% level.

Figure 7, Panel A shows the synthetic control method graph for the outcome variable - number of factories in all industries at the state level. Appendix Table 11a 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). Appendix Table 11b shows the mean of the predictors used for number of factories. I conduct a similar inference test in Figure 7, Panel B and plot the ratio of post-intervention to pre-intervention MSPE. The ratio is largest for the treated state and hence the probability of obtaining a ratio as large as is $1/24 = 0.042$.

Figure 8 and 9 show the synthetic control method graphs for total output and fixed capital. 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.

7 Conclusion

Many argue that industrial policy is a waste of taxpayers' money and governments use it to prevent competition and pick winners. Others argue that industrial policy might be needed to fix market failures such as spillovers. Whether location based tax incentives work and help in industrialization and employment generation at the local level, thus is largely an empirical question. Previously there was a paucity of rigorous studies to look at the causal effects of such policies because of clear identification strategies and data issues. 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 largely been understudied in developing countries where regional economic disparities can be large and labor mobility might be low.

In this paper, I look at such a policy in India - the New Industrial Policy for the states

of Uttarakhand and Himachal Pradesh, that provided income tax and excise tax exemption to new industrial units and existing industrial units on their expansion, starting 2003. I use state-industry data and this state-year variation to find that employment, number of firms, total output and fixed capital increase at the 3-digit industry level. I find no evidence of relocation of firms from neighboring states in to the treated states. I also use firm level data, to find that the average employment, output, fixed capital, and additions to plant and machinery increased for existing firms as a result of this policy. Hence, I am able to show that the policy change affected both the intensive (growth by existing firms) and extensive (entry of new firms) margins. I also look at heterogeneity by firm size and find that the policy led to employment and output increases for smaller firms and had almost no impact on larger firms despite increases in their fixed capital investments. These results are consistent with smaller firms being credit constrained and the tax benefits easing these constraints and suggestive of large firms gaming the system.

However, 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 (2011)). Whether or not such policies have a lasting impact (for example, agglomeration economies)[8] 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 scope of this paper. With more data available in the following years, this will be an exciting avenue for future research.

[8]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.

Banerjee, A., and E. Duflo (2012): "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program," *mimeo MIT*.

Bollard, A., P. Klenow, and G. Sharma (2011): "India's Mysterious Manufacturing Miracle," *mimeo U Missouri*.

Bronzini, R., and E. Iachini (2011): "Are incentives for R&D effective? Evidence from a regression discontinuity approach," *Bank of Italy, Economic Research and International Relations Area Economic working papers 791*.

Busso, Matias, J. G., and P. Kline (2012): "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, forthcoming.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (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*.

Fernandes, A., and G. Sharma (2012): "Determinants of Clusters in Indian Manufacturing: The Role of Infrastructure, Governance, Education and Industrial Policy," *mimeo U Missouri*.

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

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.

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.

Kline, P., and E. Moretti (2011): "Local Economic Development, Agglomeration Economies and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *mimeo UC Berkeley*.

Lach, S. (2002): "Do R&D Subsidies Stimulate or Displace Private R&D? Evidence from Israel," *Journal of Industrial Economics*, 50(4), 369–90.

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.

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.

Wang, J. (2010): "The Economic Impact of Special Economic Zones: Evidence from Chinese Municipalities," *mimeo HKUST*.

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

Variable	State	Time period	Observations	Mean	Median
Number employed	All states	Pre-2003	3604	3962.88	801
		Post-2003	6179	4162.30	822
	Treated	Pre-2003	231	589.68	274
		Post-2003	424	834.16	377
	Neighboring states	Pre-2003	709	3153.96	1057
		Post-2003	1203	3554.21	1046
	All major states	Pre-2003	2746	4952.17	1294
		Post-2003	4630	5237.32	1258
Number of factories	All states	Pre-2003	3604	107.64	26
		Post-2003	6179	112.77	26
	Treated	Pre-2003	231	15.98	9
		Post-2003	424	21.15	12
	Neighboring states	Pre-2003	709	103.90	41
		Post-2003	1203	110.14	41
	All major states	Pre-2003	2746	134.31	40
		Post-2003	4630	141.98	40
Total output	All states	Pre-2003	3604	8379.36	1420.45
(in '000,000 Rs.)		Post-2003	6179	16268.90	2471.30
	Treated	Pre-2003	231	1506.81	486.70
		Post-2003	424	4197.37	1392.50
	Neighboring states	Pre-2003	709	7322.53	2238.90
		Post-2003	1203	13135.86	3370.40
	All major states	Pre-2003	2746	10381.42	2340.95
		Post-2003	4630	20213.52	3634.80
Fixed capital	All states	Pre-2003	3604	3541.41	394.80
(in '000,000 Rs.)		Post-2003	6179	5103.75	538.60
	Treated	Pre-2003	231	735.84	108.60
		Post-2003	424	1886.03	295.40
	Neighboring states	Pre-2003	709	2423.11	536.60
		Post-2003	1203	3268.24	605.40
	All major states	Pre-2003	2746	4430.63	648.90
		Post-2003	4630	6398.87	869.75

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, West Bengal. Observations here are state-3 digit-year observations.

Table 2: Summary statistics at the firm level

Variable	State	Time period	Observations	Mean	Median
Number employed	All states	Pre-2003	93276	41.50	11
		Post-2003	195581	42.64	12
	Treated	Pre-2003	1915	39.63	9
		Post-2003	5222	43.43	14
	Neighboring states	Pre-2003	18192	32.74	10
		Post-2003	39710	36.03	11
	All major states	Pre-2003	84879	41.55	11
		Post-2003	176605	42.62	12
Total output	All states	Pre-2003	88126	92.80	9.16
(in '000,000 Rs.)		Post-2003	187796	173.00	13.50
	Treated	Pre-2003	1743	110.00	14.40
		Post-2003	4733	230.00	42.60
	Neighboring states	Pre-2003	17234	80.00	10.40
		Post-2003	38348	136.00	14.40
	All major states	Pre-2003	80680	91.90	8.89
		Post-2003	170818	170.00	12.80
Fixed capital	All states	Pre-2003	96718	35.80	1.37
(in '000,000 Rs.)		Post-2003	206686	49.50	2.03
	Treated	Pre-2003	1961	47.10	1.73
		Post-2003	5461	92.80	5.99
	Neighboring states	Pre-2003	18741	24.10	1.47
		Post-2003	42523	30.90	1.94
	All major states	Pre-2003	87607	35.90	1.33
		Post-2003	185860	49.40	1.95

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, West Bengal. Observations here are firm-year observations.

Table 3: Dependent variable: 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	X	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 4: Dependent variable: 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 5: Dependent variable: 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	2567	2567	2567	2567
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 6: Dependent variable: 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	8030	8030	8030	8030
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 7: 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 (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 (addition 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 (addition 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 8a: Firm size heterogeneity

	(<)50 workers	**50-100 workers**	**(>)100 workers**
Log (directly employed)			
post*treat	0.0902*	0.0422*	-0.0126
	(0.0505)	(0.0246)	(0.0280)
Observations	47,162	5,709	10,450
R-squared	0.907	0.828	0.899
Log (total output)			
*post*treat*	0.102*	-0.167	-0.00602
	(0.0517)	(0.162)	(0.0487)
Observations	42,133	5,495	12,745
R-squared	0.962	0.969	0.985
Log (addition to fixed capital)			
*post*treat*	0.203	0.362	0.324*
	(0.168)	(0.382)	(0.163)
Observations	35,764	5,279	11,585
R-squared	0.837	0.888	0.903
Log (addition to plant and machinery)			
*post*treat*	0.109	0.381	0.398**
	(0.157)	(0.372)	(0.198)
Observations	26,441	4,774	10,209
R-squared	0.856	0.860	0.850
Control group	Neighboring states	Neighboring states	Neighboring states
firm FE	Yes	Yes	Yes
year FE	Yes	Yes	Yes
4 digit industry FE	Yes	Yes	Yes
Age Controls	Yes	Yes	Yes

Table 8b: Firm size heterogeneity

	(<)40 workers	40-80 workers	(>)80 workers
Log (directly employed)			
*post*treat*	0.0860*	0.0238	-0.0119
	(0.0484)	(0.0492)	(0.0211)
Observations	44,666	6,352	12,168
R-squared	0.904	0.834	0.904
Log (total output)			
*post*treat*	0.113**	-0.0626	-0.0390
	(0.0482)	(0.124)	(0.0370)
Observations	39,777	6,049	14,417
R-squared	0.962	0.974	0.984
Log (addition to fixed capital)			
*post*treat*	0.169	0.734**	0.340**
	(0.169)	(0.347)	(0.165)
Observations	33,476	5,843	13,185
R-squared	0.835	0.899	0.897
Log (addition to plant and machinery)			
*post*treat*	0.128	0.650	0.405**
	(0.170)	(0.392)	(0.193)
Observations	24,452	5,206	11,669
R-squared	0.859	0.872	0.842
Control group	Neighboring states	Neighboring states	Neighboring states
firm FE	Yes	Yes	Yes
year FE	Yes	Yes	Yes
4 digit industry FE	Yes	Yes	Yes
Age Controls	Yes	Yes	Yes

Figure 1: Map of India

SOURCE: mapsofindia.com

Figure 2: Trends in employment and number of factories

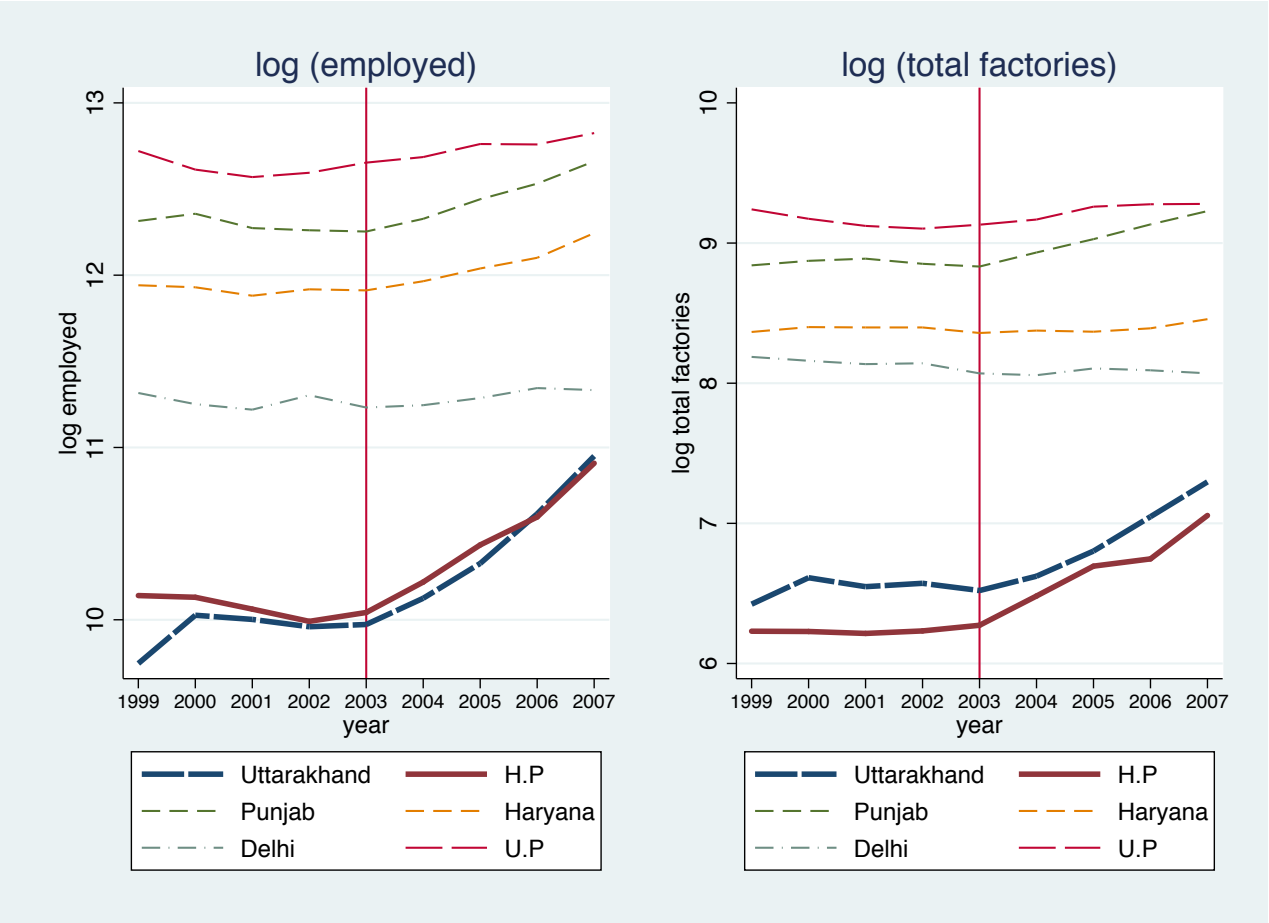


Figure 3: Trends in total output and fixed capital

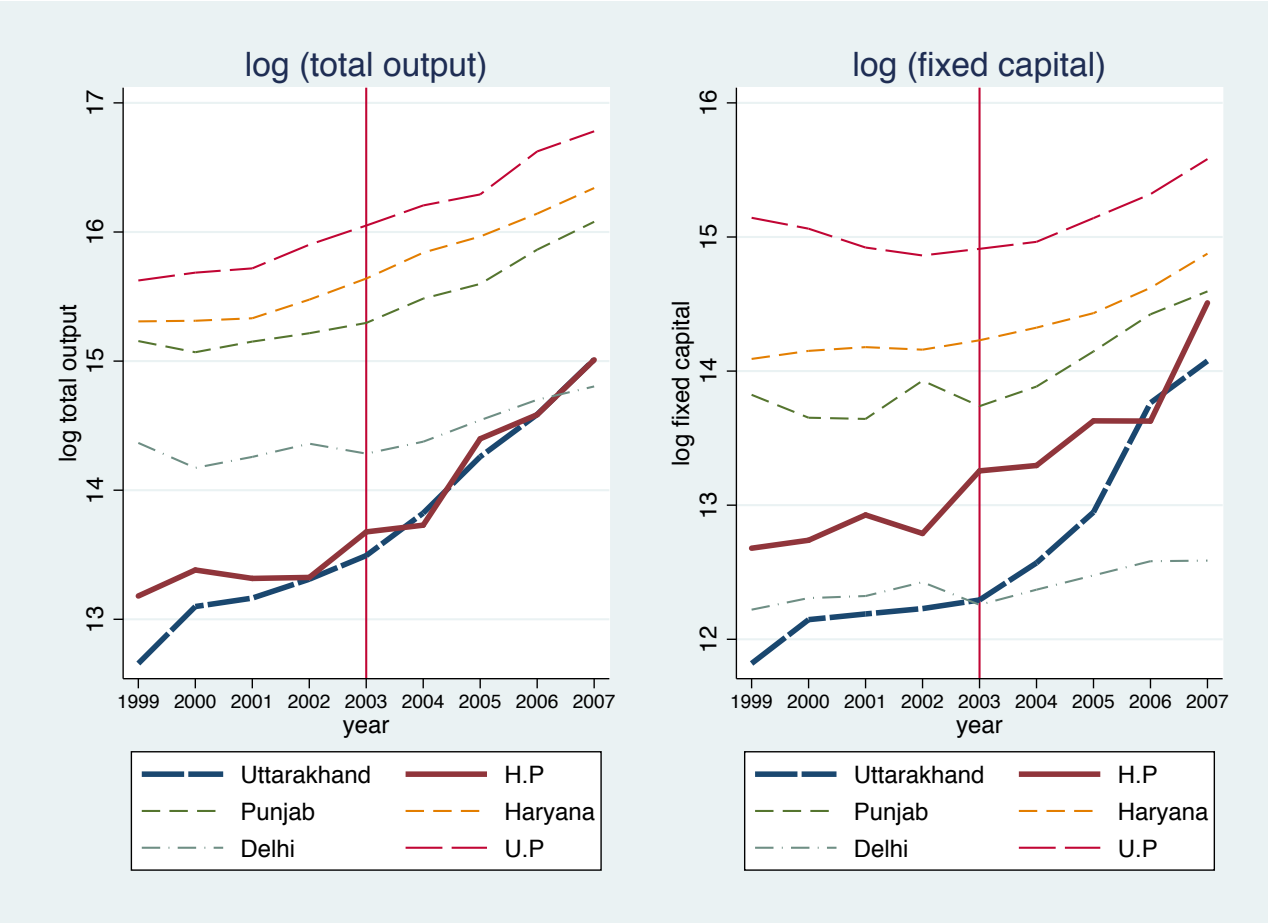
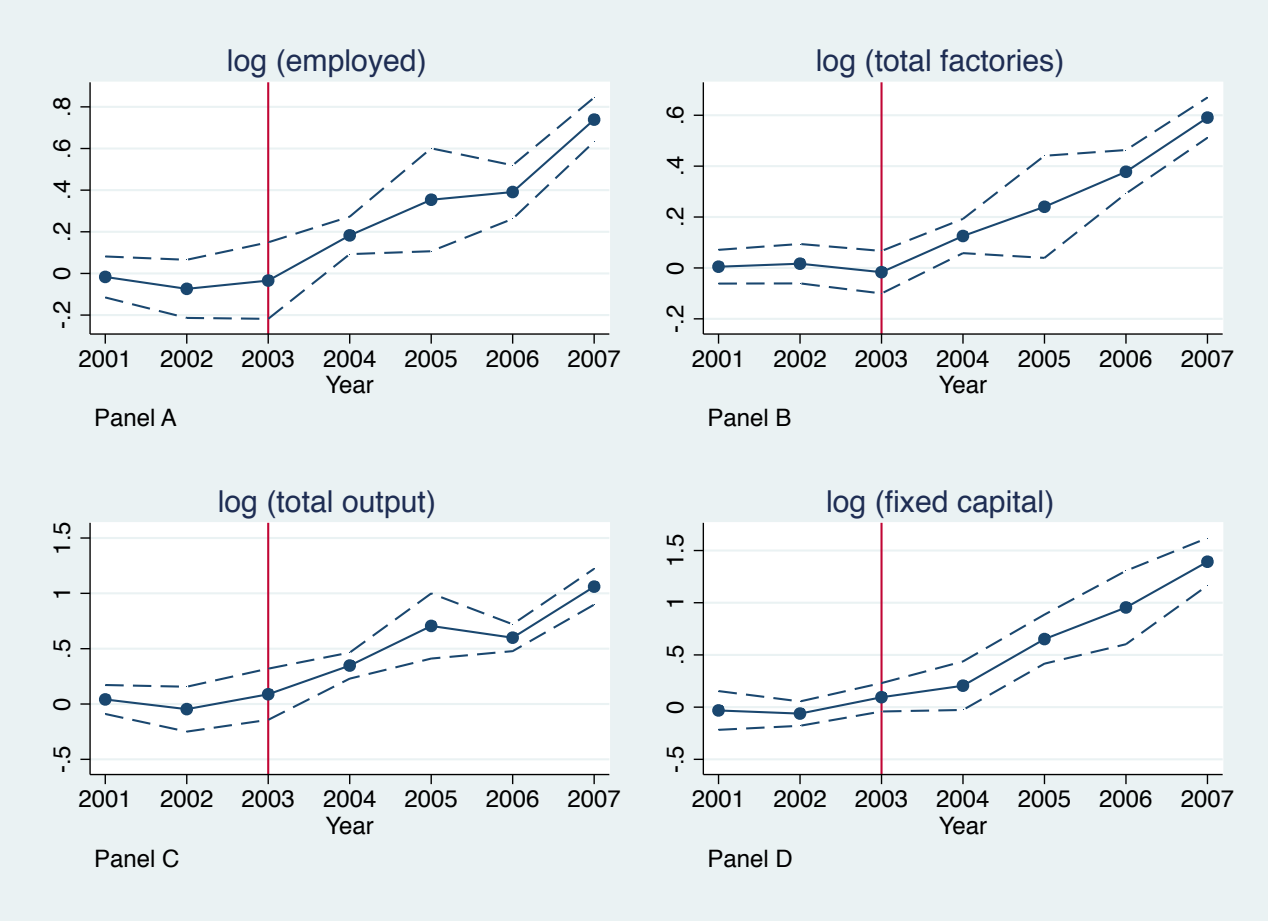


Figure 4: Estimated coefficient graphs



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

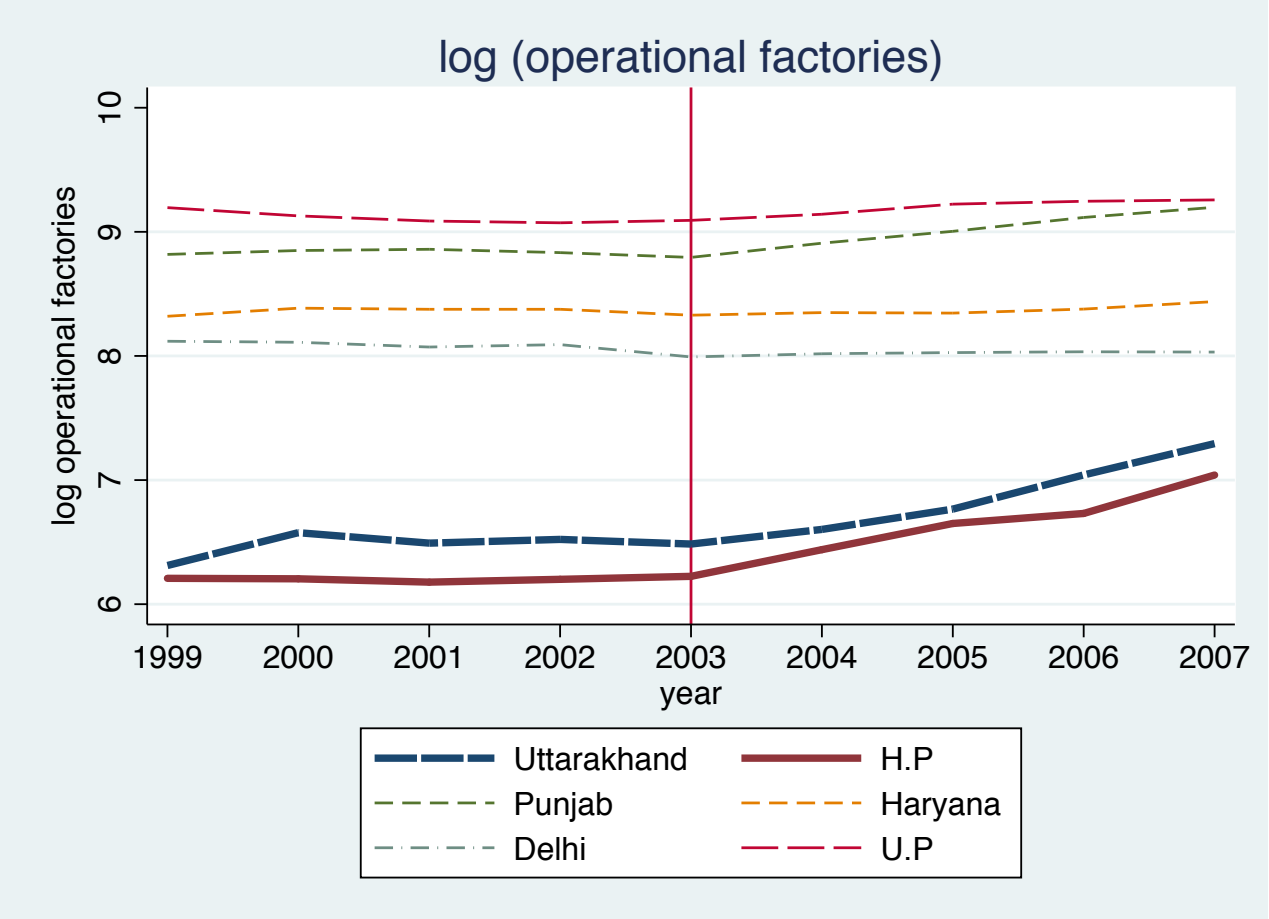
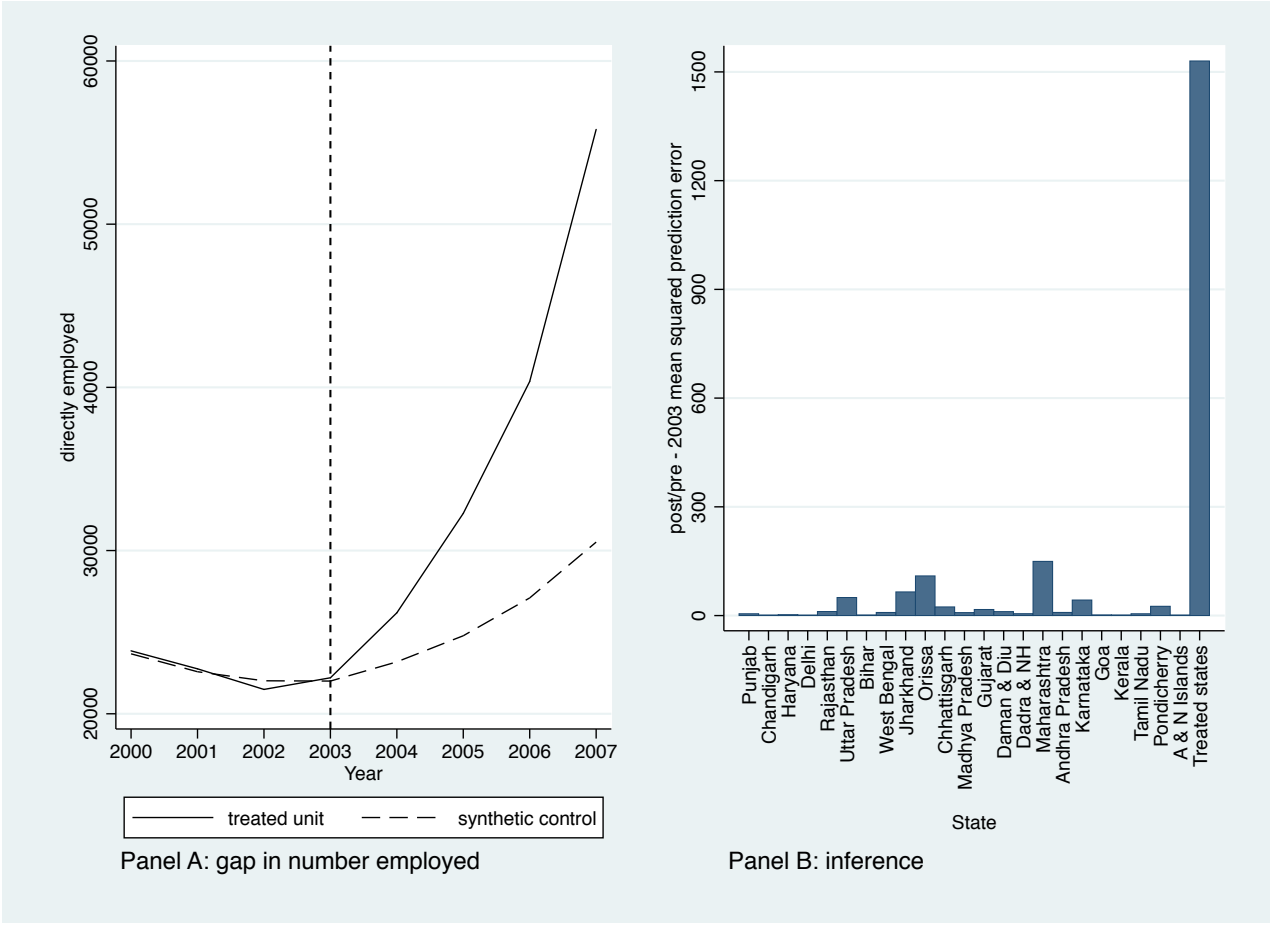
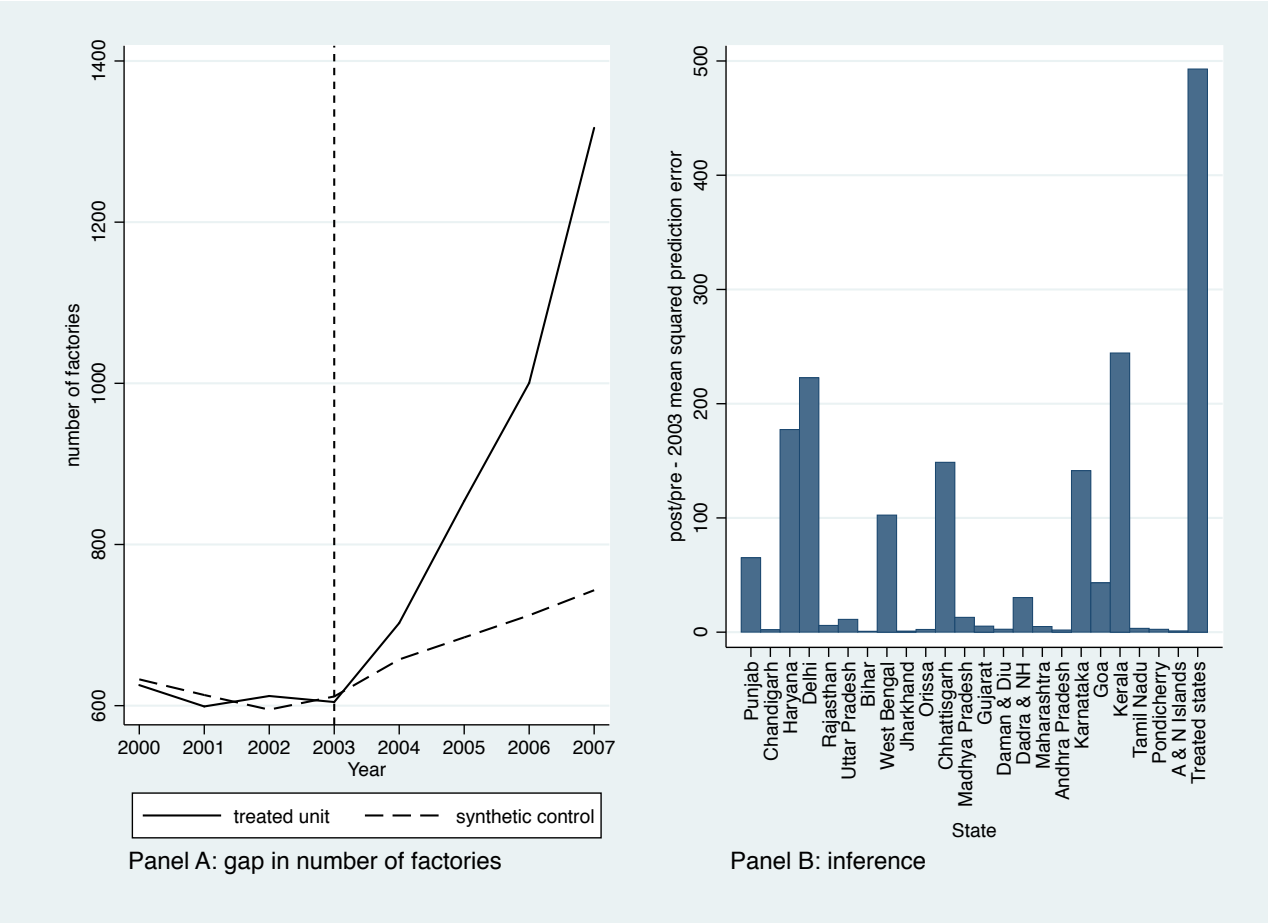


Figure 6: Synthetic Control Method - employment



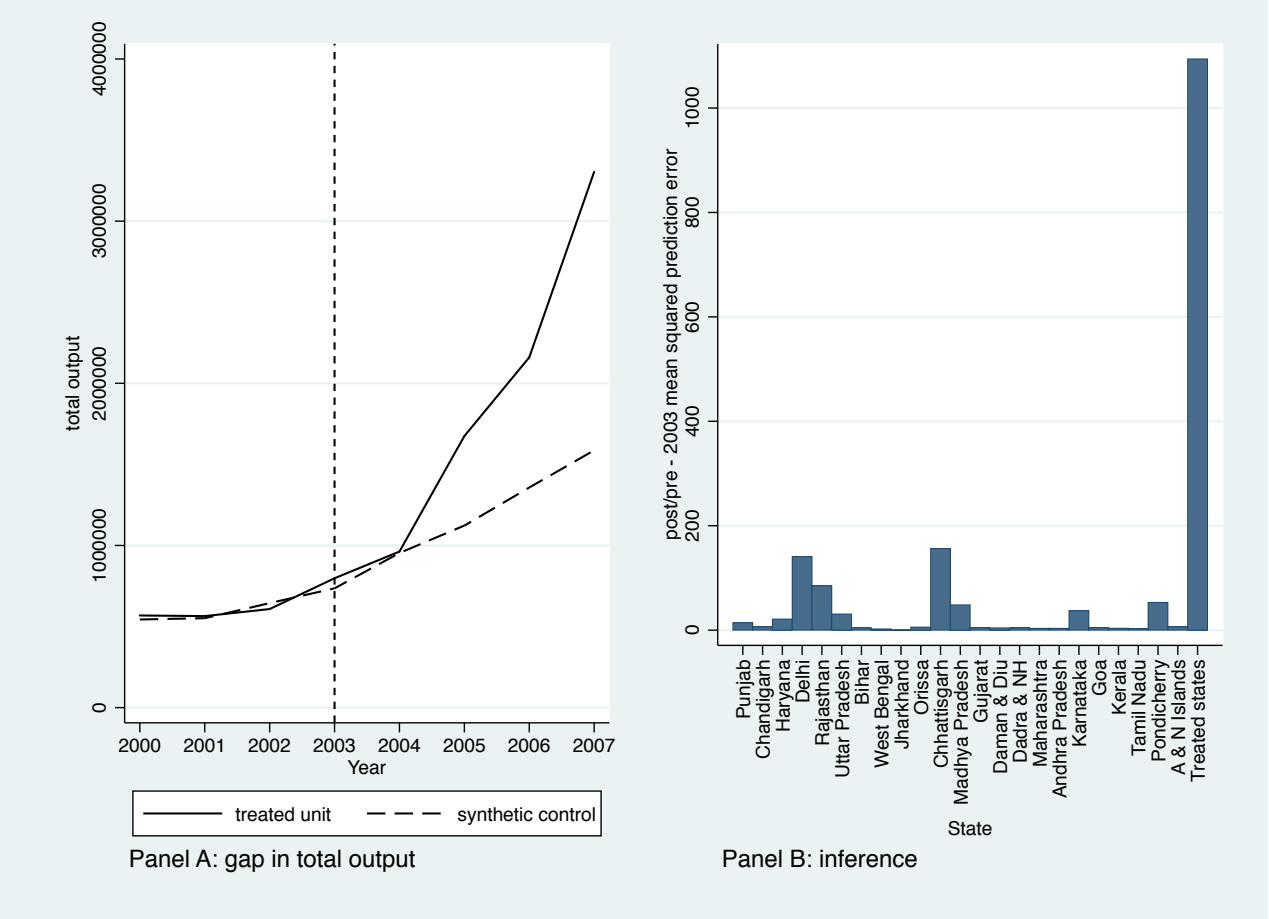
Notes: Panel A shows the gaps in employment between the treated and the synthetic control unit while Panel B shows a histogram for inference

Figure 7: Synthetic Control Method - total factories



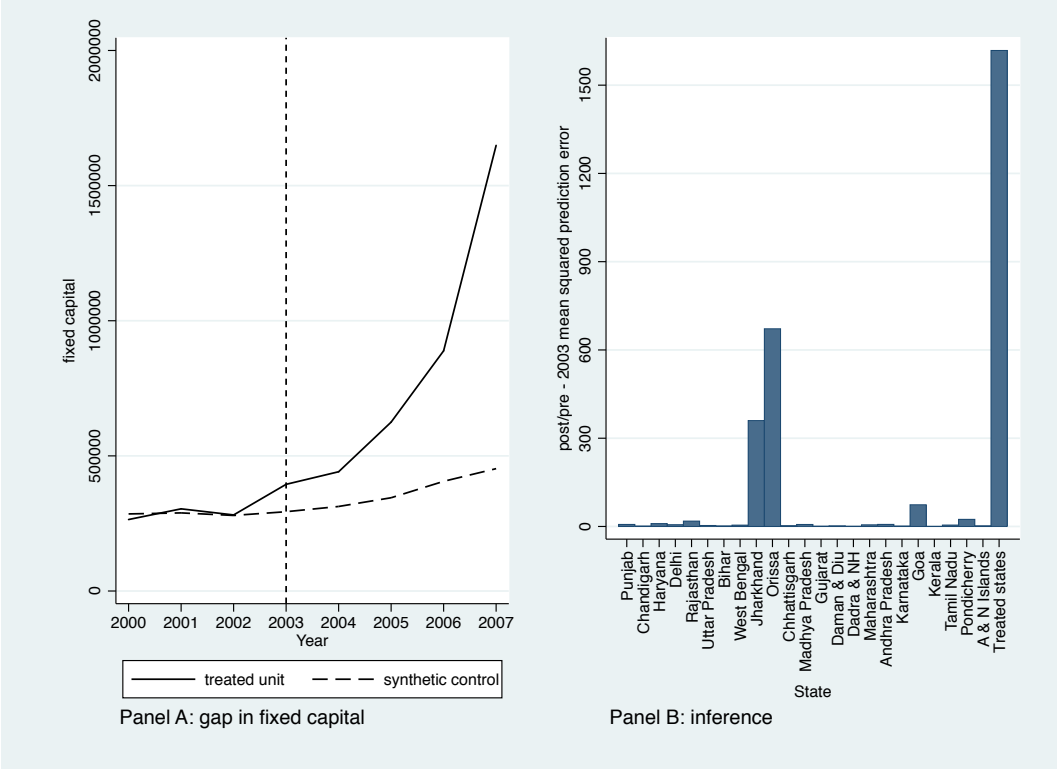
Notes: Panel A shows the gaps in total factories between the treated and the synthetic control unit while Panel B shows a histogram for inference

Figure 8: 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 9: 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.

8 Appendix

Negative list for Uttarakhand and Himachal Pradesh[9]

1. Tobacco and tobacco products including cigarettes and pan masala
2. Thermal Power Plant(coal/oil based)
3. Coal washeries/dry coal processing
4. Inorganic Chemicals excluding medicinal grade oxygen, medicinal grade hydrogen peroxide, compressed air
5. Organic chemicals excluding Provitamins/vitamins, Hormones, Glycosides, sugars
6. Tanning and dyeing extracts, tanins and their derivatives, dyes, colours, paints and varnishes; putty, fillers and other mastics; inks
7. Marble and mineral substances not classified elsewhere
8. Flour mills/rice mill
9. Foundries using coal
10. Minerals fuels, mineral oils and products of their distillation; Bituminous substances: mineral waxes
11. Synthetic rubber products
12. Cement clinkers and asbestos, raw including fibre
13. Explosive (including industrial explosives, detonators and fuses, fireworks, matches, propellant powders etc.)
14. Mineral or chemical fertilizers

[9]Source: Department of Industrial Policy & Promotion, Govt. of India

15. Insecticides, fungicides, herbicides and pesticides (basic manufacture and formulation)

16. Fibre glass & articles thereof

17. Manufacture of pulp - wood pulp, mechanical or chemical (including dissolving pulp)

18. Branded aerated water/soft drinks (non-fruit based)

19. Paper

19.1 Writing or printing paper, etc.

19.2 Paper or paperboard, etc.

19.3 Maplitho paper, etc.

19.4 Newsprint, in rolls or sheets

19.5 Craft paper, etc.

19.6 Sanitary towels, etc.

19.7 Cigarette paper

19.8 Grease-proof paper

19.9 Toilet or facial tissue, etc.

19.10 Paper & paper board, laminated internally with bitumen, tar or asphalt

19.11 Carbon or similar copying paper

19.12 Products consisting of sheets of paper or paperboard, impregnated, coated or covered with plastics, etc.

19.13 Paper and paperboard, coated impregnated or covered with wax, etc.

20. Plastics and articles thereof

Table 9: Firm level regression - testing for spillovers

	(1)	(2)	(3)
VARIABLES	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 10a: 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 10b: 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 11a: 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 11b: 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 12a: 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 12b: 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 13a: 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 13b: 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