Discussion Paper Series

2010 – 04

Department of Economics
Royal Holloway College
University of London
Egham TW20 0EX

Electricity Provision and Industrial Development: Evidence from India

Juan Pablo Rud*

August 2010

Abstract

I investigate the effect of electricity provision on industrialization using a panel of Indian states from 1965-1984. To address the endogeneity of investment in electrification, I use the introduction of a new agricultural technology intensive in irrigation (the Green Revolution) as a natural experiment. As electric pumpsets are used to provide farmers with cheap irrigation water, I use the uneven availability of groundwater at the start of the Green Revolution to predict divergence in the expansion of the electricity network and, ultimately, to quantify the effect of electrification on industrial outcomes. I present a series of tests to show that the electrification channel remains the most important one among alternative explanations that could link groundwater availability to industrialization directly or through other means than electrification. Overall, one standard deviation in my measure of electrification explains around 14 percent of the variation in manufacturing output across states in India.

1 Introduction

The adequate supply of infrastructure goods is increasingly acknowledged as one key factor in generating a conducive environment for industrial and economic development. The World Bank currently directs 35% of their lending portfolio to infrastructure projects with the idea that "infrastructure has a central role in the development agenda and is a major contributor to growth, poverty reduction and achievement of the Millennium Development Goals" (World Bank (2005)). However, average expenditure on infrastructure for developing countries is only around 3% of GDP. Among developing countries, South Asia and Sub-Saharan Africa perform particularly poorly in a range of infrastructure goods, such as water and sanitation, telecoms and electricity access. For example, around 43% of the population have access to the electricity

*Department of Economics, Royal Holloway, University of London. I thank Laura Abramovsky, Oriana Bandiera, Robin Burgess, Irma Clots-Figueres, Esteban Jaimovich, Guy Michaels, Patrick Nolen, Steve Redding and Jon Temple and seminar participants at LSE-EOPP, NEUDC '08, 23rd Annual Congress of the European Economic Association, Universidad de San Andres, Oxford University, Royal Holloway University of London, University of Essex, IV CEPR Summer School in Applied IO, XII SMYE 2007, First "Riccardo Faini" Conference on Development Economics and UNU-WIDER for helpful comments and suggestions.

network in Southern Asia, compared to at least 90% in Latin America, Eastern Europe and East Asia.

Access to reliable sources of electrification is, according to World Bank investment climate surveys, one the greatest obstacles to industrial development in India and the rest of Southern Asia. In particular, India's poor infrastructure in general, and the power sector in particular, is seen as one of the reasons behind India's slow export growth during the 1990s, limiting their comparative advantage in labor intensive products (World Bank (2000)). Public agencies, economic journalists and sector analysts[1] are among those who point to the electricity sector's poor performance as heavily affecting the growth and development possibilities of the Indian economy since independence, to the point that "the poor quality of electricity has been the single greatest deterrent to India's economic growth and development" (US DoE (2003)). Yet the history of Indian electrification is not one of uniform failure. For example, around 10% of villages on average were electrified by 1965, with some states like Tamil Nadu close to 50% and others such as Assam or West Bengal less than 3%. In 1984, the average increased to over 75% and some states like Punjab achieved full village electrification.

I use this variation across regions and over time within India to investigate the effect of electricity provision on industrial development, by examining a panel of Indian states between 1965 and 1984. India's federal political organization gives each state full responsibility for creating, expanding and managing the electricity network, from electricity generation to retail. This provides significant variation in the extent of the physical network, allowing me to test whether the gap in infrastructure provision is associated with unequal industrialization levels.

Assessing and quantifying the impact of investment in electrification on economic outcomes in the long run is a difficult task, since it is hard to determine the underlying driving force. The resulting endogeneity concerns arise because of reverse causality and unobserved state characteristics (such as the business or political climate) that might explain why some states are better prepared to provide a better electricity network than others. In such cases, OLS estimates would be biased. To address this problem, I investigate a source of exogenous variation to predict the expansion of the electricity network in order to estimate its impact on industrial development in an instrumental variable (IV) strategy. In particular, I use the start of the

[1] For example, "(..) electricity is unusable for industry and of such poor quality that power surges wreck equipment." (The Economist, 22/9/2005). See also TERI (1999).

Green Revolution – an agricultural technology intensive in irrigation introduced in India in the mid 60s – as a natural experiment. The successful introduction of the new High Yield Variety (HYV) seeds depended, among other factors, on the provision of timely irrigation. As electric pumpsets were used to provide cheap irrigation water, the uneven availability of groundwater across states – as measured by the thickness of the watertable– was an important determinant of electrification. For that reason, I use the time-varying effects of the initial groundwater availability as an instrument for the expansion of the electricity network.

This paper contributes to the growing literature on infrastructure, reinforcing the idea that geographic characteristics can be used to explain differential investments in expensive infrastructure projects, and address its endogenous placement. For example, Duflo and Pande (2007) instrument dams placement in India using information on rivers' gradients and Dinkelman (2010) instruments electrification in South Africa with land gradients.

The main concern regarding my empirical strategy is the validity of the exclusion restriction, i.e. that groundwater or HYV adoption could affect industrial outcomes through channels other than electrification. For example, greater rural incomes can increase the demand for manufacturing goods, positive shocks to agricultural productivity release cheap labor to be employed in the manufacturing sector or groundwater can be used directly by industries. Similarly, the political features of a state that determine a successful process of HYV adoption and electrification could also explain industrial growth. I present a discussion of alternative channels and a series of checks that show that, when considering these, the electrification link remains the most important of all.

Results show that an increase in one standard deviation in the measure of electrification increases around 14% the manufacturing output for a state at the mean of the distribution. Electrification is also associated with more factories and output among smaller firms. Magnitudes are substantial and underline the potential economic benefits of investing in the expansion of the electricity network.

The paper is organized as follows. Section 2 provides some background of the electricity sector in India and its relation to the Green Revolution and presents the data. Section 3 discusses the IV strategy, presenting the first stage and reduced form evidence. I then show OLS and IV results of linking electricity indicators to various manufacturing outcomes and discuss alternative explanations in Section 4. In Section 5 I show that results are robust to different measures of groundwater and electrification. Section 6 concludes.

2 Background and Data

The empirical contribution of this paper (to document the effect of electrification in industrial development) draws both from the organization of the electricity sector in India and from the uneven introduction of an agricultural technology that increased the demand for powered irrigation. In particular, I argue that groundwater availability has been a major source of electrified irrigation since the start of the Green Revolution and, thus, an important factor in explaining rural electrification across states. I then explore whether this mainly agricultural phenomenon could have benefited untargeted industrial users (in both rural and urban areas) by looking at a panel of 15 Indian states between 1965 and 1984[2].

The starting point is the Green Revolution, i.e. the introduction of high-yield varieties (HYV) in the mid 1960s, after persistent food shortages in newly independent India. Farming was mainly for subsistence, rainfed and characterized by the use of primitive techniques (Chakravarti (1972)). The Third Five Year Plan laid out by the Planning Commission, and starting in 1961, set ambitious targets in terms of agricultural production so that "the economy will become self-sufficient in the supply of foodgrains". After the failure of some programmes such as the Intensive Agricultural District Programme (IADP) and the Intensive Agricultural Areas Programme (IAAP), in 1966-7 the introduction of the High Yielding Variety Seed Programme provided the expected breakthrough, by providing farmers with hybrid seeds scientifically adapted to India's domestic conditions. The Green Revolution – in order to become a 'revolution' – depended on a series of factors, especially the adequate and timely supply of water. It is no surprise that states like Haryana and Punjab – both at the forefront of the Green Revolution – achieved a share of irrigated area around 60% and 90% respectively in the mid-1980s, while other relatively rich states like Gujarat and West Bengal were around the 25% mark. Irrigation could be provided by using canals supplying water from a dam or reservoir or the installation of deep or shallow powered tubewells, i.e. diesel or electricity operated pumps that lift the water from the water table. Bharadway (1990) notes that "the rate of increase of

[2]The data includes all Indian states according to the state configuration in 1966. I only exclude Jammu & Kashmir and smaller Union Territories because there is no consistent data available.

irrigation by wells/tubewells was higher than that by canals, and accelerated remarkably during the period 1969-1980 when there was a spurt in private tubewells, especially in the late sixties". Moreover, according to the National Commission on Agriculture, Ministry of Agriculture and Irrigation of India (1976), electricity-powered pumpsets were the cheapest option for farmers, up to half the cost of operating diesel pumpsets. It follows that places where groundwater was available and abundant were better suited for a cost-effective utilization of electric pumpsets that would provide the timely and precise irrigation needed by HYV seeds.

The data on groundwater availability and adoption of HYV seeds come from the district level dataset "India Agriculture and Climate Dataset", compiled by Sanghi, Kumar, McKinsey, Jr. (1998) for the years 1957-58 to 1986-87. I use information on the net proportion of area cropped with HYV seeds. For groundwater availability, the dataset provides a set of dummies according to the aquifers depth[3]. I aggregate this data at the state level by calculating the proportion of districts per state with the aquifers thicker than 150 mts[4]. This measure is the most important for my empirical strategy, since groundwater availability has been the major source of electrified irrigation after the Green Revolution (see McGuirk and Mundlak (1991), and Foster and Rosenzweig (2008))[5].

I subsequently investigate the process of electrification and industrial growth across states. There is a constitutional arrangement in India that ensures state independence from the central government in designing electricity policies. The Electricity Supply Act created in 1948 fully vertically integrated State Electricity Boards (SEB), in charge of coordinating electricity generation, transmission and distribution and commercialization at the state level, whose Board members are appointed by the state government (Electricity Supply Act, Chapter III, 5.2). An important aspect of the power sector organization is that during the period analyzed, there was negligible cross-state trade of electricity: on average, 97% of available electricity was generated within each state. The electricity data comes from "Public Electricity Supply - All India Statistics", published annually by the Central Electricity Authority between 1950 and 1985. The

[3]The measure of groundwater depth is divided in three: up to 100 mts deep, between a 100-150 mts and above 150 mts.

[4]Note that the dataset does not contain information on the states of Assam and Kerala. I compute their measure of groundwater availability using information on net groundwater availability in 2004 estimated by Central Ground Water Board (2006). In particular, I look at the correlation between my measure of groundwater and the 2004 measure and find a correlation of around 69%. I then compute values for each missing state such that the correlation remains at the same level. I discuss other groundwater measures below.

[5]Note that I use a measure of thickness of the water table at the outset of the Green Revolution. Subsequent use (and depletion) of the available groundwater could be endogenous to HYV adoption and electrification.

measure of rural electrification is the number of agricultural units connected to the electricity network per 1000 people[6].

Data for industrial development and performance indicators are taken from different publications from the Department of Statistics, Ministry of Planning, Government of India. Manufacturing output comes from"Estimates of State Domestic Product" and measures of the stock of fixed capital, value added and number of factories come from the "Annual Survey of Industries"[7]. The data on industrial development is only used until 1984. The main reason is that in this period there was a relatively stable institutional framework for industrial producers: the manufacturing sector was regulated by the Industries Development and Regulations Act that had been in place since 1951. The following year, in 1985, India started a wave of reforms in the manufacturing sector by delicensing around a third of all three-digit industries. From the outset, these reforms have changed considerably the dynamics of the manufacturing sector (see Aghion et al. (2006)). However, this twenty-year period from 1965 still provides me with a time span long enough to observe whether there is a persistent divergence in electrification across states that could be associated with initial groundwater after the beginning of the Green Revolution.

The evolution of electrification and manufacturing output and their link to initial groundwater availability is provided in Table 1. The table shows significant variation across states in the measure of groundwater, initial levels of electrification and industrialization and subsequent changes. Figure 1 shows a virtually flat relationship between rural electrification and manufacturing output in 1965. Figure 2 shows a positive correlation between changes in agricultural electrification and manufacturing output in the subsequent twenty years. Note that states with relatively high initial groundwater (Punjab, Haryana and Tamil Nadu) are amongst the faster growing states, and have experienced a larger expansion of their electricity network.

[6]Other measures of electrification include the connected load, i.e. maximum consumption available per consumer, measured in KW per million people, the number of industrial users connected to the network per 1000 people and the proportion of villages electrified.

[7]Other state controls include population density, proportion of literate population, and rural population, mean rural expenditure, credit, education and development expenditure and political outcomes. Sources and variables are all publicly available online from "EOPP Indian States Data", EOPP-STICERD, London School of Economics.

3 Groundwater as an Instrument

In this section I explore more systematically whether initial groundwater can be associated with an expansion of the electricity network (the first stage in the IV strategy) and with increases in manufacturing outcomes (the reduced form results). Before, I start by providing evidence of the link between groundwater availability and the diffusion of HYV seeds. This helps to back the claim that states followed either a path of progressive adoption of the new agricultural technology that created a demand for rural electricity, or a path of traditional farming. Subsequently, I present evidence towards the validation of the first stage and reduced form results.

3.1 Groundwater and HYV adoption

The adoption of HYV seeds in rural India has been subject to a number of studies. As noted by Sharma and Dak (1989) or Kohli and Sing (1997), among others, there are many institutional factors that may claim a share of responsibility in the adoption of this new technology (e.g. price incentives, land distribution or human capital.) However, since these variables might be correlated with (or determined by) industrial output there is a need for exogenous source of variation that explains HYV successful adoption. The answer lies in geographic and climatic characteristics. Foster and Rosenzweig (1996) point out that "a feature of the Green Revolution in India is that the ability to exploit the new seeds profitably was substantially different across India because of exogenous differentials in local soil and weather conditions". Similarly, Evenson and McKinsey (1999) show that climatic and soil conditions resulted in the diffusion of HYV seeds.

In this section I further investigate these claims by looking at whether groundwater availability can be associated to the evolution of HYV adoption, a necessary step to explain subsequent electrification. To that end, I first run a regression of the form

$$HYV_{st} = \beta_{0s} + \beta_{1t} + \sum_{k=1966}^{1984} \gamma_k (G_s * T_k) + \delta \mathbf{X}_{st} + \epsilon_{st} \tag{1}$$

where HYV_{st} is the proportion of cultivated land with HYV seeds in state s at time t, G_s captures states' time-invariant measure of groundwater, and T_k is a year dummy equal to 1 whenever $k = t$.

To account for the availability of abundant groundwater that could be obtained with electric pumpsets and used to provide timely and adequate irrigation, I use the measure of groundwater described above. State controls, X_{st}, are included to control for characteristics suspected to be correlated with the diffusion of HYV seeds, such as the proportion of rural population, the log of real per capita education expenditure, the log of real per capita credit availability and population density. Additionally, state and time fixed effects are included and errors are robust to heteroskedasticity and autocorrelation. The coefficients of interest in equation (1) are γ_k, where positive and significant values would indicate that states with greater access to groundwater adopted a greater proportion of HYV seeds. Increasing values of γ_t would also indicate that the difference was growing over time, i.e. the appropriate characteristics for successful adoption would trigger a path of divergence across states in HYV adoption. The estimation of regression (1) provides point estimates of the yearly impact of groundwater availability on HYV adoption, as captured by the coefficients γ_t. Results are shown in Figure 3 for specifications with and without state controls. All estimated coefficients are significant at the 1% level and show that states with better access to groundwater have become relatively more intensive in the use of HYV seeds.

3.2 Groundwater, Electrification and Industrial Output

The consistency of the IV estimator relative to OLS hinges on the power of the instrument, i.e. on how important is the correlation between the instrument (namely, G_s), and the instrumented variable, electricity provision (e_{st}). To test this, I check whether the timing at which some states have improved their electricity reach significantly coincides with the start of the Green Revolution. Additionally, I perform a similar exercise to see whether a similar patter emerges for the measure of industrialization used in the second stage of the IV strategy. As in the previous section, I use a specification where I estimate the time-varying effect of groundwater availability on outcomes (Y_{st}), whether electrification or manufacturing output:

$$Y_{st} = \beta_{0s} + \beta_{1t} + \sum_{k=t}^{1984} \gamma_k (G_s * T_k) + \delta \mathbf{X}_{st} + \epsilon_{st} \qquad (2)$$

Figure 4 shows four different specifications. In the first one, I use the measure of electrification on the left hand side, starting in 1950, with state and year fixed effects but no state controls. In the second regression, the data start in 1958 as to be able to include state controls. In the third and fourth, I use the log of manufacturing output with and without state controls, using data available (only from 1960). Results show that states with more groundwater have expanded the electricity reach in rural areas significantly only after the Green Revolution, where coefficients become significantly different from 0 at the 1% level. The slope also becomes steeper, suggesting an acceleration in the electrification process. Additionally, the reduced form results hold, since manufacturing output also seems to take off after the Green Revolution in states with abundant groundwater.

3.2.1 Evidence at the district level

In this section, I check the robustness of the first stage using information on the total number of pumpsets per capita available at the district level. In particular, I run a similar set of regressions as previously where the divergence in pumpset adoption (P_{dst}) is explained by the interaction of groundwater availability and year dummies. Even though there is no information on availability of electricity at the district level, the anecdotal evidence suggests a shift from diesel to electrified pumpsets after the start of the Green Revolution (see for example, Bharadway (1990) or McGuirk and Mundlak (1991)). To add confidence to the use of the Green Revolution as a natural experiment, I perform a similar regression where I replace the explained variable by a measure of HYV adoption.

$$P_{dst} = \beta_{0ds} + \beta_{1t} + \sum_{k=1966}^{1984} \gamma_t (G_{ds} * T_k) + \delta_t \mathbf{X}_{dst} + \epsilon_{dst} \tag{3}$$

As in the analysis for the state level, G_{ds} is a measure of groundwater availability, in this case a dummy that equals 1 if the district has aquifers thicker than 150 meters and 0 otherwise. The estimated γ_t provides the average conditional difference on pumpset use in year t between districts with an aquifer thicker than 150 mts and those with shallower aquifers. I exploit within district variation by using district fixed effects. I also include year dummies and run (3) with and without district controls X_{dst}[8].

Figure 5 plots the estimated γ_t for three specifications. As shown for the state level regressions, districts with greater groundwater abundance are associated with an increasing use of

[8]Time varying controls at the district level include literacy, population density, proportion of rural population, bullocks and tractors per hectare. Standard errors are clustered at the district level.

irrigation pumpsets and HYV seeds. Results hold in specifications with and without controls[9]. Estimates for the interaction terms are positive, significantly different from 0 only after the Green Revolution and increasing, meaning that there is a positive marginal effect for districts with more groundwater. This set of results is consistent with the findings at the state level and provides more evidence validating the choice of instruments.

An example: districts in Uttar Pradesh

It is apparent from the results on pumpsets that groundwater availability is associated with powered irrigation at the district level, a result that is consistent with the state level analysis. In Table 2 I check the link between deep aquifers and a measure of rural power availability, available at the district level, that was collected by the Ministry of Agriculture and Irrigation of India (1976) only for 1974. Columns (1) and (2) confirm the link between groundwater and rural power. To illustrate this point further I look at the state of Uttar Pradesh, since it provides a good testing ground of the link between groundwater and use of pumpsets and power availability across districts within a state. In particular, note that Uttar Pradesh's north-western districts are near the states of Haryana and Punjab, where the greatest proportion of districts with deeper aquifers can be found. At the same time, it limits to the South with Rajasthan and Madhya Pradesh and to the east with Bihar, all states with no districts with a watertable deeper than 150 mts. It is no surprise then that the only 8 districts with the deeper aquifers are all in the state's north-western region. When looking at the measure of power availability and operating pumpsets in Columns (3) to (8) in Table 2, the districts with more groundwater availability show significantly higher measures of rural power. As most districts in Uttar Pradesh (43 out of the 47) have at least 100 mts-deep aquifers, I run a regression comparing districts with the deepest aquifers and districts with the middle level of groundwater depth. Columns (5) and (8) show that the former have significantly more rural power. This evidence suggests that access to the deeper aquifers facilitated the expansion of the electricity network.

4 Main Results

In this section I explore further the whether the reduced form relationship between groundwater and industrial outcomes can be robustly ascribed to the expansion of the electricity network

[9]This statement applies to a regression of HYV adoption without controls, not reported but available upon request.

across Indian states. To that end, I start by estimating an OLS regression where manufacturing output (y_{st}) is expressed as a function of the measure of of electricity supply (e_{st}):

$$y_{st} = \alpha_{0s} + \alpha_{1t} + \lambda e_{st} + \alpha_3 \mathbf{X}_{st} + \mu_{st} \tag{4}$$

The regression includes time varying state controls (X_{st}), state fixed effects included to control for persistent and constant features within states, and year fixed effects to control for shocks common to all states. In particular, in the most parsimonious specification I control for demographic features (i.e. log of population density, proportion of rural population), credit availability (log of real per capita total credit) and a measure of development expenditure at the state level (log of real per capita expenditure in education). I discuss below a fuller specification[10].

The use of state fixed effects allows me to control for all state characteristics that remain constant over time, including unobserved variables that would contaminate the standard errors. Still, to account for short term correlation in shocks I assume standard errors to present first degree autocorrelation. However, note that results discussed throughout are robust to the presence of heteroskedasticity and to greater degrees of autocorrelation[11].

I subsequently run the IV specifications where e_{st} is replaced by $\hat{e}_{st} = f(G_s * T_t)$. That is, I instrument the electricity indicator using the measure of groundwater availability that, interacted with year dummies, has shown explanatory power with respect to states' divergence in both HYV diffusion and electricity expansion. Note that this specification does not impose a functional form on the divergence in electrification for states with more groundwater available. However, results in Figure 4 suggest that the divergence process might have followed a linear trend. To account for this, I also interact G_s with a linear time-trend in the first stage.

In the OLS estimation of the coefficient of interest, λ, endogeneity is a major concern for quantifying the effect of infrastructure on output. The presence of correlation between the error term and the explanatory variable could be explained by reverse causality (e.g. states that are industrializing faster demand investments in expanding the electricity network) or omitted

[10] All regressions use an efficient 2-step feasible GMM estimation.

[11] As pointed out by Wooldridge (2003) and Bertrand et al. (2004), clustering standard errors would not be appropriate to deal with within-state autocorrelation when the number of states is small. Results do not change when I allow for higher degrees of autocorrelation. To address this issue in a more general way I also perform successfully a wild cluster bootstrap-t estimation, as in Cameron et al. (2008), that has proven to work well with small number of clusters. Results are available upon request.

variables (e.g. unobserved changes in the institutional environment that drive both industrialization and electrification), and would introduce a bias to the estimation. However, I am agnostic about the sign of the bias, since alternative stories could work in both directions. For example, phenomena such as a pro-business environment or other political economy considerations could explain a positive correlation with both the outcome and the explanatory variable, biasing the estimate upwards. Alternatively, sources of bias towards zero could be any unobserved process driving rural electrification positively, but at the expense of industrial producers (such as skewed electricity pricing policies or rural lobbies) or the presence of attenuation bias because of measurement error in the electricity indicator (produced by illegal connections, theft of electricity, etc. that tend to be widespread in rural areas, see Tongia (2003)).

Results are shown in Table 3. Columns (1) and (2) show the OLS regressions with and without state controls, respectively, and find a positive and significant correlation between electrification and industrialization. Since the electrification process was targeted at agricultural producers, it is reassuring to find a positive and strongly significant correlation between electrification and industrial output once other time-varying factors and state fixed effects are controlled for.

Columns (3) to (6) show the key results in this paper, where I use the time varying effects of groundwater availability as an instrument for electrification. In Columns (3) and (4) the number of agricultural electricity connections is instrumented by the interactions between year dummies and the measure of abundant groundwater availability, controlling for state-time-varying characteristics only for the latter specification. In both cases, the coefficients are positive and strongly significant. Both coefficients are slightly greater than OLS estimates, suggesting the net presence of a downward bias. Columns (5) and (6) interact the measure of groundwater with a linear trend and find similar results in magnitude. Their main difference with the previous two results comes from a more powerful first stage, probably explained by the greater degrees of freedom in the latter specification's first stage.

The IV results could be interpreted as in the LATE literature: states that increased their rural electrification in 1 connection per 1000 people because of their groundwater endowment would experience around a 0.025 log points average increase in their manufacturing output per capita[12]. The size of the estimates of λ also suggest that the impact of electrification is

[12]In particular, there could be other margins that explain electrification, that are not captured by this instru-

substantial: an increase in one standard deviation in the number of rural connections implies an increase of around 13.6% in manufacturing output. It is important to note that, given the great variation in electricity provision across states, a state at the mean of the electrification distribution would need to double its rural electrification to observe such an effect. Another idea of the magnitude of the effect can be taken from asking what the benefit for a state at the 25th percentile of the distribution of electrification would be to move to the 75th percentile. For example, Bihar in 1984 had a real manufacturing output per capita of Rs 1206 and 2.64 farms connected to the network per 1000 people. If Bihar engaged in a rural electrification process to move to the top 25% states in terms of electrification, i.e. increasing the number of connections per 1000 people by 8.8[13], the model predicts that the manufacturing output would increase by 23.7%, or around 24 billion Rupees[14]. This would constitute an increase of around 4.2% of Bihar's total gross product per year, that can be used as a benchmark for how much the state could spend in a rural electrification programme that takes the state to the 75th percentile of the rural electrification distribution[15].

4.1 Electrification v. Alternative Channels

The main concern with respect to the identification strategy at hand is related to the exclusion restriction. This means that if the measure of rural electrification is strongly correlated with other state outcomes that are actually driving the industrialization process, then the results obtained in the previous section would be spurious. The use of observational data makes the task of validating the exclusion restriction very hard, since even though the characteristics of the water table in a state could be as good as random, it is not possible to rule out all alternative channels that could explain the complex process of industrialization.

In this section I discuss the likely alternative channels that might be explaining the reduced form results and I contrast them to the electrification channel proposed in this paper. To do so, I perform three different checks. I first explore whether the electrification channel holds when exposed to variation in other observable state characteristics associated to mechanisms

ment. For example, this strategy would not capture the effect of electrification in states with no groundwater that have somewhat electrified for other reasons.

[13]That is an increase of around 650,000 connections, up from 200,000.

[14]This is obtained by multiplying the estimated increase in real manufacturing output per capita times total population in Bihar in 1984 (around 74 million people). The value is in 1974 Rupees.

[15]Note that this measure does not take into account the effect of electrification in other economic activities, including agriculture.

that the economic literature has underlined as important in the process of industrial growth. I subsequently take a step back in terms of the electrification channel and put to the data the following question: could the same strategy be used to explain divergence in industrial development through other means than electrification? In short, the second set of checks consists in running OLS and IV specifications with these alternative stories as in Table 3 to see whether the first and second stages hold across specifications. The third and final set of checks implies looking at the composition of industrial outcomes, that could provide suggestive evidence against the electrification mechanism.

4.1.1 Alternative Channels: a brief discussion

There are many stories in the literature as to why shocks to agricultural productivity would spill over to the manufacturing sector. Johnson (1997) points out that a rapid increase in agricultural productivity contributed to the start of the Industrial Revolution, by releasing (cheap) labor for the manufacturing sector and creating the agricultural surplus needed to feed the growing population. This argument could be linked to a process of urbanization as the agricultural sector lays workers off. To check this I use the proportion of urban population in a state. Matsuyama (1992) formalizes these mechanisms in a model where both sectors compete for labor and where exogenous increases in agricultural productivity benefit the manufacturing sector. This mechanism works only in closed economies, since comparative advantages in the agricultural sector might deprive industries from labor in open economies. In that case, effects on industrialization could go either way[16]. Murphy et al. (1989a) propose a mechanism wherein industrialization happens because positive shocks to agricultural productivity boost rural income and – subsequently – the demand for manufactures. This result hinges on two conditions: non-homothetic preferences (the share of expenditure on manufactures increases with income) and a fair distribution of benefits from increases in agricultural productivity to sustain a sizeable demand[17]. Information on the mean expenditure of rural households allows me to control for this channel.

Other stories might imply the development of complementary markets, such as credit mar-

[16]India, being a very large federal country, could be thought as a collection of open economies. In that case, the mechanisms described by Matsuyama might be present.

[17]See also Voigtlander and Voth (2006). This could undermine the argument when applied to India, since it is well documented (see for example, Dhanagare (1989)) that the Green Revolution has increased the rift between poor and rich farmers and, in some cases, has not even increased poor peasants' real incomes.

kets. For example, the increase in rural incomes affects saving rates in rural areas and, subsequently, credit availability. To check this, I use information on non-agricultural rural credit. Similarly, Foster and Rosenzweig (1996) have shown greater private investments in schooling after the Green Revolution has increased returns to education. I use the proportion of literate population, that might also be at the heart of improvements in the manufacturing sector.

Finally, there is the political economy channel that raises two issues: the first one is why farmers' demand for electricity was satisfied in groundwater (and HYV) intensive states. Second, whether these political circumstances that facilitated electrification can also explain, independently, improvements in the manufacturing sector.

Two mechanisms could explain why farmers' need of electricity was satisfied by their state governments. The first one is linked to the preferences of the median voter. In the period analyzed, rural population in any given state is at least 65% of total population, and almost 80% on average (topping 90% in three states), in a country that has been democratic and federal since its independence in 1947. There is some evidence that voters in states intensive in HYV like Punjab, Haryana and Tamil Nadu moved away from the Congress Party and voted for regional parties. A second mechanism is supported by anecdotal evidence: the endogenous formation of lobbies. It seems important to note that "increasing food shortages and mounting concern for immediate gain in production led to the shift in developmental priorities" (Sharma and Dak, (1989)), such that expanding the production possibilities in the agricultural sector was deemed fundamental. That circumstance gave farmers, in particular big ones, an unprecedented political clout over state governments. Dhanagare (1989) estimates that "the prosperity unleashed by the Green Revolution was distributed differentially, putting the small and marginal farmers at a relative disadvantage. The high cost/high yield technology called for capital investments beyond the means of a majority of small and marginal farmers." It might be the case that a stronger economic position translated into political leverage. Tongia (2003), for example, stresses that rural electrification "has swayed in strong political winds." Gulati and Narayanan (2003) also show that what they call 'the subsidy syndrome in Indian agriculture' after the introduction of the new seeds in late 60s became a fundamental instrument of economic policy where available electricity was one of its main channels.

With respect to direct effects on manufacturing outcomes, there are many channels that could operate. For example, as noted by Laffont (2005), government inefficiencies and corruption increase the marginal cost of raising funds and constrain the ability of the executive to invest in infrastructure. Alternatively, other phenomena such as the power of industrial lobbies or unions, or the social roots of a party might explain why resources are shifted in one or another way. It follows that the indicators of electricity network development might actually be capturing these political economy features that could also drive the dependent variable.

To address the concern that the political economy channel is driving the results, I include state controls that capture political party outcomes (party allegiance of the Chief Minister and share of the votes to the main national party, the Congress Party of India). Additionally, I also use expenditure in development sectors (infrastructure, health and education) that can capture not just states' investment in physical and human capital, but also the ability and/or willingness of the state government to enhance the economic environment. Similarly, I use the labor regulation measured constructed by Besley and Burgess (2004) to capture whether the states have enacted more pro-worker or pro-business regulation over time.

4.1.2 Alternative channels as additional controls.

In Table 4 I progressively include the above-described variables as controls and check whether the electrification channel holds. In all specifications, the coefficient on electrification is not only significant but also its magnitude remains quite stable. The inclusion of these variables does not reduce significantly the explanatory power of the electricity story.

The coefficients on the control variables do not provide much additional information. In fact, the political economy story seems to be the one providing more explanatory power to differences in manufacturing output. First, results suggest that an active state, as captured by the development expenditure, is always positive and significant. Column (5) also shows that, as in Besley and Burgess (2004), more pro-worker regulation is negatively correlated with manufacturing outcomes. These two results suggest that lobbying might indeed have had an impact in industrial outcomes. A similar picture emerges when looking at electoral outcomes. States where the Congress Party gets more votes and run the state government or states that vote for a leftist government seem to industrialize more than states that vote for the Janata or other parties. Note that the use of state fixed effects imply that these results do not come only from the cross-section of states. Overall, the coefficient on electrification remains significant and very similar in magnitude, suggesting the electrification channel remains strong. It has to be noted, however, that this evidence does not preclude the possibility that alternative unobserved time-varying phenomena with enough idiosyncratic variation could be at play. In that sense, these results only suggest that, with the information at hand, the electrification channel remains robust.

4.1.3 Groundwater as an instrument for alternative channels.

In this section I investigate further if some of the alternative stories could have been explaining the reduced-form results documented in the previous section. Basically, whether the time-varying effect of groundwater availability is a strong predictor of variation in these variables in the first stage and, if so, if the predicted variation can explain manufacturing outcomes in the second stage. Even though variation in groundwater was presented as very specific to the electrification channel, because of the use of electrified pumpsets, it might well be the case that it is affecting manufacturing output through any of the channels associated to the HYV agricultural productivity shock.

Table 5 presents OLS and IV results (when using the flexible and the linear specifications). I am interested in testing whether the estimates show the expected sign and similar levels of significance for the point estimates and the first stage F-tests.

The main conclusion is that none of the alternative stories passes all tests consistently. In some cases, e.g. literacy, rural expenditure or rural credit, the OLS estimates are not significant or do not have the expected sign. However, it could well be the case that the estimates are biased and the IV regressions would reverse that pattern. The three IV specifications show that this is not the case. In particular, is not just that across the board estimates are not significantly different from 0, but that the first stage would be extremely weak. The time-varying effects of groundwater availability provide a very powerful first stage for the case of electrification, but would be a hard sell for any other alternative channel, where I get results that show no consistent pattern[18].

These results increase the confidence that the variation used in the instrumental strategy

[18] I run a similar specifications using political and urban outcomes, such as credit or mean expenditure. In all cases, the first stage yields very low values and coefficients are not significantly different from 0.

works through the proposed channel and not through others. However, it does not mean that the alternative channels are unimportant or did not reflect differences across states, but simply that the variation used to instrument for the measure of electrification cannot be used to explain variation in urbanization rates, credit availability, literacy, state expenditure in development or expenditure in rural households. The use of state fixed effects in all specifications means that I exploit within state variation, and only in the case of the electricity measure does the time-varying effect of groundwater availability has explanatory power for understanding states' divergence in industrial development.

4.1.4 Industry composition

In this section I break down manufacturing output in different ways, to test whether the effects on aggregate manufacturing outcomes can be associated with improvements in access to electrification or to other channels associated with the Green Revolution (such as an increase in food production) or to groundwater availability. In the first split of the data, I make use of the distinction in India between the formal (or 'registered') sector (i.e. firms with more than 10 workers that use power or more than 20 workers without power) and the informal (or 'unregistered') sector of smaller firms. Since the use of electricity is an important distinction between the two, I would expect the manufacturing output of the 'registered' sector to be more affected by improvements in access to electricity than the 'unregistered' sector. Columns (1) and (2) in Table 6 show that this is the case. The expansion of the electricity network is associated with increases in manufacturing output among firms in the formal sector where the coefficient is positive and strongly significant. Its magnitude is larger than the one obtained for the whole manufacturing sector, an unsurprising result since no significant effect is present for the informal sector[19].

Columns (3) to (8) make use of data by industry, only available from 1980 to 1984. Because of the little time variation, in these specifications I only use the measure of groundwater as an instrument. In Columns (3) and (4) I split industries according to their intensity in electricity usage, using an input-output matrix for all India industries, and find that the effect of electrification on industrial output is only present for industries that are above the median in electricity

[19]Unreported results using census data for two years also show that employment in the rural manufacturing sector also goes up with the instrumented measure of electrification. Overall, there is no enough data on the rural manufacturing sector to produce further checks.

usage.

Another way of checking whether the Green Revolution has a direct effect in the development of the industrial sector is by looking at the food processing industries. Since the Green Revolution has increased considerably food production only, it could be the case that this drives manufacturing output up by boosting the industrial processing of foodstuff. Columns (5) and (6) show that this is not the case. The coefficients for both food processing industries and the rest are very similar[20]. Finally, the same exercise could be done with respect to water use. A concern might be that states with abundant groundwater are better suited to increase their production in industries that are water intensive, such as textiles or pulp and paper[21]. If that was the case, the instrument would be affecting the dependent variable through means other than electrification. Columns (7) and (8) show positive results for both groups of industry water-intensity, that are not significantly different from each other. Once again, the electrification channel remains the most robust of all.

4.2 Mechanisms

There are direct and indirect mechanisms through which the expansion of the electricity network can affect industrialization. For example, access to electricity can lower costs, then prices and induce more consumption of manufacturing goods[22]. Murphy et al. (1989b) show that infrastructure can be an important component of a 'big push' industrialization process. By overcoming coordination problems that arise because infrastructure goods are used by many sectors at the same time, the public provision of infrastructure contributes to the development of other markets, since it has the effect of reducing the total production costs of the other sectors. In a setting with heterogeneous firms, a selection effect might be present, wherein smaller and less productive firms can break even only when electricity is available. In that sense, the effects of electrification should be present both in the intensive and extensive margins.

An alternative story could be that electrification in rural areas reduces home production time and increases labor supply (as in Dinkelman (2010)) or simply that the increase in electrification

[20] Splitting the sample according to whether industries have strong linkages to the agricultural sector or not provides qualitatively similar results.

[21] However, the effect could go the other way. For example, Keskin (2009) shows that industries and the agricultural sector might compete for water and that industrialization might hurt farmers.

[22] A look at the input-output coefficient matrix for industries in the early 1990s India shows that for many industries, such as heavy chemicals, cement or non-ferrous metal, electricity is more than 10% of their total costs and can be as high as 17% (Ministry of Industry, Government of India (1993)).

(and agricultural productivity) is freeing labor resources from the agricultural sector. That would translate into a reduction in wages that could spurt the manufacturing sector.

To test these mechanisms, I first make use of the Annual Survey of Industries' distinction of firms in the 'registered' sector between 'small sector' factories , i.e. those with less than 50 workers (and with more than 50 and less than 100 workers that do not use power), and 'census sector' factories, i.e. those with more than 100 employees. In particular, to tell apart the intensive and extensive margins of industrialization, I look at other indicators of industrial development (value added and fixed capital) for all firms and for the small sector. Columns (1), (2), (4) and (5) in Table 7 show positive and significant results. The coefficients for all firms are larger than for the small sector, suggesting that the effect is not all coming from smaller firms[23]. Another way of testing the extensive margin is by looking at whether new factories open. In column (3) I find that electrification is significantly associated with more factories. The magnitude suggests that 1 extra electricity connection per 1000 people would be associated with 0.003 factories per 1000 people. For a state with the average population of 1965, i.e. 30 million, that would mean 90 extra factories. These results provide evidence of effects happening both at the intensive and the extensive margins: a deepening of the industrial sector among big and small firms and the presence of a selection effect where new and smaller firms also benefit by the expansion of the electricity network.

To test the indirect channel, I use information on real and nominal wages in the manufacturing sector. In both cases, shown in columns (6) and (7) there is no evidence that wages have dropped significantly in states that electrified more. This is not consistent with the idea that manufacturing output has increased because of a drastic reduction in wages (that would have followed an stark increase in labor availability).

5 Robustness Checks

In this section I investigate further to what extent the main set of results presented in section 4 hold when using alternative measures of groundwater availability and electrification.

[23]This holds when restricting the dataset to state-year observations that are both available for all firms and for the small sector.

5.1 Other Measures of Groundwater

The evidence presented so far has linked a measure of groundwater availability that only considered the presence of aquifers deeper than 150 mts. The district level evidence in Figure 5 suggested that this measure was a good predictor of divergence in pumpset use and the example of Uttar Pradesh in Table 2 showed an example of difference in rural power when comparing districts with very aquifers and deep aquifers (i.e. between 100 and 150 mts deep). What happens with the main results when I relax the constraint imposed on the measure of groundwater availability?

I consider four measures to check the robustness of results in Table 3. In the first one, I construct a measure of groundwater that weights each district by area, relative to the state area. In the second and third measures I use, respectively, the proportion of districts with aquifers deeper than 100 mts and the proportion of districts with any positive groundwater depth. Finally, the fourth alternative measure uses data on net groundwater availability (in cubic meters per squared km), as estimated by Central Ground Water Board (2006) for the year 2004.

Table 8 shows results when using the flexible year-specific and the linear time trends. Across specifications the second stage shows a positive and significant effect of electrification on manufacturing outcomes. However, with the exception of the area-weighted measure, the first stage looks weaker for the three other specifications (even though when instrumenting with the interaction between each groundwater measure and the linear time trend, the first stage performs significantly better.) This suggests that using information on the deepest aquifers only is a better predictor for subsequent electrification, or that that the same amount of water but distributed in shallower watertables would not be as effective[24].

5.2 Other Measures of Electrification

The use of alternative measures of electrification provide a similar picture. I use information on the percentage of villages electrified, on the state technical capacity both to provide (i.e. connected load) and to generate electricity and on the number of industrial connections. Finally,

[24]To check to what extent results are driven by a "lucky draw" in the measure of groundwater, I run 2000 Monte Carlo simulations, in each case randomly assigning a value between 0 and 1 to each state and see use that measure as an instrument. Results, available upon request, show that the probability of obtaining the main results by a random draw are in the order of 0.15%

it might be the case that the expansion of the network has also brought a reduction or an increase in quality that might explain part of the results. To test this I use information on Transmission and Distribution losses (as a percentage of total electricity available).

Results in Table 9 show that the link between electrification and manufacturing output holds for the four alternative measures of network expansion. The first-stage statistics are lower than for the measure of agricultural connections, suggesting the stronger link between groundwater and rural electrification[25]. The result on the quality measure, in column (5) does not provide evidence of a positive or negative link with manufacturing output.

There is no consistent information on electricity prices by type of consumers for the period analyzed, but the anecdotal evidence (see Tongia (2003), for example) suggest a significant cross-subsidy to rural and household consumption from industrial and commercial users across most Indian states. However, I can use information on average electricity tariffs and average cost, provided from 1974 by the Planning Commission (2001). When running the instrumented measure of agricultural electricity connection on the ratio average tariff-average cost I find that states that have electrified more have, on average, a significant 2 percentage points lower ratio[26]. That means that pricing policies seem to be less sustainable in states with more rural electrification. Furthermore, a regression of the tariff-cost ratio on manufacturing output shows a negative and significant coefficient. That means that states with more sustainable tariffs (and probably less cross-subsidies to agricultural and household consumers) have observed a lower industrial growth. Overall, this would suggest that, if these higher prices of electricity for industrial users affect significantly manufacturers' outcomes, the main results would underestimate the effects of network expansion in a setting without cross-subsidies.

6 Conclusion

The role that infrastructure provision plays in improving economic and social outcomes in developing countries is part of the conventional wisdom among academics, policy makers and the population at large. Despite this consensus, the academic literature has not provided a substantial amount of empirical investigations aimed at understanding and quantifying the effects of large scale infrastructure provision. The challenge of this paper is to document the

[25] However, the First-stage is substantially stronger when using the interaction with the linear time trend.

[26] Results available upon request.

effect of electrification in industrial output by looking at a panel of Indian states from 1965-1984. To do so, I address the endogenous placement of infrastructure by looking at the time-varying effect of time-invariant geographic characteristics on electrification after the start of the Green Revolution in India. The need for timely irrigation that could be cheaply supplied by pumping water from the water table using electrified pumpsets generated a growing demand for electricity and subsequent expansion of the electricity grid in some states. Instrumental variables estimations show that one standard deviation difference in electrification can explain around 14% of the differential level in manufacturing output across Indian states. The expansion of the network also helped the entry and performance of smaller firms. These results are robust to a series of tests regarding the power and validity of the instruments and support the idea that the lack of infrastructure is a major constraint on economic activity in developing countries. The magnitude of the results suggests that expanding the electricity network – including rural areas – should be considered seriously as a policy option for promoting industrial development. Still, a degree of caution is needed here: electrification might have had these effects because it was a significant binding constraint for manufacturing firms across Indian states. However, this might not always be the case, in particular if the population is subject to other constraints, such as transport infrastructure or credit availability that would make the provision of electricity less effective.

References

[1] Aghion, P., Burgess R., Redding, S. and Zilibotti, F. (2008).The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India, American Economic Review, 98(4), 13971412

[2] Bharadway, K. (1990), Irrigation in India: Alternative Perspectives, Indian Council of Social Science Research

[3] Bertrand, M., Duflo, E. and Mullainathan, S. (2004), How Much Should We Trust Differences-in-Differences Estimates?, Quarterly Journal of Economics, 119(1), 249-75.

[4] Besley, T. and Burgess R. (2004), Can Labor Regulation Hinder Economic Performance? Evidence from India, Quarterly Journal of Economics, 119(1), 91-134

[5] Burgess, R and Pande, R. (2005), Can Rural Banks Reduce Poverty? Evidence from the Indian Social Banking Experiment, American Economic Review, 95 (3), 780-795

[6] Cameron, C; Gelbach J and Miller, D (2008), "Bootstrap-Based Improvements for Inference with Clustered Errors", Review of Economics and Statistics, 90, 414-427

[7] Chakravarti, A.K. (1972), Green Revolution in India, Annals of the Association of American Geographers, 63 (3), 319-330

[8] Central Ground Water Board, Ministry of Water Resources, Government of India (2006), Dynamic Ground Water Resources of India (as of March 2004), Faridabad.

[9] Department of Energy, USA (2003), An Energy Overview of India, mimeo.

[10] Dhanagare, D.N. (1989), The Green Revolution and Social Inequalities in Rural India, in Green Revolution and Social Change, edited by Sharma and Dak, Ajanta Publications, Delhi.

[11] Dinkelman, T. (2010), The effects of Rural Electrification on Employment: New Evidence from South Africa, University of Michigan, mimeo.

[12] Duflo, E. and Pande, R. (2007), Dams, Quarterly Journal of Economics, 122 (2), 601–646

[13] Estache, A. (2004), Emerging Infrastructure Policy Issues in Developing Countries: A Survey of the Recent Economic Literature, INFVP, The World Bank.

[14] Evenson, R.E. and McKinsey, J.W. (1999), Technology-climate interactions in the Green Revolution in India, Economic Growth Center Discussion Paper 805, Yale University.

[15] Foster, A and Rosenzweig, M.(1996), Technical Change and Human Capital Returns and Investments: Evidence from the Green Revolution,, American Economic Review 86(4), 931-953.

[16] Foster, A and Rosenzweig, M. (2003b), Agricultural Development, Industrialization and Rural Inequality, mimeo, Harvard University.

[17] Foster, A and Rosenzweig, M. (2008), Inequality and the Sustainability of Agricultural Productivity Growth,: Groundwater and the Green Revolution in Rural India, mimeo.

[18] Gulati, A. and Narayanan, S. (2003), The Subsidy Syndrome in Indian Agriculture, Oxford University Press.

[19] Johnson, D.G. (1997), Agriculture and the Wealth of Nations, American Economic Review, 87(2), 1-12

[20] Keskin, P. (2009), Thirsty Factories, Hungry Farmers: Intersectoral Impacts of Industrial Water Demand, mimeo, Harvard Kennedy School

[21] Kohli, D. S. and Singh, N. (1997), The Green Revolution in Punjab, India: The Economics of Technological Change, mimeo.

[22] Laffont, J-J (2005), Regulation and Development, Federico Caffe Lectures, Cambridge University Press.

[23] Matsuyama, K. (1992), Agricultural Productivity, Comparative Advantage and Economic Growth, Journal of Economic Theory 58, 317-334

[24] McGuirk, A and Mundlak, Y. (1991), Incentives and Constraints in the Transformation of Punjab Agriculture, International Food Policy Research Institute.

[25] Ministry of Agriculture and Irrigation, Government of India (1976), Report of the National Commission on Agriculture, Volume X

[26] Ministry of Industry, Government of India (1993), Handbook of industrial statistics

[27] Ministry of Power, Government of India (2001), Blueprint for Power Sector Development, mimeo

[28] Murphy, K., A. Shleifer and R. Vishny (1989a), Income Distribution, Market Size, and Industrialization, The Quarterly Journal of Economics, 104(3), 537-564

[29] Murphy, K., A. Shleifer and R. Vishny (1989b), Industrialization and the Big Push, The Journal of Political Economy, 97(5), 1003-1026

[30] Planning Commission, India (1961), Third Five Year Plan.

[31] Planning Commission, India (2001), Annual Report on the working of State Electricity Boards and Electricity Departments, mimeo.

[32] Sims, H. (1988), Political Regimes, Public Policy and Economic Development: Agricultural Performance and Rural Changes in Two Punjabs, Sage Publications, New Delhi.

[33] Sharma, M.L. and Dak, T.M (1989), Green Revolution and Social Change, Introduction, edited by Sharma and Dak,Ajanta Publications, Delhi.

[34] Staiger, D. and Stock, J.H. (1997), Instrumental Variables Regression with Weak Instruments, Econometrica, 65, 557 – 586.

[35] Stock, J and Yogo, M. (2003), Testing for Weak Instruments in Linear IV Regression, manuscript.

[36] TERI (1999), Indian Power Sector: change of gear, mimeo.

[37] Tongia, R. (2003), The Political Economy of Indian Power Sector Reforms, WP 4, Program on Energy and Sustainable Development, Stanford University

[38] Voigtländer, N. and Voth, H.-J. (2006), Why England? Demographic factors, structural change and physical capital accumulation during the Industrial Revolution, Journal of Economic Growth, 11(4), 319-361

[39] Wooldridge, J. (2003), Cluster-Sample Methods in Applied Econometrics, American Economic Review, XCIII, pp 133-138.

[40] World Bank (2000), India: Policies to Reduce Poverty and Accelerate Sustainable Development, World Bank.

[41] World Bank (2005), Infrastructure and the World Bank: A Progress Report, World Bank.

[42] World Bank Report (2006), Infrastructure: Lessons From the Last Two Decades of World Bank Engagement, World Bank.

Table 1: States: Groundwater, Electrification and Manufacturing Output

State	Groundwater	Agricultural Connections (per 1000 people)						Real Manufacturing Output (per capita)					
		Level	Change					Level	Change (%)				
	1965	**1965**	**1965-70**	**1970-75**	**1975-80**	**1980-84**	**1965-85**	**1965**	**1965-70**	**1970-75**	**1975-80**	**1980-84**	**1965-84**
Andhra Pradesh	**0.05**	1.48	2.98	1.77	2.20	2.25	**9.21**	675	16.57	18.56	17.87	14.02	**85.75**
Assam	**0.11**	0.00	0.00	0.05	0.05	0.04	**0.14**	1095	-41.26	29.75	-36.18	113.43	**3.81**
Bihar	**0.00**	0.21	0.98	0.74	0.39	0.31	**2.43**	610	-9.58	44.66	-17.15	82.44	**97.72**
Gujarat	**0.00**	0.74	1.95	1.57	2.71	1.15	**7.39**	1727	5.46	4.84	33.14	18.47	**74.41**
Haryana	**0.50**	1.56	7.34	3.95	5.13	1.47	**17.89**	1142	6.57	18.55	42.68	30.89	**135.95**
Karnataka	**0.00**	1.66	2.93	2.44	1.64	2.78	**9.79**	812	79.59	-4.12	1.56	14.98	**101.08**
Kerala	**0.12**	0.38	0.53	0.91	1.86	1.07	**4.37**	738	25.80	-0.09	50.75	-9.85	**70.82**
Madhya Pradesh	**0.00**	0.22	1.22	2.45	3.08	1.94	**8.68**	552	21.36	4.85	60.48	7.05	**118.59**
Maharashtra	**0.00**	1.10	3.30	2.82	3.69	3.08	**12.89**	2304	19.15	-7.32	57.81	-12.34	**52.76**
Orissa	**0.08**	0.01	0.01	0.13	0.46	0.32	**0.92**	718	-14.23	-19.63	103.34	4.40	**46.35**
Punjab	**0.60**	1.86	5.04	3.08	7.01	6.25	**21.38**	1063	13.08	42.99	5.56	29.60	**121.21**
Rajasthan	**0.00**	0.30	1.13	1.83	3.00	1.16	**7.12**	600	14.68	-8.19	19.68	21.96	**53.68**
Tamil Nadu	**0.08**	6.93	5.03	4.61	1.88	2.18	**13.70**	1559	14.07	-12.73	78.03	11.66	**97.87**
Uttar Pradesh	**0.17**	0.18	1.19	1.05	0.68	0.45	**3.36**	558	33.70	1.17	12.65	28.60	**95.95**
West Bengal	**0.13**	0.02	0.01	0.19	0.22	0.03	**0.46**	2037	-8.70	12.11	28.76	-4.78	**25.49**

Figure 1: Agricultural Connections and Manufacturing Output: Initial Levels

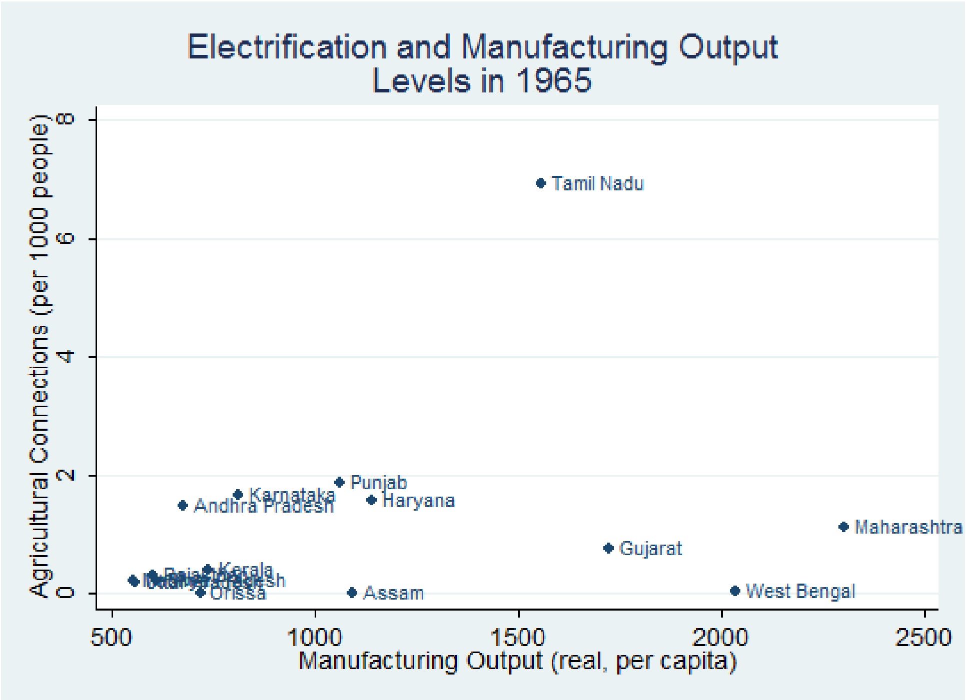

Figure 2: Agricultural Connections and Manufacturing Output: Changes

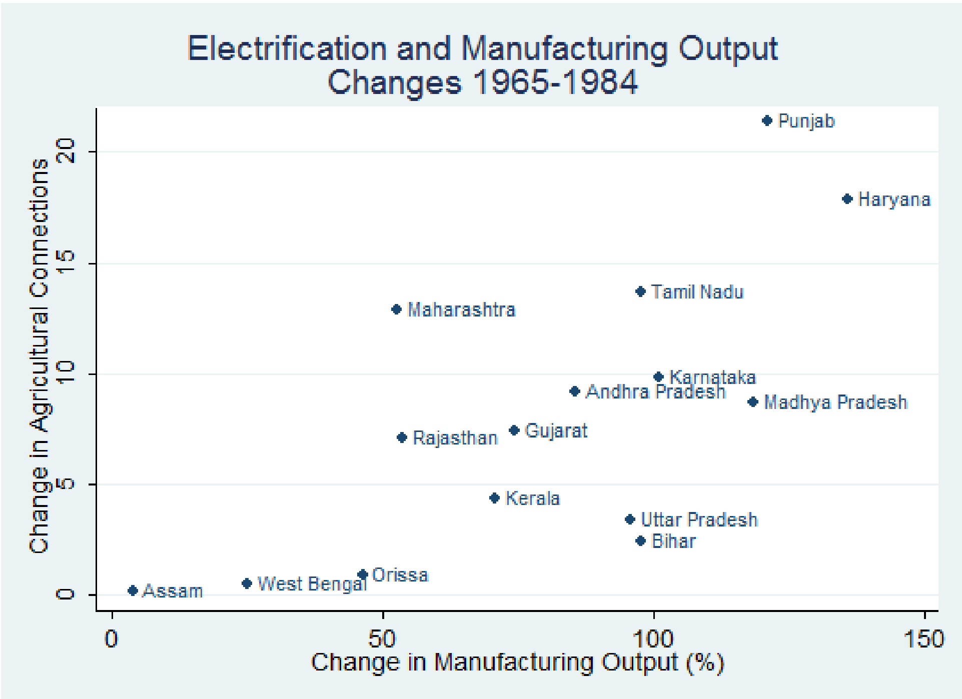

Figure 3: Groundwater and HYV adoption

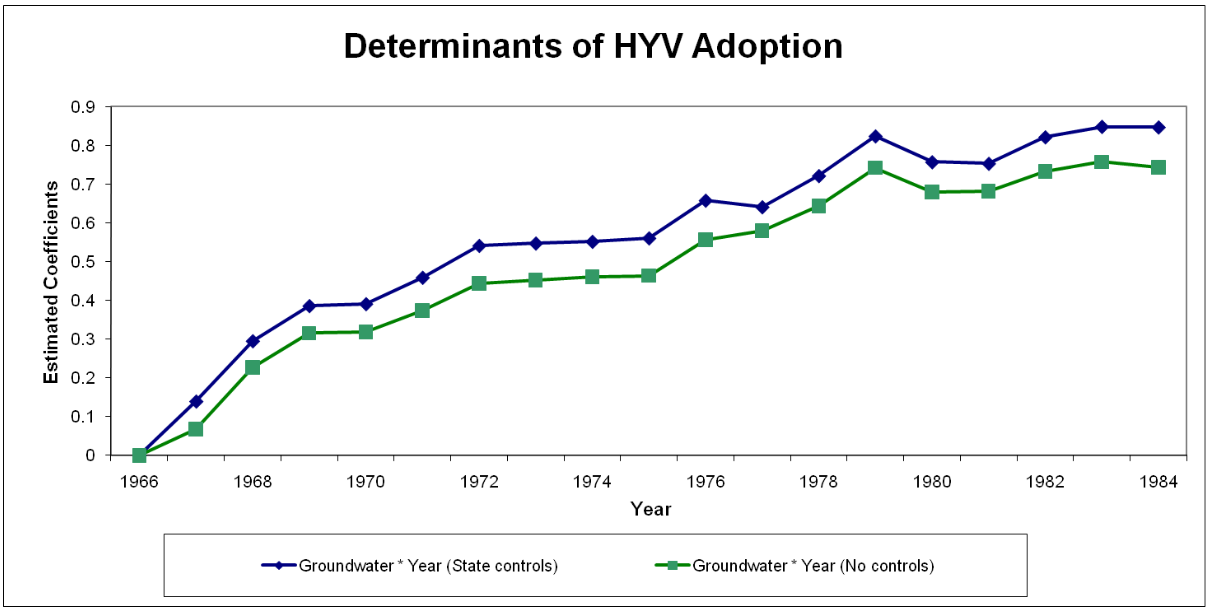

Figure 4: Groundwater, Electrification and Manufacturing Output

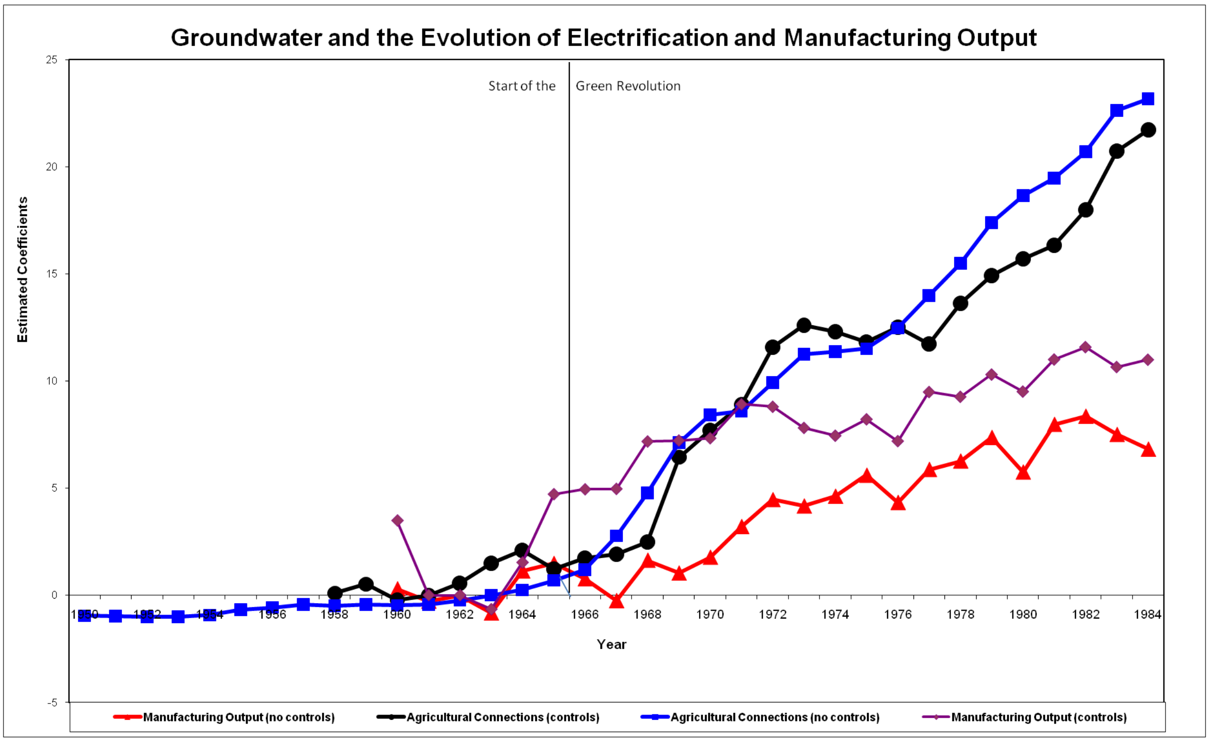

Figure 5: Differential pumpset use and HYV adoption in districts with deep groundwater

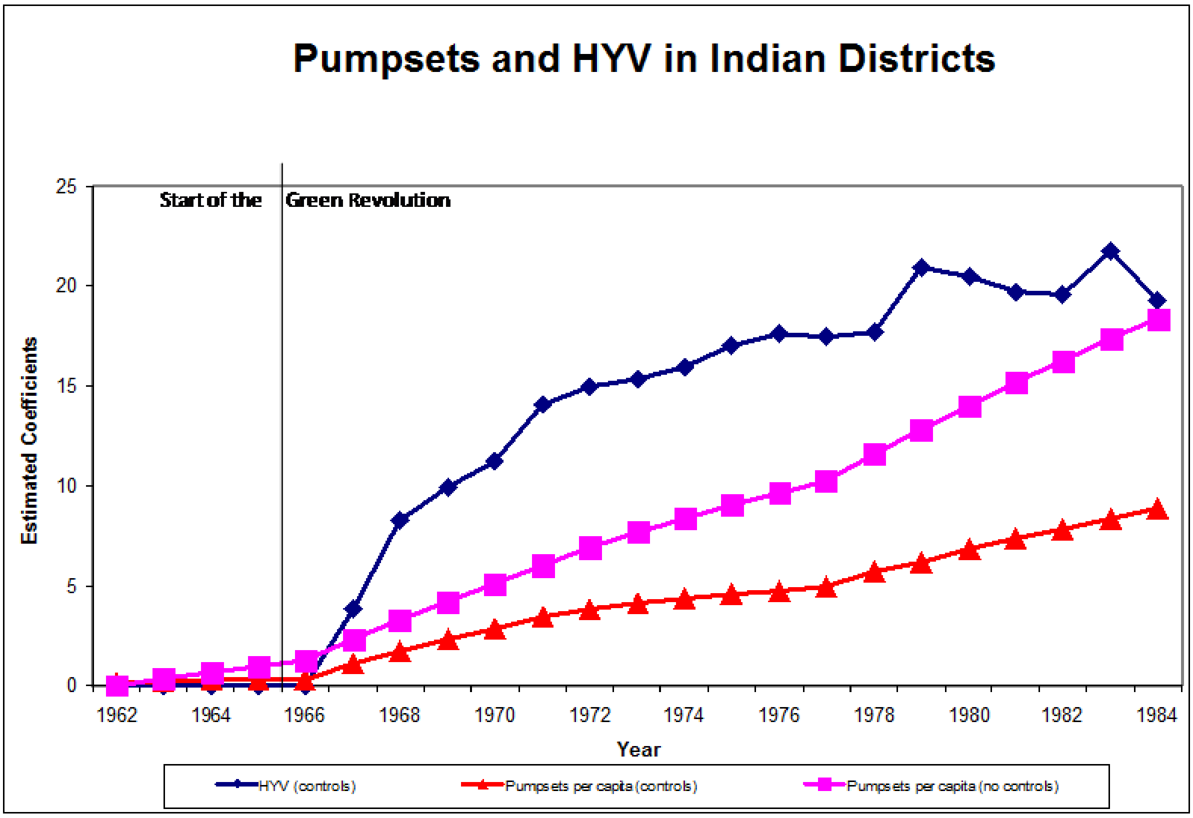

Table 2: Groundwater and Power Availability at the District Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Power Availability (1974)		Power Availability (1974)			Total Pumpsets (1981)		
	All India		Uttar Pradesh					
Groundwater	**0.237**	**0.195**	**0.150**	**0.131**	**0.152**	**0.013**	**0.007**	**0.007**
	(0.104)**	(0.09)**	(0.034)***	(0.053)**	(0.062)***	(0.003)***	(0.003)**	(0.003)*
District Controls	No	Yes	No	Yes	Yes	No	Yes	Yes
Observations	270	270	47	47	43	47	47	43

Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. "Power availability" is a measure, in horses per hectare, of power in rural areas. "Groundwater" is a dummy equal to 1 if the district has aquifers thicker than 150 mts. Controls are: population density, literacy, roads and bullock per hectare. In columns (3) and (6) I exclude districts with aquifers shallower than 100mts.

Table 3: Main Results: OLS and IV

OLS and IV

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Manufacturing Output					
	OLS		IV			
Agricultural Connections	**0.029**	**0.022**	**0.030**	**0.025**	**0.032**	**0.024**
	(0.006)**	(0.007)***	(0.006)***	(0.007)***	(0.007)***	(0.009)***
Instruments	n/a	n/a	Groundwater * Year Dummies		Groundwater * Time Trend	
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State Controls	No	Yes	No	Yes	No	Yes
Observations	300	300	300	300	300	300
Over-id P-value	n/a	n/a	0.98	0.95	n/a	n/a
First Stage F-Test	n/a	n/a	8.51	12.69	67.39	133.25

* significant at 10%; ** at 5%; *** at 1%. Standard errors in parentheses, robust to heteroskedasticity and autocorrelation. IV estimations use 2-step efficient GMM estimation. "Agricultural Connections" is the number of agricultural units connected to the electricity network per 1000 people. "Log Manufacturing Output" is the log of real per capita manufacturing output. "Groundwater" is the proportion of districts per state with aquifers thicker than 150 mts. Controls are log of education expenditure, proportion of rural population, population density and log of total credit per capita. F-test: Staiger and Stocks (1997) rule of thumb is that instruments are weak if the first-stage F is less than 10, the Stock-Yogo Weak ID test critical value for 2SLS bias being less than 10% of OLS bias is 11.46 for Columns (3) and (4), 16.38 for columns (5) and 19.93 (6). All data are for 1965-1984.

Table 4: Main Results with Additional Controls

	(1)	(2)	(3)	(4)	(5)
	Log Manufacturing Output				
	IV using Groundwater * Year Dummies				
Agricultural Connections	**0.022**	**0.022**	**0.023**	**0.020**	**0.024**
	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***
Development Expenditure	**0.282**	**0.280**	**0.283**	**0.277**	**0.251**
	(0.069)***	(0.071)***	(0.071)***	(0.072)***	(0.070)***
Literacy		**-0.132**	**-0.078**	**-0.051**	**-0.073**
		(1.299)	(1.299)	(1.289)	(1.258)
Mean HH Rural Expenditure			**-0.041**	**-0.050**	**0.016**
			(0.124)	(0.124)	(0.118)
Non Agricultural Rural Credit				**0.043**	**0.013**
				(0.048)	(0.050)
Labor Regulation					**-0.048**
					(0.020)**
Vote Share Congress Party					**0.005**
					(0.001)***
Chief Minister Congress Party					**0.056**
					(0.030)*
Chief Minister Hard Left Party					**0.113**
					(0.060)*
Chief Minister Janata Party					**0.008**
					(0.037)
Observations	300	300	300	300	300
First Stage F-Test	19.28	13.87	15.77	18.45	16.80

* significant at 10%; ** at 5%; *** at 1%. Standard errors in parentheses, robust to heteroskedasticity and autocorrelation. IV estimations use 2-step efficient GMM estimation. All regressions include state and year fixed effects. "Agricultural Connections" is the number of agricultural units connected to the electricity network per 1000 people. "Log Manufacturing Output" is the log of real per capita manufacturing output. "Groundwater" is the proportion of districts per state with aquifers thicker than 150 mts. F-test: Staiger and Stocks (1997) rule of thumb is that instruments are weak if the first-stage F is less than 10, the Stock-Yogo Weak ID test critical value for 2SLS bias being less than 10% of OLS bias is 11.46 for all columns. All data are for 1965-1984.

Table 5: Groundwater as an Instrument for Alternative Explanations

Specification		(1) Agricultural Connections	(2) Rural Expenditure	(3) Urban Population	(4) Rural Non Agric. Credit	(5) Literacy	(6) Development Expenditure
OLS	Estimate	0.02	0.02	0.06	0.05	-0.06	0.27
		(0.007)***	(0.13)	(0.04)*	(0.05)	(1.19)	(0.07)***
IV using Groundwater * Year Dummies	Estimate	0.02	-0.17	0.07	0.06	-1.57	0.21
		(0.007)***	(0.23)	(0.07)	(0.08)	(4.77)	(0.12)*
	F-test Statistic	18.45	3.67	2.26	5.74	0.69	4.90
IV using Groundwater * Time-Trend	Estimate	0.02	-0.17	0.13	0.09	0.35	0.08
		(0.007)***	(0.60)	(0.12)	(0.12)	(2.21)	(0.45)
	F-test Statistic	66.70	9.49	15.40	39.49	0.02	1.73

* significant at 10%; ** at 5%; *** at 1%. Standard errors in parentheses, robust to heteroskedasticity and autocorrelation. IV estimations use 2-step efficient GMM estimation. "Agricultural Connections" is the number of agricultural units connected to the electricity network per 1000 people. "Log Manufacturing Output" is the log of real per capita manufacturing output. "Groundwater" is the proportion of districts per state with aquifers thicker than 150 mts. "Rural Expenditure" is the log average expenditure in rural households, "Rural Non Agric. credit" is a log measure of real per capita rural credit not used in agriculture and "Development Expenditure" is the log of real per capita expenditure by the state government in health, education and other development projects. "Literacy" and "Urban Population" are proportions of total population. F-stat critical levels are 11.46 in row 3 and 19.43 in row 5. All data are for 1965-1984.

Table 6: Industry Composition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Manufacturing Output							
	Registered Sector	Unregistered Sector	High Electricity	Low Electricity	Food Sector	No Food	High Water Use	Low Water Use
Agricultural Connections	**0.044**	**-0.005**	**0.060**	**0.016**	**0.046**	**0.038**	**0.051**	**0.033**
	(0.008)***	(0.014)	(0.021)***	(0.016)	(0.026)*	(0.017)**	(0.017)***	(0.017)**
Instruments	Groundwater * Year Dummies		Groundwater					
State Fixed Effects	Yes	Yes	No	No	No	No	No	No
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	300	300	75	75	75	75	75	75
First Stage F-Test	12.69	12.69	53.00	53.00	53.00	53.00	53.00	53.00

* significant at 10%; ** at 5%; *** at 1%. Standard errors in parentheses, robust to heteroskedasticity and autocorrelation. IV estimations use 2-step efficient GMM estimation. "Agricultural Connections" is the number of agricultural units connected to the electricity network per 1000 people. "Groundwater" is the proportion of districts per state with aquifers thicker than 150 mts. "Registered sector" includes factories with more than 10 workers and power or more than 20 workers without power. "Electricity Use" and "Water Use" are high if the input use is above median for all Indian industries and low otherwise. Controls are as in Tables 2 and 3. F-test: Staiger and Stocks (1997) rule of thumb is that instruments are weak if the first-stage F is less than 10, the Stock-Yogo Weak ID test critical value for 2SLS bias being less than 10% of OLS bias is 11.46 for columns (1) and (2); 16.38 for the remaining columns. Data is 1965-1984 for columns (1) and (2) and 1980-84 for all other columns.

Table 7: Mechanisms: small sector, factories and wages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Value Added		Factories	Fixed Capital		Manufacturing Wage	
	All	Small Sector		All	Small Sector	Real	Nominal
	IV using Groundwater * Year Dummies						
Agricultural Connections	**0.045**	**0.014**	**0.003**	**0.071**	**0.034**	**-0.001**	**0.004**
	(0.006)***	(0.003)***	(0.001)**	(0.02)***	(0.005)***	(0.005)	(0.005)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	300	195	293	280	224	300	301
First Stage F-Test	12.69	11.09	12.07	9.95	10.39	12.69	12.69

* significant at 10%; ** at 5%; *** at 1%. Standard errors in parentheses, robust to heteroskedasticity and autocorrelation. IV estimations use 2-step efficient GMM estimation. "Fixed Capital" and "Value Added" are per capita for the manufacturing sector. "Agricultural Connections" is the number of agricultural units connected to the electricity network per 1000 people. "Groundwater" is the proportion of districts per state with aquifers thicker than 150 mts. All outcome variables are in real terms and per capita. "Small sector" are factories in the 'registered sector' with less than 50 workers using power or more than 50 and less than 100 workers without power. "Wages" are log of wages for workers in the manufacturing sector.

Table 8: Alternative Measures of Groundwater

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Manufacturing Output							
	IV using Groundwater * Year Dummies				IV using Groundwater * Time-trend			
Agricultural Connections	**0.026**	**0.032**	**0.035**	**0.036**	**0.024**	**0.043**	**0.048**	**0.035**
	(0.008)***	(0.010)***	(0.011)**	(0.014)***	(0.008)***	(0.015)***	(0.016)***	(0.015)**
Measure	Area-weighted	>100 meters thick	All aquifers	Quantity of groundwater	Area-weighted	>100 meters thick	All aquifers	Quantity of groundwater
Observations	300	300	300	300	300	300	300	300
First Stage F-Test	9.93	1.37	1.30	0.76	108.03	9.35	11.50	9.50

* significant at 10%; ** at 5%; *** at 1%. Standard errors in parentheses, robust to heteroskedasticity and autocorrelation. IV estimations use 2-step efficient GMM estimation. "Agricultural Connections" is the number of agricultural units connected to the electricity network per 1000 people. "Log Manufacturing Output" is the log of real per capita manufacturing output. In columns (1) and (5) "Groundwater" is the proportion of districts per state with aquifers thicker than 150 mts, weighted by the relative size of the district. In columns (2) and (6), "Groundwater" is proportion of districts per state with aquifers thicker than 100 mts, and in columns (3) and (7), the proportion with any positive groundwater depth. Columns (4) and (8) use the net groundwater available (in cubic meters per squared km), as estimated by Central Ground Water Board (2006). All regressions use state and year fixed effects and controls (as in Table 3).

Table 9: Alternative Measures of Electrification

	(1)	(2)	(3)	(4)	(5)
	Log Manufacturing Output				
	All regressions use Groundwater * Year Dummies as Instruments				
Percentage Villages Electrified	**0.405**				
	(0.156)***				
Connected Load (Kwpc)		**0.003**			
		(0.0008)***			
Industrial Connections (per 1000 people)			**0.261**		
			(0.069)***		
Generation Capacity (Kwpc)				**0.005**	
				(0.002)***	
T&D losses					**-0.897**
					(0.731)
Observations	300	300	300	300	268
First Stage F-Test	7.47	10.69	2.46	7.10	1.59

* significant at 10%; ** at 5%; *** at 1%. Standard errors in parentheses, robust to heteroskedasticity and autocorrelation. IV estimations use 2-step efficient GMM estimation. "Agricultural Connections" is the number of agricultural units connected to the electricity network per 1000 people. "Log Manufacturing Output" is the log of real per capita manufacturing output. "Groundwater" is the proportion of districts per state with aquifers thicker than 150 mts. "Connected Load" and "Generation Capacity" are measured in KW per million people. "Industrial Connections" is measured per 1000 people. "T&D losses" are Transmission and Distribution loses as a proportion of total electricity produced. All regressions use state and year fixed effects and controls (as in Table 3).

Table 10: NOT FOR PUBLICATION: Standard Errors: Alternative Treatments

	Log Manufacturing Output (IV results)				
	All regressions use Groundwater * Year Dummies as Instruments				
Agricultural Connections	0.024 (0.008)***	0.026 (0.008)***	0.025 (0.008)***	0.023 (0.008)***	0.024 Tests reject the null at 5%
Change in SE estimation	Autocorrelation of degree:			Bootstrap	Wild Bootstrap
	2	3	4	200 reps	1000 reps
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
State Controls	Yes	Yes	Yes	Yes	Yes
Observations	300	300	300	300	300
First Stage F-Test	10.87	9.86	9.70	16.20	not reported

Table 11: NOT FOR PUBLICATION: Additional Specification Checks

	Log Manufacturing Output							Rural Manuf Workers	
								% Rural Workers	% Total Workers
Agricultural Connections	0.025 (0.008)***	0.022 (0.006)***	0.025 (0.007)***	0.034 (0.010)***	0.025 (0.007)***	0.026 (0.013)***	0.023 (0.008)***	0.001 (0.0003)***	0.0005 (0.0003)**
Rural Literacy	-3.280 (1.69)**								
Neighboring States' Electrification		0.047 (0.011)***							
Other Controls/Specification	Rural Literacy (GW*Timetrend)	Connections in neighboring states	Initial Road Density Time trend	Initial Literacy, Roads, Manuf Output Timetrend	No Rajasthan or AP (GW*Timetrend)	No Punjab (GW*Timetrend)	LIML: First stage 10% critical value: 3.18	Only observations for 2 census years	
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	210	300	300	300	260	280	300	24	24
First Stage F-Test	102.8	10.70	11.60	9.33	70.24	23.16	12.70	30.91	30.91