FACTORIES AND FARMS: HOW DOES ECONOMIC GROWTH IMPACT RURAL INCOMES AND EDUCATION?

REENA BADIANI[†]

JOB MARKET PAPER

Abstract. This paper investigates the effects of rising incomes and returns to education on educational investment in rural India between 1983 and 1999. To do so, I first develop a model of household education choice that illustrates the effects of increases in agricultural productivity and policy induced expansions in skilled and unskilled manufacturing labor demand on returns to education and income for skilled and unskilled workers. Second, I exploit plausibly exogenous variation in skill-specific labor demand caused by increases in regional agricultural productivity and policy driven variation in manufacturing employment for different skill levels. The results show that an increase in unskilled labor demand decreases returns to education for all workers, but has income effects that vary by skill and land ownership. An increase in skilled labor demand increases returns to education for all workers but only increases income for skilled households. The results on educational investment show that investment for boys responds positively to both income and returns to education. For girls, investment in education only responds to increases in income, but not to returns.

[†]Department of Economics, Yale University, Box 208268, New Haven, CT 06520-8268
E-mail address: reena.badiani@yale.edu.

I am indebted to Joe Altonji, Mike Boozer, Mark Rosenzweig and Chris Udry for their guidance and comments. I thank David Atkin, Abhijit Banerjee, Prashant Bharadwaj, Gharad Bryan, Rahul Deb, James Fenske, Rachel Heath, Stefan Hunt, Pinar Keskin, Katrina Jessoe, Mushfiq Mobarak, Nancy Qian, Gus Ranis, Steve Redding, Maher Said, Peter Schott, Paul Schultz, TN Srinivasan and Vis Taraz for suggestions and support. I thank Padmaja Mehta at the Ministry of Agriculture for access to data. Saba Joshi provided excellent research assistance. All remaining errors remain mine.

1. Introduction

Economic growth is associated empirically with rising investment in education. This empirical regularity could be driven by an income effect as well as by a returns to education effect. Growth may directly raise the returns to schooling, for example skill biased technological change that raises the demand for educational skills (Nelson and Phelps, 1966; Schultz, 1975). Growth also alters the level and distribution of income through changing the returns to unequally distributed productive assets (Kuznets, 1955). The literature on the demand for education in developing countries has found that education responds positively to both incomes (Jacoby and Skoufias, 1997) and returns (Kochar, 2004; Foster and Rosenzweig, 1996). The relative contribution of these channels is however unclear since the literature has typically focused on one determinant of education at a time. Focusing on one channel may confound the two effects if growth raises wage and income levels as well as alters returns to education.

This paper makes two principal contributions to understanding how economic growth alters household education choices and incomes. First, it jointly investigates the effects of rising incomes and returns to education on educational investment in rural India between 1983 and 1999. The two effects are separated by examining *two* determinants of wage and income growth, which induce independent variation in household incomes and returns to education. Second, the sources of growth examined induce exogenous movements of labor between the agricultural and non-agricultural sectors. The paper examines the response of the level and distribution of incomes in rural areas to an expanding manufacturing sector.

The approach taken by this paper is to examine the response of wages, household income and education to two determinants of growth: region-time varying agricultural productivity and policy driven region-time variation in skilled and unskilled manufacturing employment. Agriculture and manufacturing differ in their demand for skill and land. Hired labor in agriculture conducts low skill tasks such as harvesting. Labor in manufacturing conducts skilled and unskilled tasks ranging from managerial to manual. Exogenous variation in the size of these sectors shifts skilled and unskilled labor demand differently, and changes the return to land. Therefore, sector specific shocks differ in their effect on returns to education, as captured by the ratio of skilled to unskilled wages, as well as on household incomes, according to their skill and land assets. I investigate how these two sources of growth alter returns to education and household incomes, and how changes in returns and household incomes alter household educational investment.

To frame my empirical analysis, I build a small-country general equilibrium model of household education choices in rural labor markets. The model illustrates the effects of increases in agricultural productivity and policy induced expansions in skilled and unskilled manufacturing labor demand on returns to education and households income, according to their skill and land endowments. The model highlights a source of endogeneity bias: regional manufacturing employment reflects local wages and unobserved determinants of education. The model motivates an instrumental variables strategy to overcome this bias: industries are attracted to locations endowed with the raw materials they use. Changes to industry level import tariffs and regulations therefore alter labor demand in regions endowed with the resources the industry uses.

The predictions of the model are tested using a district and household level data-set covering over 250 districts in 15 major Indian states between 1983 and 1999. The data-set includes district-level mineral, metal and energy endowments and is compiled from maps. India presents the perfect setting for studying how shocks to the agricultural and manufacturing sectors alter equilibrium outcomes. India in 1983 was split into 412 districts and migration between districts is low. The empirical specifications examine wage, household income and education responses to within district over-time variation in agricultural productivity and instrumented skilled and unskilled manufacturing employment. Agricultural productivity is measured using a Lespeyres index of regional agricultural potential: I apply the national time-varying yield frontier in seven crops to the fraction of each region physically suited to each crop. Changes in the national crop frontier occur due to the introduction of new varieties of seeds, for example.

In the first stage instrumental variables specification, I use the predictions of the model to induce variation in manufacturing employment that is uncorrelated with unobserved determinants of income and education. I exploit the observation that industries are attracted to districts endowed with the materials they use intensively. For example, cement industry employment is concentrated where limestone is found, while plywood industry employment is higher in densely wooded districts (figure A.1). Changes to industry policy are prediced to directly affect employment in districts endowed with the materials the industry uses. India relaxed industrial regulations and reduced import tariffs during a series of crises starting in 1984 (Topalova, 2004; Srinivasan, 2002). The timing and depth of policy changes varied across industries. THe paper exploits within-district over-time responses of manufacturing employment to policy changes, where the identifying variation comes from differences across industries in the broad types of raw materials used and in industry policy over time. Employment is broken apart by skill by exploiting variation in the timing and depth of policy changes across industries with different skill intensities.

The wage and income results indicate that the two drivers of growth differ in their impact on the returns to education and household incomes. An increase in agricultural productivity or instrumented unskilled manufacturing employment decreases returns to education, but has income effects that vary by skill and land ownership. An increase in skilled employment increases returns to education for all workers but only raises income in educated households. Households with different initial land and education endowments exhibit heterogeneous income responses to different drivers of growth.

The education specifications examine whether children aged between 5 and 9 start school. Unskilled manufacturing growth increases the probability a child starts school in unskilled landless households, while it reduces it in landed households. The intuition lies in the income effect: unskilled manufacturing growth raises income in unskilled landless households, while it has no effect in landed households.

The income and returns to education effects on educational investment are estimated using minimum distance. Since agricultural productivity and manufacturing employment alter educational investment through incomes and returns to education, the coefficients on these two terms in the education specification reflect weighted combinations of income and returns to education effects, where the weights are wage and income responses to the two determinants of growth. I use the estimated coefficients from the education, wage and income specifications to jointly estimate the two effects for rural landless households.

I find that income growth accounted for a larger fraction of the rise in education seen between 1983 and 1999 than growth in returns to education. Together, growth in incomes and returns accounted for 85% of the observed education growth. Male educational investment responds positively to rising household incomes and labor market returns to education. In contrast, investment among girls only responds to rising incomes. This finding fits with the observation of low female participation in complex occupations in rural areas. I test whether raising the child wage reduces education through the opportunity cost channel. I find no effect, consistent with the low incidence of child labor among young children.

The estimates shed insight into a fundamental question about the process of development: how does an expanding non-agricultural sector alter the level and distribution of incomes in predominantly agrarian areas? Between 1983 and 1999, unskilled manufacturing growth reduced income inequality, while agricultural productivity growth raised it. However, manufacturing growth contributed little to raising poverty relative to agricultural productivity growth. Manufacturing growth explains a tenth of the reduction in poverty and unskilled wages growth over the period, while agricultural productivity growth explains just under half. The small relative impact of manufacturing is partly attributable to its skill composition: skilled manufacturing contributes little to reducing poverty or raising unskilled wages. If manufacturing employment growth had been purely unskilled, the impact on poverty, wages and incomes would have been double. The small impact of manufacturing growth on rural poverty has repeatedly been asserted in the policy literature (Kocchar, 2004). The estimates in this paper are the first to empirically validate this observation and to provide an explanation for why this is the case.

The results suggest that, in terms of male educational investment, a rising tide lifts all boats. Regardless of the source of growth, education rises among the poorest households in India - rural, unskilled, landless households. Sources of growth that raise incomes but reduce returns to education, such as agricultural technological change and unskilled manufacturing growth, increase educational investment substantially. Skill biased growth raises returns but has little impact on incomes; it raises education to a lesser degree through the returns to education channel. The same is not true for girls, whose education was found to only respond to rising income: skill biased growth has little impact on their education.

This paper contributes to an extensive literature on the demand for education in developing countries.[1] The two determinants of education investment have not previously been captured in one paper; a contribution of this paper is to estimate them jointly and capture their relative effects. The paper also contributes to a growing literature on non-agricultural growth in rural areas.[2] I deviate from the literature by breaking manufacturing apart by skill and finding an exogenous source of variation in it.

The paper is structured as follows. The next section describes the setting of rural India and provides descriptive evidence supporting my main results. In section 3, I put forward a model to frame the empirical analysis. I test the model's 4 sets of predictions in sections 4 to 8. Section 9 concludes.

[1]To name but a few, the demand for education has been shown to respond to rising returns to education (Kochar, 2004; Rosenzweig and Foster, 1996), unanticipated and anticipated growth in household incomes (Edmonds, 2006; Edmonds et al, 2004; and Jacoby and Skoufias, 1997), the opportunity costs of schooling (Rosenzweig and Foster, 2004) and to changes in the direct cost of schooling (Schultz, 2004; Kremer et al. 2003).

[2]Prominent papers include Lanjouw and Murgai, 2009; Eswaran et al., 2009; Foster and Rosenzweig, 2004; Datt and Ravallion, 1999.

2. Setting: The Rural Indian Economy

In this section, I provide descriptive evidence supporting the central arguments of the paper. Section 2.1 shows that occupations vary substantially in their complementarity with education across the agricultural and manufacturing sector. Hired workers in agriculture conduct unskilled tasks, while those in manufacturing conduct skilled and unskilled tasks. Section 2.2 shows that wages in the rural labor market rise with education and reflect the skill levels of tasks conducted. Finally, section 2.3 shows trends in incomes, labor market returns to education and education participation. The data indicate a strong positive correlation between incomes, returns to education and education participation. Incomes and the proportion of children starting school have risen over time, while the average ratio of skilled to unskilled wages has fallen. Thus the descriptive statistics point towards income driving rising educational investment.

The descriptive statistics covering India's 15 major states are from four rounds of household data collected in 1983, 1987, 1993 and 1999 by the National Sample Survey Organization (NSSO). In addition, I make use of the 1982 and 1999 rounds of data from the Rural Economic and Development Surveys (REDS) collected by the National Council of Applied Economics Research (NCAER). The Data Appendix describes the data in greater detail.

2.1. **Occupations and Education in Rural India, by Sector.** In this section, I provide evidence that agriculture hires workers to conduct unskilled tasks while manufacturing hires workers to conduct both skilled and unskilled tasks. Workers conducting skilled tasks have higher education levels.

Table 1-a presents descriptive statistics that divide the rural Indian workforce into agricultural and non-agricultural sectors in 1983 and 1999. Agriculture employed a large fraction of the rural workforce in 1999 - 65% of males and 34% of females. The size of the agricultural workforce varies substantially across India: in Madhya Pradesh 85% of the workforce was employed in agriculture compared to 41% in Kerala and 61% in Tamil Nadu. The manufacturing sector employs approximately 31% of non-agricultural workers, making it the largest non-agricultural sector of employment (Government of India, 1999). In 1992, manufacturing accounted for 17% of GDP while agriculture accounted for 23% (Panagariya, 2008).

Agricultural workers can broadly be divided into those from landed households, who cultivate their own land and may work on the wage labor market, and those from landless households, who predominantly work in the wage labor market. Table 1-a shows that approximately half of the agricultural workforce worked for wages, 28.5% of rural males. In a weekly time recall, hired workers report spending 99% of their time on unskilled tasks such as weeding and sowing. These wage laborers are disproportionately illiterate - in 1983, 71% of males were illiterate, compared to 53% in the working age population as a whole. Illiteracy among females wage laboreres is substantially higher - 94% of females were illiterate, compared to 83% in the population as a whole.

Skilled work is more prevalent amongst individuals who cultivate their own land - in the REDS data, 13% of household agricultural time among landed households is devoted to supervisory activities. Foster and Rosenzweig (1996) present evidence that education raises the productivity of individuals in landed households who undertake managerial tasks, such as deciding what varietals of seeds to sow. Tables 1-a and 1-b show that individuals who cultivate their own household's land have higher levels of education than hired labor. Household heads, who are likely to be making input decisions, are slightly more educated than own-household cultivators who are less likely to be making these choices.

Labor in the manufacturing sector conducts both skilled and unskilled activities. Table 2-a examines education levels by worker category: skilled (white-collar and blue-collar skilled) and unskilled (blue-collar unskilled). Unskilled work is that in which purely physical tasks were reported.[3] Approximately 35% of male and 55% of female occupations in manufacturing fall into this category. Education decreases in the skill category of the occupation conducted. White-collar and blue-collar skilled workers are comparable in their levels of education - 78% of male white collar workers are literate compared to 72% of blue-collar skilled workers. Blue-collar unskilled workers have lower levels of literacy relative to the more skilled manufacturing groups - 49% were literate in 1987.[4]

Table 2-b describes the different types of tasks conducted in manufacturing and provides an indication of why education varies by skill category. Task definitions come from the US Dictionary of Occupational Titles, 1977. Since there are likely to be technological differences between manufacturing industries in the US and India, this data provides some intuition for why workers conducting skilled occupations have higher levels of education than those conducting unskilled occupations.[5] Each task variable is assigned a value ranging from zero to ten, with higher values representing greater intensity of a skill used. The five broad measures of tasks are routine manual activity, non-routine manual activity, routine cognitive tasks, non-routine interactive tasks (direction, control and planning) and quantitative and analytical tasks (mathematical reasoning). White collar occupations use quantitative and non-routine cognitive skills more than both sets of blue-collar occupations. The blue-collar occupations are quite similar in all attributes, with the exception of routine cognitive tasks which are conducted to a far higher degree in the skilled blue-collar bracket.

2.2. **Wages by Sector, Occupation and Complexity of Task.** The previous section showed that skilled tasks in both manufacturing and cultivation are associated with higher levels of education than unskilled tasks. This section presents evidence showing that rural wages vary with education and the complexity of task conducted.

Table 1-c displays descriptive statistics on wage earners by sex. Literate individuals working in the wage labor market earn a 20% premium over illiterate workers. Separating workers by sector as well as by literacy, illiterate manufacturing workers earn a 10% wage premium over illiterate workers in other sectors

[3]Occupations were classified using descriptions in the National Classification of Occupation (Government of India, 1968). Occupations described as "work doing purely physical activity" were classified as blue collar unskilled while more technical descriptions - working with machines, mixing paint using the right proportions of different chemicals - were defined as blue collar skilled. White collar occupations include managers, secretaries and production-floor supervisors. The codes and categorizations are available upon request.

[4]This is shown in table 1-b, which conditions on village fixed effects to alleviate concerns that the descriptive statistics reflect decisions by the manufacturing sector to establish itself in more educated regions.

[5]A crosswalk was created between the 1968 Indian National Classification of Occupations and the 1960 US Census of Occupation Titles. This is available upon request. The data on occupational skills comes from Autor et al (2003). I thank David Autor for sharing this data with me.

(column c). This premium appears to be driven by the complexity of the task performed rather than by education level - an illiterate person conducting an unskilled task in manufacturing earns the same wage as an illiterate person in the other non-agricultural sectors. Non-manual workers in agriculture earn a premium over agricultural laborers. This should however be taken with a note of caution since the number of workers falling into this industry-occupation cell is extremely small - 0.7% of hired labor in agriculture reported conducting these occupations.

2.3. **Education, Incomes and Wages.** The descriptive statistics presented in this section point towards incomes driving the rise in educational investment: incomes and the proportion of children starting school have risen over time, while the ratio of skilled to unskilled wages has fallen. Descriptive statistics on educational and child labor are presented in table 4-a and 4.b.

Between 1982 and 1999, average real household income grew by approximately 70% in the REDS data while in the NSSO real per capita consumption grew by 22%. Agricultural wages grew by approximately 52% over this period.[6] The real wages received by literate and primary educated workers in the non-agricultural sector grew by 43% and 42% respectively. The wage received by manual labor in agriculture can be considered as the reservation wage for uneducated workers in rural areas (Deaton and Dreze, 2002). I therefore use the agrarian wage to proxy for the unskilled wage, and the wages of literate/primary educated workers as the skilled wage. The skilled wage ratio is defined as the ratio of the skilled wage relative to the unskilled wage. The literate wage ratio has declined by 9%, from 1.72 in 1983 to 1.58, while the primary wage ratio has declined by 14% over the period. Therefore, rural areas in India have experience substantial rises in the level of incomes between 1983 and 1999, although the skilled wage ratio has decreased on average over time.

The education margin examined in this paper is whether to send children to school or not. This is an important margin of choice in India, despite free tuition fees: 30% of boys and 52% of girls aged between 5 and 9 in 1983 never started primary school.[7]School attendance has increased sharply between 1983 and 1999 - by 1999, 9% of boys and 14% of girls hadn't started school.[8] I examine the entry decisions in retrospect, where children aged between 5 and 9 in 1983 are aged between 10 and 14 in the 1987-88. Education investment varies substantially across states - from near universal male enrollment in Kerala as early as 1983 to 45% of boys out of school in Bihar. Kerala is excluded from the analysis since the decision to start schol doesn't appear to be a relavant margin of choice in this state.

Child labor is low among children aged 5 to 9. In 1987, 2.5% of boys and 8% of girls reported conducting domestic or income generating activities. Among children not attending school, the majority report no activity. These patterns suggest that children who are not attending school are likely to be engaged

[6]The NSS data indicate that wages grew by 59% between 1983 and 1999, while the REDS data point to a much sharper rise of agrarian wages of 68%.

[7]Tuition fees in government primary schools are free across all states (Mehta, 1996), but additional expenditures such as uniforms and books, account for approximately 300 Rs per annum on average, or approximately 10 days of male agrarian wage labor (Gov of India, 1995).

[8]The startling trends seen in the NSS employment-unemployment rounds are confirmed in the National Family Health Survey which has a more detailed education module (Kingdon, 2007).

in leisure activities.[9] The proportion of children working in either an income generating or domestic capacity is strongly negatively correlated with wage levels, and in particular with the unskilled wage. This correlation is driven by households with below median landholdings, who report the highest levels of child labor.

Table 4-b indicates that the proportion of children starting school in a district is strongly positively correlated with the level of skilled and unskilled wages. Areas with higher levels of skilled and unskilled wages therefore see higher levels of educational investment. The raw correlation between the ratio of skilled and unskilled wages and education suggests that districts in which the returns to education are high also experience higher investment in education. The statistically significant positive correlation between levels of wages and proportions of children attending school holds in each cross-section, but the statistically significant relationship between returns and schooling holds only in 1993 and 1999.

3. Theoretical Model

This section presents a model of household education choices that builds upon Foster and Rosenzweig (1996, 2004) and Ellison and Glaeser (1999). Te assumptiosn of the model draw upon the descriptive statistics presented in the previous section. The model provides a setting for understanding household education choices in the context of a two-sector rural labor market. The purpose of the model is two fold. First, it delivers 3 sets of predictions that are tested in section 4. Second, it highlights a source of endogeneity bias and puts forward exclusion restrictions to overcome it.

3.1. **Model Environment.** Households live in a region whose borders define a closed labor market. A country is made up of D regional economies. There are two sectors in each economy: agriculture and manufacturing, which are modeled using a specific factors framework. The two sectors overlap in two locally traded inputs - skilled and unskilled labor. Equilibrium wages, household incomes and education choices are determined within the model, and vary with the regional economic environment. Economy and time subscripts are omitted until necessary.

3.2. **Households: Endowments, Preferences and Maximization Problem.** Households are modeled over two periods and consist of two generations. In the first period, each household includes an adult a and a child y. In the second period, the adult dies and the child becomes a young adult.[10] Household subscripts are suppressed in this section. Households have three endowments in both periods: land (A), adult education (s^a) and a unit of adult labor market time. In the first period, households are additionally endowed with 1 unit of child time. This can either be devoted to education or to producing a domestically consumed good. Adult education takes the value of 0 if the adult is uneducated or 1 if educated. Adult education in period 2 is the only endogenously determined endowment. Adults choose the child's education in period 1. The choice is discrete - the child either attends school or does not.

[9]These patterns do not appear to be an artifact of the broad questions on an individual's principal usual activities since they are both seen in the weekly time recall data and in the more detailed REDS data-set.
[10]I abstract from the child-bearing decisions that follow.

Households have time separable preferences over consumption in period 1 and 2 - c_1 and c_2. Utility is concave in consumption - $0 < \rho < 1$.[11]

$$V(c_1, c_2) = u(c_1) + \beta u(c_2) = c_1^{\rho} + \beta c_2^{\rho} \tag{1}$$

Consumption goods are of two perfectly substitutable varieties: home produced by the child using a unit of child time l_y or purchased at price p_c:

$$\begin{aligned} c_1 &= c_{1d} + c_{1p} = \alpha l_y + c_{1p} \\ c_2 &= c_{2p} \end{aligned}$$

Households are unable to borrow or save, implying that their budget constraints clear every period.[12] In the first period, income is divided between purchased consumption and schooling. In the second period, income is spent only on purchased consumption. For tractability, there is no leisure; in the empirical framework, I discuss how incorporating leisure alters the testable predictions.[13]

Adults split their labor market time between four activities: unskilled work in agriculture, unskilled work in manufacturing, skilled work in manufacturing and skilled work in agriculture. Unskilled work in agriculture and manufacturing are activities which both educated and uneducated individuals can conduct; skilled work is defined as activities that only educated individuals can do. Skilled work in agriculture consists of supervisorial activities, which can only be conducted by members of landed households. In the first period, children either go to school or produce the domestic consumption good. The time constraints of adults and children in period 1 and adults in period 2 are given by:

$$\begin{aligned} 1 &= l_{u1}^A + l_{u1}^M + \mathbf{1}(s_1^a = 1)(l_{s1}^M + \mathbf{1}(A_h > 0) l_{s1}^A) \\ 1 &= (1 - s_1^y) l_{y1} + s_1^y l_{s1} \\ 1 &= l_{u2}^A + l_{u2}^M + \mathbf{1}(s_2^a = 1)(l_{s2}^M + \mathbf{1}(A_h > 0) l_{s2}^A) \end{aligned}$$

where l_{ut}^A and l_{ut}^M is unskilled time in agriculture and manufacturing in period $t = 1, 2$; l_{st}^A is skilled time in agriculture and manufacturing. l_{y1} denotes a unit of child home production time in period 1, while l_{s1} denotes a unit of time spent at school.

Households earn income from supplying labor to the wage labor market and from cultivation if landed. They earn income m from an exogenous source.

$$\begin{aligned} y_1 &= w_{u1}(l_{u1}^A + l_{u1}^M) + \mathbf{1}(s_1^a = 1) w_{s1} l_{s1}^M + \mathbf{1}(A > 0)\Pi^A(s_1^a) + m_1 \\ y_2 &= w_{u2}(l_{u2}^A + l_{u2}^M) + \mathbf{1}(s_2^a = 1) w_{s2} l_{s2}^M + \mathbf{1}(A > 0)\Pi^A(s_2^a) + m_2 \end{aligned}$$

[11] There is no symbolic consumption of education - i.e. there is no "warm glow" of giving to one's child (Andreoni, 1989).

[12] Credit markets in developing countries offer limited opportunities to finance educational investment (Banerjee, 2004), motivating the assumption that households are credit constrained.

[13] Jayachandran (2004) presents evidence indicating that labor supply responses to a productivity shock are more inelastic in districts in which credit constraints are greater, in which a greater proportion of workers close to subsistence and in which migration costs are high. In the empirical section, I examine how allowing labor supply elasticity to vary according to labor market alters the interpretation of my results.

The household problem is given by:

$$(2) \qquad \begin{array}{rll} \max_{s_1^y} & & c_1^\rho + \beta c_2^\rho \\ \text{s.t.} & 1 & p_c c_{1p} + s_1^y C^s \leq y_1(s_1^a) \\ & 2 & p_c c_{2p} \quad \leq y_2(s_2^a) \end{array}$$

where $s_1^y = s_2^a$ - the schooling of the youth in the first period is the educational endowment of the household adult in the second period. Since education choices are discrete, households educate their children if the indirect utility from obtaining schooling is greater than that from not doing so:

$$V(s_1^y = 1) \quad \geq \quad V(s_1^y = 0)$$

$$(3) \qquad \begin{aligned} \Theta \quad = \quad & \left(\frac{1}{p_c}\right)^\rho [(y_1 - C_s)^\rho + \beta(w_{u2}l_{u2} + w_{s2}l_{s2} + \Pi^A(s_2^a = 1))^\rho] \\ & - \left(\frac{1}{p_c}\right)^\rho [(y_1 + \alpha)^\rho + \beta(w_{u2} + \Pi^A(s_2^a = 0))^\rho] > 0 \end{aligned}$$

3.2.1. *Testable Prediction 1: Education.* This model leads to a series of testable predictions for education:

1.a *Income:* A child is more likely to be educated as household income rises, ceteris paribus. This prediction is driven by the assumption that the budget constraint clears every period.

$$\frac{\partial \Theta}{\partial m_1} = \left(\frac{1}{p_c}\right)^\rho \rho((y_{h1} - C_s)^{\rho-1} - (y_{h1} + p\alpha)^{\rho-1}) > 0$$

1.b *Labor Market Returns to Education:* A child is more likely to be educated as the ratio of the skilled and unskilled wage rises, ceteris paribus.

$$\frac{\partial \Theta}{\partial (\frac{w_s^2}{w_u^2})} = \left(\frac{1}{p_c}\right)^\rho \rho w_u (\frac{w_s^2}{w_u^2} + \Pi_A(s_h^{a2} = 1))^{\rho-1} > 0$$

1.c *Returns to Education in Agriculture:* A landed household is more likely to invest in education as agricultural productivity rises, in the land wealth of the household and in their cross partial. Intuitively, agricultural productivity raises the return to education of landed households. Since landless households don't cultivate land, their returns to education do not change.[14]

$$\frac{\partial \Theta}{\partial \theta} > 0, \quad \frac{\partial \Theta}{\partial A_h} > 0, \quad \frac{\partial \Theta}{\partial \theta \partial A_h} > 0$$

1.d *Opportunity Cost:* Educational investment is decreasing in the opportunity cost of schooling, lost domestic production.

$$\frac{\partial \Theta}{\partial \alpha} = -\rho(y_h + p_c\alpha)^{\rho-1} < 0$$

Education decision rules are a function of income, as well as the returns to education. To understand how education choices respond to growth, we need to capture the responses of incomes and skilled and

[14]These predictions have been previously been tested by Rosenzweig and Foster (1996), who find that the returns to schooling increased during the green revolution in India. Areas with the greatest agrarian technical change witnessed the greatest average increases in school enrollment, conditional on the availability of a school.

unskilled wages. Households are therefore nested in a two-sector, small country general equilibrium model of rural labor markets.

3.3. **Economy Environment and Endowments.** Each region is endowed with four immobile factors of production - a working age population (P), land (A), two non-renewable raw materials (T_1 and T_2). The population consists of P households of four kinds: landed educated, landed uneducated, landless educated and landless uneducated. Land is distributed unevenly over landed households. Total factor productivity in agriculture, θ_{dt}, varies across regions and over time according to a region's agro-climatic conditions. Each region is a small, open economy - final agricultural and manufacturing goods are freely traded across regions within the country.

The marginal cost of extracting the immobile raw material decreases in the resource endowment.[15] This implies a mechanical relationship between the price of raw materials and a region's endowments. The marginal cost of extracting raw materials is denoted by $c^{rj} = c(r_j, T_j)$ for $j = 1, 2$, where r_j denotes extracted resources and T_j denotes the total regional endowment. The raw material is priced at marginal cost, $p^{rj} = c^{rj}$. The cost of extracting the resource is decreasing in the total regional endowment - $p_2^{rj} < 0$. For simplicity, the marginal cost of extracting the resource is the same across all units extracted: $p_{12}^{rj} = 0$.[16] The rents from raw materials flow outside the region.

3.4. **Agricultural Production.** Agricultural production combines labor with land to produce a single agricultural output according to a concave, CRS technology. In the absence of a land market, only landless households cultivate land.[17] Two types of labor are demanded in agriculture: unskilled and skilled. Unskilled labor undertakes physical tasks such as weeding and harvesting. This labor can be conducted by both educated and uneducated workers. Hired and family labor are perfect substitutes in unskilled work.[18] Skilled labor undertakes managerial tasks, such as deciding which varietals of seeds to sow. Skilled labor can only be conducted by landed household members.[19] Labor inputs are chosen to maximize profits, given the household's land and education endowments and the agricultural productivity faced:

$$\begin{aligned} \Pi^A &= \max_{d_u, d_s} \quad p^A F(d_u, d_s; \theta, A, s^a) - w_u d_u \\ F(d_u, d_s; \theta, A, s^a) &= \theta(d_s(\delta_1 + \delta_2 s^a))^{\gamma_1} d_u^{\gamma_2} A^{\gamma_3} \end{aligned} \tag{4}$$

where: Π^A denotes profits from cultivating land, net of hired labor expenses; d_u and d_s denotes unskilled and skilled labor; p^A denotes the (externally set) national price of agricultural output; θ is the agricultural

[15]The immobility of raw materials is for tractability. An alternative assumption, with similar testable predictions, is that prices increase in the distance from extraction, due to transportation and trade costs across regions.

[16]The simplest structure for the raw materials market is presented. The assumptions are modifiable with minimal repercussions to the main testable predictions. A more complex structure would allow for a fixed total quantity of raw materials to be consumed every period, and for convexity in the cost of extracting the materials.

[17]The REDS data support this: 5.5% of landless households cultivated land in 1999. The distinction between landed and landless households changes little over time: 5% of landless households from 1971 owned land by 1982 and 8% of adults born in landless households owned land in 1999.

[18]This assumption is made for tractability. Bharadwaj (2009) shows that hired and family labor can be considered to be perfect substitutes in some, but not all, agrarian tasks.

[19]The data motivate this assumption - less than 0.5% of hired workers in agriculture conduct managerial work.

productivity, which varies by region and over time; A is household land; s^a is the household adult's education and w_u is the labor market clearing wage of unskilled labor.

Education of the household adult raises the marginal product of skilled labor - $F_{d_s}(s^a = 1) - F_{d_s}(s^a = 0) > 0$. The productivity of unskilled labor is not rising in the education levels of individuals conducting these tasks. This implies that educated and uneducated individuals conducting unskilled work are paid the same wage, i.e. there is no return to education in the hired labor market in agriculture.[20]

I simplify the problem faced by households by assuming an interior solution for supervisorial time.[21] A household's demand for unskilled and skilled labor is given by:

$$\begin{aligned} d_u &= d_u(w_u, w_s; p^A, \theta, A, s^a) \\ &= A\left(\frac{\gamma_1^{\gamma_1}\gamma_2^{(1-\gamma_1)} p^A \theta(\delta_1 + \delta_2 s^a)^{\gamma_1}}{w_u^{1-s^a\gamma_1} w_s^{s^a\gamma_1}}\right)^{\frac{1}{\gamma_3}} \\ d_s &= d_s(w_u, w_s; p^A, \theta, A, s^a) \\ &= A\left(\frac{\gamma_1^{(1-\gamma_2)}\gamma_2^{\gamma_2} p^A \theta(\delta_1 + \delta_2 s^a)^{\gamma_1}}{w_u^{\gamma_2(1-s^a)} w_s^{s^a(1-\gamma_2)}}\right)^{\frac{1}{\gamma_3}} \end{aligned}$$

where w_s is the wage of a skilled worker in the manufacturing sector. In the absence of a manufacturing sector, this is the shadow wage of educated supervisors in the agricultural sector. The agrarian profit function can therefore be written as:

$$\Pi^A(w_u, w_s; p^A, \theta, A, s^a) = p^A F(w_u, w_s; p^A, \theta, A_h, s^a) - w_u d^u(w_u, w_s; p^A, \theta, A, s^a)$$

Household profits increase in agricultural productivity, adult education, land, the price of agricultural products and decrease in wages.

Total unskilled and skilled labor demand is given by summing labor demand across landed households:

$$\begin{aligned} D_u^A &= \sum_{h \in P_A} d_{uh}(w_u, w_s; p^A, \theta, A_h, s_h^a) \\ D_s^A &= \sum_{h \in P_A} d_{sh}(w_u, w_s; p^A, \theta, A_h, s_h^a) \end{aligned}$$

Under the assumption of a CRS technology, the distribution of land will not affect the demand for labor unless education is unevenly distributed across landed households.

3.5. **Manufacturing Production.** The manufacturing sector uses four inputs in production: unskilled labor, skilled labor and two raw materials. Only educated workers can conduct skilled manufacturing work. Manufacturing production is modelled using a Cobb-Douglas technology, which is increasing and concave in all inputs. Labor and raw materials are complements in production. There are two manufacturing industries that vary in the output shares of inputs. Therefore two industries faced with the same vector of input prices will vary in their optimal input choices.

[20]Foster and Rosenzweig (1993) find this to be the case using piece rates data from the Philippines.

[21]If the marginal product of supervisorial time is greater than the relevant wage, landed adults will only conduct own-farm supervisorial activity. This is more likely as land holdings increase and as the agrarian technological frontier shifts out.

The manufacturing sector competes with the agricultural sector for labor. Unskilled and skilled labor is remunerated at the same rate across the two sectors. Both industries choose labor and raw material inputs to maximize profits, given the prices and wages faced:

$$\begin{aligned}\Pi_j = \max_{N_{uj},N_{sj},r_{1j},r_{2j}} & \quad p^M_{Dj}F_j(N_{uj},N_{sj},r_{1j},r_{2j}) - w_uN_{uj} - w_sN_{sj} - p^{r1}r_{1j} - p^{r2}r_{2j} \\ F_j(N_{uj},N_{sj},r_{1j},r_{2j}) & = N_{uj}^{a_{1j}}N_{sj}^{a_{2j}}r_{1j}^{a_{3j}}r_{2j}^{a_{4j}}\end{aligned}$$

where Π_j denotes profits in industry j, N_{uj} and N_{sj} are unskilled and skilled labor in industry j, r_{1j} and r_{2j} denotes raw material one and two's inputs in manufacturing; p^M_{Dj} is the output price facing industry j; w_u and w_s are the unskilled and skilled wages; and p_{r1} and p_{r2} are the time invariant prices of raw materials 1 and 2.[22]

Total demand for labor and raw materials in the manufacturing sector is given by:

$$\begin{aligned}N^M_u &= \textstyle\sum_{j=1}^2 N_{uj}(w_u,w_s,p^{r1},p^{r2};p^M_{Dj}) &&= \textstyle\sum_{j=1}^2 a_{1j}w_u^{-1}y_j\prod_{s=1}^4 p_s^{a_{sj}} \\ N^M_s &= \textstyle\sum_{j=1}^2 N_{sj}(w_u,w_s,p^{r1},p^{r2};p^M_{Dj}) &&= \textstyle\sum_{j=1}^2 a_{2j}w_s^{-1}y_j\prod_{s=1}^4 p_s^{a_{sj}} \\ r^M_1 &= \textstyle\sum_{j=1}^2 r_{1j}(w_u,w_s,p^{r1},p^{r2};p^M_{Dj}) &&= \textstyle\sum_{j=1}^2 a_{3j}p_{r1}^{-1}y_j\prod_{s=1}^4 p_s^{a_{sj}} \\ r^M_2 &= \textstyle\sum_{j=1}^2 r_{2j}(w_u,w_s,p^{r1},p^{r2};p^M_{Dj}) &&= \textstyle\sum_{j=1}^2 a_{4j}p_{r2}^{-1}y_j\prod_{s=1}^4 p_s^{a_{sj}}\end{aligned}$$

Manufactured goods have a perfect substitute produced in the world market. I place myself in a small open economy setting in which the world price of the perfect substitute ties down the domestic price faced by industry j, p^M_{Dj}. τ_{jt} denotes time and industry varying tariff barriers. C_{Ljt} represents a time and industry varying per-unit cost of complying to labor and industrial regulations. Regulations and tariff barriers are the same in all industries across all economies within the country.

$$p^M_{Djt} = p^M_{Wjt} + \tau_{jt} - C^M_{Ljt} \tag{5}$$

Zero profits conditions imply that industries choose inputs until the price of output is equal to the marginal cost of production.

3.6. **Equilibrium Conditions.** Using the two labor market clearing conditions for skilled and unskilled labor, the ten first order conditions from production and the utility maximization condition, it is possible to solve for the equilibrium quantities and prices of interest, notably the wages of skilled and unskilled workers, the allocation of labor across sectors, raw material inputs and the schooling choices of households.

Equilibrium wages are determined by the labor market clearing conditions:

$$P = D^A(w^*_u,w^*_s,Z^A) + G^A(w^*_u,w^*_s,Z^A) + N^M_u(w^*_u,w^*_s,Z^M) + N^M_s(w^*_u,w^*_s,Z^M) \tag{6}$$

$$P_s = G^A_s(w^*_u,w^*_s,Z^M) + N^M_s(w^*_u,w^*_s,Z^M) \tag{7}$$

where P is the total working age population and P_s is the educated working age population. G_s is educated skilled (supervisorial) labor in agriculture. Z^A denotes determinants that only shift agricultural labor demand and only alter manufacturing labor demand through the labor market channel, such as

[22] Intuitively, this requires the decrease in the stock of raw materials due to consumption during the period to be small.

agricultural productivity. Z^M denotes determinants that only shift manufacturing labor demand, such as industrial policy.

The wages of skilled and unskilled workers are identical if the supply of educated workers at the skilled wage is greater than or equal to the demand for skilled workers at the unskilled wage. Intuitively, this is because educated workers can conduct unskilled work. If the demand for skilled workers is greater than the supply of workers at the unskilled wage, the skilled wage rises above the unskilled wage. Under these conditions, educated workers no longer conduct unskilled tasks. I focus on the more interesting case in which the wages are different and educated workers only work in the skilled labor market.

There are two regions of interest for industries in the manufacturing sector: the corner solution where production is zero and the interior solution. The corner solution occurs where the existing unskilled wage, the shadow skilled wage and raw material prices are sufficiently high that the marginal cost of the first unit of production is greater than the price of output, or where the supply of literate workers is zero. In the interior solution, manufacturing output and inputs are chosen such that price equals to marginal cost. From an initial situation where a unit of good is produced at a price greater than marginal cost, manufacturing output expands pulling labor out of agriculture and into manufacturing. This raises the marginal product of the existing labor in agriculture. This process continues until the zero profit conditions are fulfilled.

3.7. **Testable Predictions from the Model.**

3.7.1. *Prediction 2 - Manufacturing Employment.* Regional manufacturing labor demand is a function of locally determined prices, as seen in 3.5. To empirically examine how exogenous shifts in skilled and unskilled manufacturing labor demand raise wages, incomes and education, we will need to find a source of variation in labor demand that is uncorrelated with local determinants of the outcome variables. Since manufacturing labor demand is a function of locally determined prices, empirically the challenge reduces to finding variation in manufacturing employment that is uncorrelated with local determinants of wages, incomes and education.

The model suggests a source of within-district over-time variation that can be used - since industries differ in their raw material intensities, two industries faced with the same changes in policy over time will experience different changes in employment response within a region. Since agriculture and manufacturing only overlap in the labor market, changes to the wages comes through greater competition between agriculture and manufacturing in the local labor market. The model suggests a way in which the effects of skilled and unskilled shifts can be separately identified. Since industries differ in their skill intensity, the skill composition of manufacturing demand varies across regions. Variation in the timing of policy changes across industries imply that the skilled and unskilled manufacturing labor demand curves are shifted at different moments in time. These predictions constitute the base of my identification strategy. The intuition behind them is put forward in comparative statitics below.

2.a *Raw material prices:* capture the static distribution and industrial composition of skilled and unskilled manufacturing employment. Output decreases as raw material prices increase, ceteris paribus. The demand for unskilled and skilled labor in manufacturing changes as the price of a raw material increases.[23] The response of manufacturing labor demand to a change in raw material prices increases in the raw material's share in total output.[24]

$$\frac{\partial N_{uj}}{\partial p_r^n} \neq 0, \qquad \frac{\partial N_{sj}}{\partial p_r^n} \neq 0, \qquad \left(\frac{\partial N_{uj}}{\partial p_r^n}\right)\frac{\partial}{\partial a_{nj}} > 0, \qquad \left(\frac{\partial N_{sj}}{\partial p_r^n}\right)\frac{\partial}{a_{nj}} > 0$$

2.b *Policy Changes:* induce changes in skilled and unskilled manufacturing employment that vary by region. An increase in industry level tariffs raises the domestic price of a good and increases industry output. Regions vary in their output and labor demand responses to changes in industry prices due to their initial raw material endowments. Within a region, labor demand responses to changes in policy varies across industries. This implies that regions that vary in their raw material endowments see different employment shifs at a given moment in time. Output and labor demand responses to industry policy vary across districts with different raw material prices for skill level $k = s, u$ and industry $j = 1, 2$:

$$\frac{\partial N_{kj}}{\partial \tau_j} \neq 0, \qquad \frac{\partial^2 N_{kj}^M}{\partial \tau_j^M \partial p_r^n} \neq 0, \qquad \left(\frac{\partial^2 N_{kj}}{\partial \tau_j \partial p_r^n}\right)\frac{\partial}{\partial a_{kj}} \neq 0$$

3.7.2. *Prediction 3: Wage Responses to Changes in the Aggregate Economic Environment.*

2.a *Shifts in Skilled and Unskilled Manufacturing Labor Demand* driven by tariff/deregulation policy changes raise the skilled and unskilled wage, ceteris paribus. Since the agricultural and manufacturing sector only overlap in the labor market, industrial policy only affects agriculture through this market channel. A shift in unskilled manufacturing labor demand pulls unskilled labor out of agriculture, raising the marginal product of the existing labor. A shift in unskilled labor demand reduces skilled wages, since it reduces the marginal product of the remaining skilled labor in agriculture. Vice-versa for exogenous shifts in skilled labor demand.

$$\frac{\partial w_u}{\partial N_u}\frac{\partial N_u}{\partial p_{Dj}^M} > 0, \qquad \frac{\partial w_u}{\partial N_s}\frac{\partial N_s}{\partial p_{Dj}^M} < 0, \qquad \frac{\partial w_s}{\partial N_u}\frac{\partial N_u}{\partial p_{Dj}^M} < 0, \qquad \frac{\partial w_s}{\partial N_s}\frac{\partial N_s}{\partial p_{Dj}^M} > 0$$

2.b *Agricultural TFP:* raises both skilled and unskilled wage, since it raises the marginal product of labor in agriculture:

$$\frac{\partial w_u}{\partial \theta} > 0, \qquad \frac{\partial w_s}{\partial \theta} > 0$$

3.7.3. *Prediction 4: Incomes.*

[23] Two effects operate. First, an increase in raw material prices implies that, at the labor market wage, the zero profit condition no longer holds prompting a decrease in manufacturing output. Second, a change in the relative price of the raw material and labor induces substitution amongst inputs. If the former effect dominates, an increase in the price of raw materials reduces the demand for skilled and unskilled labor. This occurs if the increase in the price of raw materials reduces output sufficiently to overwhelm the increase in labor demand due to an increase in output. This is the case for both industries if the second order effect of a change in equilibrium wages due to a shift of labor out of the agricultural sector doesn't overwhelm the direct price effect.

[24] In the extreme case where the share of a raw material in total output is zero the labor demand response to a raw material price change will be zero, ceteris paribus.

3.a *Agricultural TFP:* raises both skilled and unskilled wages, as well as the returns to land. It thus raises the incomes of both landed and landless households. In the absence of land markets, cultivation profits reflect the returns to land therefore agricultural TFP raises cultivation profits.

$$\frac{\partial y_h}{\partial \theta} > 0$$

3.b *Shifts in Skilled and Unskilled Manufacturing Labor Demand:* induced by changes in industrial policy raise the incomes of landless households. The income responses of landed households vary with land endowments. The manufacturing sector competes with the agrarian sector for labor. Shifts out in manufacturing labor demand reduce labor in agriculture and decrease the marginal product of land. The net effect is negative for net importers of labor ($A > A_X$), while it is likely to be positive for small landowners.

$$\frac{\partial y_{LL}}{\partial w_u}\frac{\partial w_u}{\partial N_u}\frac{\partial N_u}{\partial p^M_{Dj}} > 0, \qquad \frac{\partial y_{L,(A>A_X)}}{\partial w_u}\frac{\partial w_u}{\partial N_u}\frac{\partial N_u}{\partial p^M_{Dj}} < 0,, \qquad \frac{\partial y_{L,(A<A_X)}}{\partial w_u}\frac{\partial w_u}{\partial N_u}\frac{\partial N_u}{\partial p^M_{Dj}} \neq 0$$

3.c *Inequality:* The difference in incomes between landless and landed households decreases in unskilled manufacturing labor demand and increases in agricultural productivity.

$$\frac{\partial (y_A - y_{LL})}{\partial \theta} > 0, \qquad \frac{\partial (y_A - y_{LL})}{\partial N_u}\frac{\partial N_u}{\partial p^M_{Dj}} < 0$$

3.8. **From Theory to Empirics.** The 4 sets of predictions put forward in the theoretical model trace out an intuitive order for approaching the empirical analysis. Since education choices vary with wages and incomes, these specifications appear last in the empirical section. In step 0 of the empirical strategy, I draw upon prediction 2 to pursue an instrumental variable strategy. In step 1, I use the estimates from my first stage regression to evaluate whether growth in agricultural productivity and predicted manufacturing employment increases unskilled and skilled wages (prediction 3). In step 3, I examine whether the two sources of growth alter household incomes differently according to their inital assets (prediction 4). In step 4 I examine the response of education to changes in wages and incomes (prediction 1). I use the estimated coefficients from the education regressions to disentangle the income and returns to education effects. Throughout the analysis, I break predicted manufacturing apart by skill to test whether the relationship between manufacturing growth, wages and incomes varies with the skill composition of the manufacturing sector.

The data used in this paper can be divided into six categories: industry-region level data on employment, wages, regional raw material endowments, industry varying factor intensities, time and industry varying policies and household education. The data will be presented as it is used. A more detailed description of the data can be found in the appendix 2.

4. Empirical Strategy

This section translates the predictions of the model into an empirical strategy.

4.1. **Empirical Specification for Wages and Income.** The model makes predictions on wage and income responses to agricultural productivity growth and shifts in manufacturing labor demand induced by changes in industry policy. This section derives the empirical specification used to test these predictions.

The model applies to a locally clearing labor market. The appropriate empirical analogue in the Indian context is a district. Migration across districts and state boundaries within India is low both in absolute terms and relative to other comparable developing countries (Munshi and Rosenzweig, 2009). In the REDS data, 8% of adult males had migrated from their villages of birth, and of these over 80% went iwthin the district. 5.9% of male respondents in the NSSO from 1999 had moved between 1990 and 1999, 55% of these movements occurred within the district. NSSO figures on temporary migration indicate that these flows are also low - in 1999, 3.06% of males had temporarily migrated for work purposes.[25]

4.1.1. *Wage Specification.* Equilibrium unskilled and skilled rural wages are determined by setting district level skilled and unskilled labor demand of equal to their labor supply. For convenience, I repeat the labor market clearing conditions from the model:

$$\begin{aligned} P &= D_u^A(w_u^*, w_s^*, Z^A) + D_s^A(w_u^*, w_s^*, Z^A) + N_u^M(w_u^*, w_s^*, Z^M) + N_s^M(w_u^*, w_s^*, Z^M) \\ P_s &= D_s^A(w_u^*, w_s^*, Z^M) + N_s^M(w_u^*, w_s^*, Z^M) \end{aligned}$$

The log-linearized supply and demand equations in the agricultural and manufacturing sectors are given by the following system of equations:

$$\begin{array}{lcllllll} \text{Labor Demand}_u^M & = & a_0 & +a_1 w_{udt} & +a_2 w_{sdt} & +a_3 X_{dt} & +a_4 Z_{dt}^M & +u_{dt}^a \\ \text{Labor Demand}_s^M & = & b_0 & +b_1 w_{udt} & +b_2 w_{sdt} & +b_3 X_{dt} & +b_4 Z_{dt}^M & +u_{dt}^b \\ & & & & & & & \\ \text{Labor Demand}_u^A & = & c_0 & +c_1 w_{udt} & +c_2 w_{sdt} & +c_3 X_{dt} & +c_4 Z_{dt}^A & +u_{dt}^c \\ \text{Labor Demand}_s^A & = & d_0 & +d_1 w_{udt} & +d_2 w_{sdt} & +d_3 X_{dt} & +d_4 Z_{dt}^A & +u_{dt}^d \\ & & & & & & & \\ \text{Labor Supply}_u & = & P_u & & & & & \\ \text{Labor Supply}_s & = & P_s & & & & & \end{array}$$

w_{udt} and w_{sdt} denote unskilled and skilled wages in district d at time t, X_{dt} denotes all common determinants of agricultural and manufacturing labor demand, Z_{dt}^M are all variables that only shift the manufacturing demand equation, conditional on equilibrium wages and X_{dt}, such as the price of raw materials and industrial policy. Z_{dt}^A refers to variables that only shift agricultural labor demand, such as agricultural TFP or the price of agricultural products.

Setting demand for skilled and unskilled workers equal to supply, we find equilibrium wages. I evaluate labor demand for manufacturing at the equilibrium wages and solve for equilibrium wages as a function of manufacturing employment:

[25]The survey asks whether individuals had temporary migrated for 60 days or more for work purposes. This question therefore elicits responses on longer-term migration episodes and may miss shorter-term higher frequency trips.

$$
\begin{aligned}
(8) \quad w^*_{udt} &= \alpha_0 + \alpha_1 E^M_{unskilled,dt}(Z^M_{dt}) + \alpha_2 E^M_{skilled,dt}(Z^M_{dt}) + \alpha_3\theta_{dt} + X_{dt}\gamma + Z^A_{dt}\eta + \delta_d + \delta_t + \epsilon_{dt} \\
(9) \quad w^*_{sdt} &= \beta_0 + \beta_1 E^M_{unskilled,dt}(Z^M_{dt}) + \beta_2 E^M_{skilled,dt}(Z^M_{dt}) + \alpha_3\theta_{dt} + X_{dt}\xi + Z^A_{dt}\lambda + \delta_d + \delta_t + \varepsilon_{dt}
\end{aligned}
$$

where w^*_{udt} and w^*_{sdt} denote equilibrium unskilled and skilled wages in district d at time t, $E^M_{unskilled,dt}$ and $E^M_{skilled,dt}$ denote equilibrium employment of unskilled and skilled workers in the manufacturing sector, θ_{dt} denotes the agricultural productivity frontier. .

Adding and subtracting $\alpha_2 E^M_{skilled,dt}$ and $\beta_2 E^M_{skilled,dt}$ from these equations and suppressing the notation indicating equilibrium outcomes, we obtain the final wage regressions:

$$
\begin{aligned}
(10) \quad w_{udt} &= \alpha_0 + \alpha_1 \hat{E}^M_{total,dt} + (\alpha_2 - \alpha_1)\hat{E}^M_{skilled,dt} + \alpha_3\theta_{dt} + A_{dt}\kappa + \delta_d + \delta_t + \epsilon_{dt} \\
(11) \quad w_{mdt} &= \beta_0 + \beta_1 \hat{E}^M_{total,dt} + (\beta_2 - \beta_1)\hat{E}^M_{skilled,dt} + \beta_3\theta_{dt} + A_{dt}\psi + \delta_d + \delta_t + \varepsilon_{dt}
\end{aligned}
$$

Appendix B.1 shows the derivation of equations (10) and (11) in greater detail and defines the α and β terms as a combination of agricultural labor demand parameters. For example, α_1 captures a non-linear combination of the own and cross-price elasticities of skilled and unskilled agricultural labor demand. The derivation of the equations in the appendix is in a more general form, in which labor supply is responsive to changes in equilibrium wages.

4.1.2. *Income Specification.* Prediction 4 implies that household incomes are affected by changes in the aggregate economic environment through two components of income: (a) labor market earnings and (b) profits from cultivation. The empirical specification used to test prediction 4 is given by the following specification, derived in appendix B.2:

$$
y_{hdt} = \gamma_0 + \gamma_1 \hat{\mathrm{E}}^M_{total,dt} + \gamma_2 \hat{\mathrm{E}}^M_{skilled,dt} + \gamma_3\theta_{dt} + \gamma_4 A_{dt} + \gamma_5 HH_{dt} + \mu_d + \delta_t + \phi_{hdt}
$$

Household level factors, such as the household's land and skill endowment, are captured by HH_{dt}.

4.1.3. *Issues with the Empirical Wage and Income Specification.* There are two reasons for keeping manufacturing employment directly in the empirical specification, rather than directly estimating the reduced form specifications. First, it sheds light on the distributional implications of an expanding non-agricultural sector in rural agrarian areas. A key characteristic of the classic models of economic growth is a movement out of agriculture and into the non-farm sector (Kuznets, 1955; Ranis-Fei, 1961; Ahluwalia, 1975). This model predicts that this transition will alter the distribution of incomes in rural areas.

Second, this approach provides insight into whether promoting the agricultural or non-agricultural sector has a greater impact on the level of wages and incomes. The ideal conceptual experiment would be to consider whether moving a portion of the workforce from agriculture to manufacturing, and vice-versa, has a greater impact on rural wages. To investigate which movement raises wages to a greater extent, we need to estimate the own and cross-price wage elasticities of labor demand in agriculture and manufacturing.

The strategy I use approaches the question from a different angle - given the tools through which policy makers promote the agricultural and manufacturing sector, do wages and incomes rise more through channels that promote the agricultural or manufacturing sector? The promotion of agricultural technical change and manufacturing jobs are major channels through which policy alters rural outcomes in India. For example, policy makers directly target job creation through manufacturing employment subsidies in West Bengal and Orissa to attract industries to rural areas. The Government of India invests extensive resources into the research and development of new seed varietals and outreach programs which promote the adoption of new technologies.[26]

Estimating these equations directly is likely to lead to biased estimates of the parameters of interest. The simultaneous determination of manufacturing employment and wages implies that any locally unobserved common determinants of agricultural and manufacturing labor demand could be driving the relationship between wages and manufacturing employment. For example, changes in the quality of electricity services may increase the demand for labor in both sectors. Therefore the OLS estimates of equations (10) and (11) will yield biased estimates of the parameters of interest.

5. Step 0: IV Strategy

This section describes the two-stage least squares strategy used to identify the causal effect of shifts in manufacturing labor demand on rural wages and incomes. The strategy used is motivated by prediction 2 of the model, which states that the manufacturing employment response to a change in industrial policy varies according to the region's endowments of the resources the industry uses intensively, i.e. if the industry is expected to initially have been located in the district. The instruments used are the interactions between the region's resource endowments, the industry's raw material intensity and the industry's policy.

Section 5.1 explains why this strategy has predictive power and section 5.2 motivates their validity. The results from the first stage estimations are presented in section 5.5.

5.1. **Why are these instruments likely to be correlated with manufacturing employment?** I combine two different literatures to explain why the set of instrumental variables are likely to be correlated with manufacturing employment. First, industry locations have been found to be related to resource and labor-market natural advantages across US states (Ellison and Glaeser, 1999; Kim, 1999). Ellison and Glaeser (1999) find that location responses to resources vary according to the intensity with which an industry uses a resource.[27] Second, reductions in industrial regulations and trade barriers between 1985 and 1997 have been found to have had large effects on manufacturing output and employment in India (Aghion et al 2009; Hasan et al, 2007; Chari, 2008; Topalova, 2004). The first-stage empirical model is built on the overlap between these two literatures. If the static distribution of industial employment in

[26]An alternative approach is to directly examine how wages respond to a change in agricultural relative to manufacturing employment. Under labor market clearing, the coefficients on agricultural and manufacturing employment will capture a measure of the labor supply elasticity which is the same for both variables.

[27]For example, the highest concentration of aluminium production, an electricity intensive industry, is found in Washington, the state with the lowest electricity prices.

Indian districts can be explained by the interaction between an industry's resource usage and regional resources, the interaction of these variables with industrial policy captures variation in employment responses across districts with different initial industrial concentrations.

5.2. **Validity of Instrumental Variables.** In this section, I discuss the validity of the TSLS strategy. Prediction 2.b states that within district employment responses across industries to changes in policy vary according to the industries' technologies. This implies the following empirical specification:

$$\begin{aligned} E^M_{total/skilled,idt} &= \beta_0 + \beta_1\tau_{it} + \beta_2\tau_{it} * s_i + \beta_3\tau_{it} * s_i * n_d + \beta_4\tau_{it} * n_d + \beta_5 s_i * n_d + \\ &\quad + \beta_6 s_i + \beta_7 n_d + \beta_8 X_{dt} + \delta_d + \delta_t + u_{idt} \end{aligned} \tag{12}$$

where $E^M_{total/skilled,idt}$ is total or skilled manufacturing employment in industry i, district d at time t, τ_{it} are industry level policies at time t, s_i is a measure of an industry's use of a given input, and n_d is a measure of the district level endowment of that input.[28] I treat the triple interaction term as excludable. [29] To consistently identify the coefficients on total and skilled manufacturing employment in the wage and income specifications, the excluded variables should be uncorrelated with unobserved determinants captured by the error terms in the outcome regressions. To separately identify the coefficients on total and skilled manufacturing employment, the instruments must also generate independent variation in these two variables. I discuss this further in section 5.3.

The framework implies that the only channel through which changes in industrial policy affects agrarian wages and incomes is through competition between agriculture and manufacturing in the rural labor market. Empirically, this implies that all policy interactions can be treated as excludable from the wage and income regressions. In the empirical strategy however all lower order terms are included in the second stage regression and only the triple interaction terms are excluded. Intuitively this allows industrial policy to alter agricultural productivity through *nationally* traded agricultural inputs produced by the manufacturing sector, such as fertilizers. This effect is allowed to vary according to the region's raw material endowments and the region's distance from trading hubs (major ports/state capitals). The identifying assumption needed is that labor is the only *locally* traded overlapping input between the agricultural and manufacturing sectors, and that output markets for neither agricultural nor manufacturing commodities clear at a local level. I examine the robustness of my results to these assumptions in section 6.5.

[28]A conceptually equivalent approach would be to aggregate industry-district level employment across industries to arrive at a district level strategy:

$$\begin{aligned} E^M_{total/skilled,dt} &= \gamma_0 + \gamma_1\overline{\tau}_t + \gamma_2\overline{\tau s}_t + \gamma_3\overline{\tau s}_t * n_d + \gamma_4\overline{\tau}_t * n_d + \\ &\quad + \gamma_5\overline{s} * n_d + \gamma_6\overline{s} + \gamma_7 n_d + \gamma_8 X_{dt} + \delta_d + \delta_t + v_{dt} \end{aligned}$$

where $\overline{\tau}_t = \frac{1}{J}\sum_{i=1}^J \tau_{it}$ captures the average level of the policy of interest in the economy at time t, $\overline{\tau s}_t = \frac{1}{J}\sum_{i=1}^J \tau_{it} * s_i$ is a weighted average of industry tariffs at time t, and $\overline{\tau s}_t * n_d = \frac{1}{J}\sum_{i=1}^J \tau_{it} * s_i * n_d$ is the interaction of resource-use weighted policy with the raw material resource endowment of the district. Aggregating tariffs across industries compresses the variation in timing and magnitude, although the conceptual source of variation is the same. For this reason, the district-industry regressions are my preferred specification. I report results from both sets of regressions.

[29]In the first stage specification, I include own-industry policy changes and omit cross-industry policy channels. For these omissions to invalidate the identification strategy, the variables capturing how industry i's tariff change affects j's employment need to be correlated with the triple interaction term *as well as* enter directly into the wage regression. Appendix B.3 discusses these concerns in greater detail and suggests alternative approaches which help validate the strategy.

The next paragraphs explain in greater detail why the lower order terms are not excludable from the wage and income regressions but the triple interaction terms are. First, district level time-invariant resource endowments are likely to be correlated with time invariant determinants of agrarian wages. For example, the mineral endowments of a region are correlated with its soil quality, an unobserved input in the agricultural production function.[30] These time invariant district level terms are thus captured by district fixed effects.

Second, average effects of policy changes are captured by time dummies. The removal of import tariff barriers has been found to alter both the price and number of imported intermediate goods sold, as well as the number of final goods produced by Indian firms (Goldberg et al, 2009). If the market for final manufacturing products is at a national rather than regional level, the agricultural input response to a change in goods markets can be captured using year fixed effects. To capture variations in a region's exposure to imported goods, I include the interaction of average tariff changes and the distance from trading hubs.

Finally, the interactions between policy and endowments capture how employment response to policy changes vary with regional resource endowments. A concern is that aggregate policy changes may affect regions differently according to their raw material endowments.[31] In addition, time trends in wages may vary with initial conditions. For example, densely wooded regions may have fewer urban settlements or the baseline size of the urban population may be associated with differential time trends in rural wages. The policy and endowment interactions are therefore include in the second stage specification. I allow for regional time trends in my specification to address the concern that districts may be experiencing different trends in wages and industrial growth.

Conditional on the lower order terms, the interaction between resource weighted industrial policy and regional resources, $\tau_{it} * s_i * n_d$, captures variation in regional employment driven by technological variation across industries and variation in the timing and depth of policy across industries. By definition, the variables inducing the identifying variation are location invariant. Therefore, these terms are unlikely to capture any local unobserved variation in baseline conditions that may be correlated with agrarian wages nor are they likely to be correlated with time trends or aggregate policy responses that vary according to district level baseline conditions. I verify whether this is the case by examining whether change in the excluded variables between 1983 and 1999 is partially correlated with baseline characteristics. The instrumental variables are not systematically and jointly correlated with the level of wages in 1980, male or female literacy or the size of the population. To address the second concern, I examine the robustness of my results to including region level growth trends and time trends interacted with baseline characteristics.

A key identifying assumption is that the agricultural and manufacturing sectors only overlap in the local labor market. The coefficients on manufacturing employment in the wage regressions are identified if

[30] Chemical and mechanical weathering at the Earth's surface decomposes rocks and the minerals within those rocks to form soil (Kessler, 1996). The soil's nutrient content, profile and composition will depend on the texture, structure, chemical and mineralogical composition of the parent material. Therefore, the underlying quality of the soil for agricultural purposes is likely to be highly correlated with the mineral composition of ores in a district.

[31] Extending the soil quality example above, if changes in aggregate policy reduce the price of rice seeds, we would expect profits in areas to be differently affected by policy changes according to whether rice is cultivable.

unobserved district-time varying determinants of agricultural productivity (such as pesticide use) are uncorrelated with the instrumental variables. The income regressions require the additional assumption that the instruments are uncorrelated with income from sources other than cultivation and hired work in manufacturing. The instrumental variables capture all industry-region level responses to a change in tariff changes, such as changes in output, input demand for all inputs other than labor and prices. If agriculture and manufacturing are linked through other locally clearing markets, the exclusion restrictions may be violated. I test the robustness of my results to other potential channels in section 6.5.

5.3. **Identification of the Effects of Skill Composition.** Shifts in labor demand are broken apart by skill by exploiting variation in the timing and degree of tariff movements across industries with different proportions of literate workers. For example, assume there are two industries. Industry 1 employs 50% skilled workers and experiences a drop in tariff barriers 3 years before industry 2, which only employs unskilled workers. The direct effect of changes in industry 1's tariffs will shift both the demand for skilled and unskilled labor in locations in which industry 1 is predicted to be present 3 years before industry 2's unskilled labor demand curves are shifted.[32]

This strategy separately identifies variation in skilled versus unskilled employment if industries vary in their technological requirements for skilled labor. The education and skill composition of the manufacturing workforce in India varies substantially by industry. While part of the variation may be attributed to variations in tasks by industry, it is also likely to depend on the relative skilled and unskilled wages in a given location. To examine how much of the variation in the proportion of skilled workers in a region-industry cell is industry specific and how much is due to regional variation in relative wages, I conduct an Analysis of Variance of region-industry level skill and education proportions using region and industry fixed effects. Industry level dummies capture variation in the proportion of skilled workers that is common to the industry across all locations (e.g. technology, type of work conducted). District level factors capture variation in the skilled proportion of the workforce that is common to all industries within that district. For example, in areas where unskilled wages are low, both relative to the labor market wage for literate individuals in that district and in absolute terms, all firms within that district will be likely to employ a greater proportion of illiterates.

Table 3 reports the results from the variance decomposition for 1983 and 1999. The industry level dummies capture 21.5% of the variance in the proportion of literates and skilled workers employed and are jointly statistically significant at a 1% level with an F-Statistic of 17.32 for literates. Therefore industry level factors explain a large and statistically significant fraction of the variance.

5.4. **Data.** I use data on industry-time level policies, industry-level raw material intensity, industry-region-time level skilled and unskilled employment and regional resources. The data are discussed in greater depth in appendix 2.

[32]Both industries will of course be affected by both sets of tariff reforms through general equilibrium effects, the most simple of which works through equilibrium wages at a district level. Focusing on the direct effect of own-industry reforms will not violate the exclusion restrictions as long as labor is the only district level interaction between agriculture and manufacturing.

Resources groupings follow the main categories of industrial usage in the Mineral Atlas of India and various NCAER Economic Plans. Raw materials are grouped into construction, metals, chemicals, ceramics, wood and energy. Mineral and metal endowments were captured by geocoding the National Mineral Atlas of India (Gov of India, 1995). Industry level policies from Aghion et al (2008) are captured using a continuous import tariff measure and a 0-1 industry deregulation measure. Industrial employment is measured using 3-digit National Industrial Classification codes (Government of India, 1987). Industry resource use is measured using the Input-Output matrix of India (1971 and 1993). I impute the share of total costs for a resource group and convert it into an above-median dummy and quantile measure.

5.5. **First Stage Regressions: Results.** In section 5.5.1, I verify the framework by examining whether regional resources and industry resource intensity predicts the static distribution of industry-region level employment. In section 5.5.2 I discuss the results from the first stage regressions.

5.5.1. *Do resource endowments help to explain the static geographic distribution of employment?* The results from the static specification are presented in column (a) of table 5. All specifications include lower order interaction terms and region fixed effects. Standard errors are clustered on a region level. The dependent variable is the log of total regional industry employment; factor intensity is measured using a dummy indicator. The variables are presented in descending order of transportability. Appendix table A.2.3 presents the density and price per tonne of raw materials. The cost to the consumer of minerals with a low unit value increases with distance to the place of use and the difficulty of transporting the material (the lower the density). Consequently, low unit value commodities are of little value unless processed close to their source. I focus on these relatively immobile inputs.

The interaction between an industry's input intensity and the region's resource endowment captures prediction 2-a: employment is more responsive to a raw material in industries that use the material more intensively than those which use it less intensively. The results indicate that the availability of bulky materials such as wood, construction minerals and ceramics minerals are positively and significantly correlated with employment. For example, in ceramics intensive industries, employment is 38% greater in areas in the top quartile of ceramics endowments than in the bottom quartile. In regions within the top quartile of ceramics materials, employment in ceramics industries is 23.1% greater than in non-ceramics industries. Similarly, in regions within the top quartile of forest density, employment in wood intensive industries is 48% higher than in non-wood intensive industries. The coefficients are largest for the commodities which are bulkiest and least cost effective to transport. By contrast, the availability of materials whose price and density combination make them more worthwhile to transport don't have a positive impact on the location of industries that use those materials most intensively. The results support the hypothesis that raw material endowments contribute to the explanation of the static distribution of employment.

5.5.2. *Does the interaction capture differential employment responses to policy changes?* Tables 5 present Fixed-Effect specifications of changes in region-industry level employment over time. Columns (b) through (e) present estimates of employment responses to changes in industrial regulations while columns

(f) through (i) presents results capturing the response to changes in import tariffs. All specifications include lower order interaction terms, region fixed effects and time variant district characteristics including a region and time varying measure of agricultural productivity and the logarithm of the population, split by landless and landownership quartiles.

The coefficients on the triple interactions confirm the hypothesis that employment responses to changes in industry time varying import tariffs and licensing reforms vary by industry resource use and a region's resource endowments. The lower order interaction terms are not shown for reasons of parsimony, therefore the heterogeneous employment response of industries across regions and industries is best illustrated by examples. In response to a 65% drop in tariffs (the average decrease between 1983 and 1999), densely wooded areas saw on average a 4% greater decline in employment than non-wooded areas. The results are driven by wood intensive industries: in densely wooded areas, employment in wood intensive industries saw employment drop by 34% more than in non-wood intensive industries. Similarly, a 65% drop in tariffs in the ceramics industry resulted in a 12% decrease in employment in areas in the top quartile of the cermaics endowment, compared to a 3% increase in areas in the bottom quartile of the ceramics endowment.

6. Step 1: Skilled and Unskilled Wage Regressions

In this section, I use the IV strategy from step 1 to estimate the response of wages to changes in instrumented skilled and unskilled manufacturing employment. In section 6.4, I interprete the results and in section 6.5, I examine the robustness of the results to alternative explanations.

6.1. **Empirical Specification.** Prediction 2 suggests that both skilled and unskilled wages increase in agricultural productivity. Unskilled manufacturing employment is predicted to raise the unskilled wage, and to do so by more than skilled employment. The skilled wage is predicted to increase in skilled employment and to do so by more than in unskilled manufacturing employment.

The empirical strategy derived in section 5 is restated below for convenience:

$$\begin{aligned} w_{udt} &= \alpha_0 + \alpha_1 \hat{E}^M_{total,dt} + (\alpha_2 - \alpha_1)\hat{E}^M_{skilled,dt} + \alpha_3\theta_{dt} + A_{dt}\kappa + \delta_d + \delta_t + \epsilon_{dt} \\ w_{mdt} &= \beta_0 + \beta_1 \hat{E}^M_{total,dt} + (\beta_2 - \beta_1)\hat{E}^M_{skilled,dt} + \beta_3\theta_{dt} + A_{dt}\psi + \delta_d + \delta_t + \varepsilon_{dt} \end{aligned}$$

where w_{udt} and w_{sdt} denote equilibrium unskilled and skilled wages in district d at time t, $\hat{E}^M_{total,dt}$ and $\hat{E}^M_{skilled,dt}$ denote instrumented total and skilled employment in manufacturing; θ denotes agricultural productivity, A_{dt} captures all other determinants of wages. District dummies absorb all location invariant factors while year dummies capture all common year effects.

As shown in appendix B.1, the structural parameters indicate that $\alpha_1 > 0$, $\alpha_1 > \alpha_2$ and $\alpha_3 > 0$; and $\beta_2 > 0$, $\beta_2 > \beta_1$ and $\beta_3 > 0$. Therefore I anticipate that the estimated coefficient on skilled manufacturing employment is negative in the unskilled wage regression, while it is positive in the skilled wage regression.

6.2. **Data.** Prediction 2 is tested using measures of wages, agricultural productivity, characteristics of the rural population and instrumented manufacturing employment. Agrarian daily wages are used to measure unskilled rural wages.[33] The median district level daily wage earned by illiterate workers in the NSSO survey is used as an alternative measure. Skilled wages are measured as the median daily wage of literate non-agricultural workers. Alternative measures include the wage of workers with primary and above primary education.[34] Skilled manufacturing employment is measured as literate workers. Alternative measures include workers with at least primary education and those in skilled occupations.

I put forward a new approach for measuring agricultural technological change. Agricultural TFP growth constitutes a moving out of the agricultural production frontier. Changes in agricultural productivity can be captured using growth in agricultural yields over time (Foster and Rosenzweig (2004), Jayachandran (2006), Lanjouw and Murgai (2009)). Yields however reflect a combination of TFP growth as well as endogenous responses of inputs to TFP growth (such as labor and irrigation).[35] Therfore using actual yields to measure agricultural production will result in biased estimates of the parameters of interest.

Agricultural TFP is measured by combining information on the technological frontier of crops across India with variation in the type of crops that can be grown across districts. The technological frontier in India is captured by the maximum of the India-wide yield in a particular crop. For example, Punjabi districts are at the rice yield frontier. The national frontier is combined with a district-level time-invariant measure of the physical suitability of crops to local agro-climatic conditions. The Global Agro-Ecological Zones (GAEZ) project run by the Food and Agriculture Organization (FAO) combines information on climatic conditions, soil quality and terrain to make an index of the suitability of a region to growing rice, maize, wheat, cotton, sugar, pulses and sorghum. For example, sorghum is not physically suited to parts of Rajasthan but is well suited to most of Andhra Pradesh. I combine the physical suitability of land to a crop with the crop level yield frontier in India to capture the maximum potential productivity of a district at a point in time, where c represents crops, d districts and t time:

$$\begin{aligned} \theta_{dt} \quad &= \quad \sum_{c}^{C} \text{Maximum National Yield in Crop}_{ct} * \\ &\quad *\text{Proportion of District Most Suited to Crop}_{cd} * \text{Crop Price}_{c1980} \end{aligned} \tag{13}$$

Weather shocks are computed using weekly rainfall data. The measures used are total monsoon rainfall, the square of total monsoon rainfall and a weather "shock" variable taking the value of -1 if rainfall is 50% below it's long term mean, and 1 if it is 50% above.

6.3. **Results.** Wage estimates are presented in tables 6 and 7. The distinction between skilled and unskilled manufacturing employment is put aside in Table 6, panel A. The results for skilled wages are

[33]Wage workers in the agricultural sector are disproportionately illiterate relative to the population average and conduct manual physical tasks, as discussed in section 2.1. This measure therefore constitutes a good proxy for the wage obtained by individuals working in unskilled occupations.

[34]Measures imputed from the NSSO data exclude manufacturing sector workers since the instrumental variables will be directly correlated with the marginal product of workers in this sector through various input channels.

[35]Foster and Rosenzweig (1996) measure technical change by estimating the agricultural technology before and after the onset of the green revolution. This approach however requires detailed cultivator level data on agricultural production over time, which is unavailable for the years I examine.

presented in table 7. In all tables, the Fixed-Effect (FE) specification is presented in column (a) and column (b) presents the Fixed-Effect Instrumental Variables (FE-IV) specification. Columns (c) through (g) present specification and data checks using the FE-IV specification. The FE coefficient captures two effects. Firstly as manufacturing employment grows, labor is drawn out of agriculture raising the marginal product of the remaining agrarian workforce. Secondly, manufacturing is attracted to areas with lower wages and wage growth profiles. The FE-IV strategy isolates the first effect; therefore the FE-IV coefficient is expected to be greater than the FE coefficient.[36]

The FE coefficient on manufacturing employment reported in column (a) of table 6-A shows a small statistically significant relationship between manufacturing employment and unskilled wages. The coefficient estimated using the FE-IV strategy in column (b) is greater than the FE coefficient, confirming the hypothesis that the estimated FE coefficient was downward biased. A 10% increase in manufacturing employment raises wages by between 1.1% and 1.7%, confirming the prediction of the model. An increase in agricultural productivity is predicted to have a positive impact on agrarian unskilled wages. I find this to be the case: a 10% increase in the technological frontier increases wages by approximately 4%. The slight increase in the coefficient on agricultural productivity between columns (a) and (b) suggests that manufacturing employment orientates itself towards areas with low agricultural productivity, where the unskilled wage is relatively low. Conditioning on region time trends in column (d) increaces the estimated wage impact slightly - this suggests that growth paths in manufacturing employment and wages are negatively correlated. The estimated coefficients are of a similar magnitude to the coefficient of 0.9% found by Rosenzweig and Foster (2004) using village level data and a different IV strategy.

Table 6-B breaks manufacturing employment down by skill. The model predicts that the coefficient on total manufacturing employment is positive, while that on skilled employment is negative - $(\alpha_2 - \alpha_1) < 0$. The FE-IV coefficient estimates in column (b) confirm these predictions: a 10% increase in purely unskilled (illiterate) manufacturing employment raises agrarian wages by 3% while an increase in purely skilled (literate) employment raises wages by 0.6%. The larger change in the coefficient estimate for unskilled manfacturing employment relative to the change in the coefficient on skilled manufacturing employment indicates that selection on the basis of the unobserved determinants of unskilled wages occurs most predominantly in the unskilled manufacturing sector. Similar results are found when using other measures of skilled employment.

Table 7 reports estimates of equation (11) in which the dependent variable is log skilled wages. In the FE specification, skilled manufacturing employment has a positive and statistically significant impact on skilled wages - a 10% increase in skilled employment raises skilled wages by 1.2% - while an increase in unskilled employment has no statistically significant effect. In the FE-IV specification, a 10% increase in skilled employment rises to 2%, while that of unskilled employment remains unchanged. The estimated coefficients confirm the predictions of the model. The direction of movement of coefficients between the FE and FE-IV specification indicates that purely unskilled manufacturing employment does not select into locations on the basis of the unobserved determinants of skilled wages, mirroring the observation for

[36]For example, in the Indian context, cross subsidization between agricultural and industrial electricity prices implies that the two rates are negatively correlated. Manufacturing is deterred from places with low agricultural electricity prices, which are in turn likely to raise productivity in agriculture.

skilled employment and unskilled wages. Agricultural productivity is positively correlated with skilled wages, although the estimated coefficient is not statistically significant.

6.4. **Interpretation of Results from Wage Regressions.** Between 1983 and 1999, agricultural potential increased by 54% or at an annualized rate of 2.51%. The estimates suggest that this raised real unskilled wages by 21.6% and accounted for just under half of the total wage growth of 52%. The impact on skilled wages was much smaller - it raised them by 7.6%, or accounted for 18% of the total skilled wage growth of 42%.

Total manufacturing employment increased by 43% over the same period, although the proportion of individuals in manufacturing has increased by only 14%. The majority of employment growth consists of unskilled employment, which grew by 67% compared to an increase of 24% in unskilled employment. Manufacturing employment growth raised the wage of rural unskilled workers by 8% between 1983 and 1999 and raised the wages of skilled workers by 5%, accounting for 15% and 14% of unskilled and skilled wage growth respectively. Had the increase in manufacturing employment consisted of only unskilled growth, the wages of rural unskilled workers would have increased by 13%.

The estimates suggest that drivers of growth vary in their impact on skilled and unskilled wages. Low-skill biased growth, such as agricultural technical change, has a greater effect on unskilled wages than on skilled wages. Skill-biased growth, such as skilled manufacturing job growth, has a greater effect on skilled wages. The estimates suggest that growth in agricultural potential and unskilled manufacturing employment *decreased* the ratio of skilled to unskilled wages by 12% and 9% respectively. Skilled manufacturing growth *increased* the ratio by 4%. Therefore, sources of low-skilled biased growth raise the levels of wages while reducing the return to education.

The relative impact of agriculture and manufacturing are opposite to those found in the existing literature. Using household panel data between 1982 and 1999, Foster and Rosenzweig (2004) find that a 74% increase in agricultural productivity resulted in a 47.7% increase in village level agricultural wages. A 900% increase in factory workers resulted in a 93% increase in wages. The point estimates in the two papers are of similar magnitudes; the primary explanation for the differences estimated is that the trends in manufacturing employment seen in the REDS data do not match those seen in any other nationally representative data-sets.

6.5. **Robustness Checks.** In this section, I verify that the coefficients estimated in the wage specifications are indeed driven by movements in labor between the agricultural and manufacturing sector and don't capture alternative explanations. The coefficients may capture overlaps between the manufacturing and agricultural sector in other locally clearing input and output markets, such as local markets for agricultural or manufacturing products (Adhvaryu et al, 2009) or groundwater inputs (Keskin, 2009). In this section, I discuss the robustness of my results to these alternative explanations.

6.5.1. *Agricultural Outputs.* If agricultural product markets clear at a district level, a reduction in industrial production may effect the marginal revenue product of labor in agriculture through agricultural prices. Demand for agricultural products may be affected through two primary channels. Firstly, there may be a direct demand effect from the agro-processing industry. Secondly, demand for products may be affected by a household level income effect.[37] To address these concerns, I examine whether the excluded variables are conditionally correlated with district level farm-harvest prices for rice, wheat, maize, cotton and sugar at a district level. Appendix table A.1.2 reports the estimated coefficients results. F-Tests over the excluded instruments confirm that they are jointly unable to explain a significant fraction of the variation in crop prices.

6.5.2. *Other Overlapping Input or Output Markets.* If manufacturing and agriculture compete for other locally clearing inputs, the excluded instruments will be correlated with unobserved local prices that directly enter into labor demand. For example, if manufacturing uses groundwater (Keskin, 2008), the instrumental variables will be correlated with the equilibrium price of water.[38] In the absence of data on groundwater prices, I examine whether my instruments are correlated with groundwater usage. Groundwater use is measured as the proportion of villages within a district whose water table is below 5 different threshold depths. An increase in the proportion of villages whose water table lies below 10 meters captures whether groundwater is extracted faster than it is being replenished. I include region time trends to capture variations with districts in groundwater usage over time in areas with common rock types.[39] The results are presented in columns (f) through (j) of appendix table A.1.2. A F-Test over the excluded set of instruments confirms that they are jointly unable to explain a significant fraction of the variation in groundwater depletion.[40] If industries vary in their intensity of water usage, an alternative test is whether the ratio of employment in water intensive industries is able to explain away the effect of predicted employment on wages. I follow Keskin (2008) in separating industries into water and non-water intensive industries.[41] The results are robust to this specification and are available upon request.

If the market for goods produced by the manufacturing sector clears at a local level, productivity in the agricultural sector will be directly affected by the policy reforms due to changes in the price of inputs. In the absence of input price data, I use the Input-Output matrix of India to divide manufacturing into industries which are backwardly linked to the agricultural sector, and those that are not. Including the ratio of predicted employment in industries linked to the agricultural sector slightly raises the magnitude of the coefficient on total manufacturing employment from 0.30 to 0.34.

[37] Atkin (2008) argues that movement restrictions for certain agricultural products imply that a district is the relevant market for these products.

[38] The bias generated by other competing inputs runs is likely to be negative - intuitively, a policy shock that shifts out industrial output and input demands in manufacturing is likely to (but does not necessarily) increase the equilibrium price of all locally clearing inputs. A decrease in the amount of land/groundwater employed in agriculture will pull down the marginal product of labor, while a decrease in labor employed will increase the marginal product of labor.

[39] Trends in groundwater depletion are likely to vary over time according to the rock type in the region (Jessoe, 2009), which is strongly correlated with mineral composition. This would normally be captured by the lower order interaction terms.

[40] Three rounds of data are required to estimate whether the excluded variables are correlated with local groundwater prices or measures of local groundwater depletion. The rationale behind this test is that in areas in which trade reform is most likely to decrease demand for water from the manufacturing sector, groundwater depletion should be lower (Keskin, 2009). Since groundwater measures are hard to come by, I use two waves of the Minor Irrigation Census of India (1993 and 2000).

[41] Cotton, wool, silk, jute textile production and paper and paper production industries are defined as water-intensive.

6.5.3. *Alternative Channels.* The estimates may also be biased due to non-market channels, for example through negative externalities such as air and water pollution. If the volume of the pollutant is positively correlated with manufacturing output, and agricultural production decreases in the pollutant, this channel will bias the coefficient downwards. It is harder to think of examples of positive externalities of industries which could be drive the results seen.

A series of specification checks are conducted to ensure the results aren't driven by the level of aggregation used. NSSO regions are made up of 4 to 7 districts similar in their agro-climatic and geographic conditions, there is also a high degree of correlation in their mineral and natural resource endowments. Therefore regional growth in manufacturing employment and its skill composition are likely to be highly correlated with their district level counterparts. In column (b) of table 6, the log of industry-region level employment is instrumented using industry-region level measures. Predicted industry-region level employment is then aggregated to a NSSO region level. Column (d) uses the district level specification, in which the log of district level manufacturing employment is instrumented using the interaction of local resource endowments with factor use weighted tariff changes as instruments. The two methods give results of similar magnitudes.[42] I use district level industry-employment data from the Economic Census. The magnitude and statistical significance of the estimated coefficient remains robust to the use of district rather than region level data - 10% increase in manufacturing employment is predicted to increase male agrarian wages by 1.4%.

Alternative wage measures are examined in column (g) of tables 6 and 7. The unskilled wage is measured using the log of the median wage of illiterate workers in the district. By including all agricultural and non-agricultural occupations, this measure provides an indication that both non-agricultural and agricultural wages for unskilled workers respond in similar ways to changes in the structure of production.[43]

Finally, to show that the mechanisms through which shifts in manufacturing labor demand affect wages correspond to my model, I additionally examine the time devoted by households to agricultural production.[44] Landless illiterate households are the most responsive to shifts in unskilled employment, while landless literate households are the most responsive to shifts in skilled employment. The total days worked in agriculture decreases among landless households in response to an increase in unskilled manufacturing employment. I find that large landed households increase own-household labor supply as unskilled manufacturing employment expands.

[42] I do a Hausman test to verify whether the results are within sampling error of each other, where the null-hypothesis is that the district-industry exogeneity assumption is correct and that this approach is more efficient. Under the null hypothesis the test statistics is distributed as a chi-squared statistics with 20 degrees of freedom; the critical value at a 5% level is therefore 31.4. The Hausman test statistic is 14.09, therefore I do not reject the null hypothesis.

[43] An alternative approach is to look at the ratio of non-agricultural and agricultural wages of unskilled workers, to examine whether they co-move. The coefficient on the shocks is statistically insignificantly different from 1, confirming this hypothesis. Results available on request.

[44] Results available upon request.

7. Step 2: Income Regressions

The results from the previous section indicate the two sources of growth induce positive responses in skilled and unskilled wages, and that they alter the return to skilled and unskilled labor. This section examines the response of the level and distribution of incomes to the two drivers of growth.

7.1. **Empirical Specification.** Prediction 4 states that household incomes responses to aggregate economic change come through two components: (a) labor market earnings and (b) cultivation profits. Agricultural productivity is predicted to raise both components, while manufacturing growth will raise labor earnings and reduce cultivation profits. Therefore the two drivers of growth are predicted to induce heterogeneous income effects according to the household's initial education and land endowments. The predictions are tested using the following specification, derived in section B.2:

$$y_{hdt} = \gamma_0 + \gamma_1 \hat{\mathrm{E}}^{Manu}_{total,dt} + \gamma_2 \hat{\mathrm{E}}^{Manu}_{skilled,dt} + \gamma_3 \theta_{dt} + \gamma_4 A_{dt} + \gamma_5 HH_{dt} + \mu_d + \delta_t + \epsilon_{hdt}$$

where HH_{dt} captures household factors, such as the household's land endowment and skill endowment. District dummies absorb location invariant factors and year dummies capture all common year effects.

Households are separated into groups according to their landholdings and the education of the household head. As such, I divide households according to the probability that they are working in the skilled and unskilled labor market, as well as whether they are earning cultivation profits. The coefficients on agricultural productivity and predicted manufacturing employment capture average income responses across households within an endowment category.

7.2. **Data.** In the absence of income data, I use consumption as a proxy. This is likely to be a better proxy for households at the bottom than at the top tail of the income distribution. Consumption per capita, district level poverty and inequality are measured using the NSSO consumption modules.[45] I follow Deaton (2003a, 2003b) in correcting the estimates for changes in the recall period used and use an alternative set of poverty lines to those put forward by the Indian Planning Commission. Poverty is measured as the head-count ratio and poverty gap and inequality is measured as the gini coefficient.[46]

Households are split into landholding groups, defined using extensive data on agricultural production in the REDS survey. Landless households are defined as those who own less than 0.1 acres of land. Landed households are separated into net exporters and importers of labor, where the cut-off is set at 2 acres.[47]

7.3. **Results.**

7.3.1. *Consumption.* The results are presented in table 8 and 9. An increase in agricultural productivity is found to raise the consumption of all households, as predicted by the model. A 10% increase in

[45] I thank Petia Topalova for sharing her district level poverty data with me.

[46] District level identifiers were not entered into the 1983 data therefore these years are not used.

[47] Since the NSSO survey has no information on inputs into cultivation, I use the 1982 wave of the REDS data to identify the land cut-off at which labor is imported. The REDS data indicate that 75% of households who own less than 2 acres of land don't hired labor, while 90% of household who own more than 2 acres of land do.

agricultural productivity raises the consumption of large landed households by approximately 1.5%, while it raises the consumption of all other households by approximately 2%.

Manufacturing growth has a heterogeneous effect on household consumption across land and skill groups. A 10% increase in unskilled manufacturing employment raises consumption in illiterate and literate households by 3.6% and 2.5% respectively. The results from section 6.3 indicate that unskilled manufacturing employment has no effect on skilled wages. The consumption response of literate landless households to changes in unskilled manufacturing employment indicates that these households earn labor income in both the skilled and unskilled labor market. In contrast, skilled manufacturing employment only has a positive consumption effect on literate landless households - skilled growth raises their consumption by 4.1%. These results indicate that illiterate households don't earn income from skilled activities, while literate landless households earn income in both the skilled and unskilled labor markets.

The model predicts that unskilled manufacturing growth unambiguously lowers the incomes of net importers of labor. The results corroborate this prediction - in contrast to other household groups, they see no positive consumption effect from manufacturing growth. By contract small landed households see a positive consumption response to unskilled manufacturing growth. As predicted, the estimated effect is smaller than that of landless households who don't experience reductions in cultivation income.

7.3.2. *Poverty.* Households declaring manual, unskilled labor as their primary source of income constitute a substantial fraction of the rural poor (Eswaran et al, 2009). Therefore determinants of unskilled wage growth are expected to decrease poverty. Columns (a) and (b) of table 10 present results from a FE specification, columns (c) through (f) present results from the FE-IV specification. In panel A the dependent variable is the headcount rate; in panel B it is the poverty gap. Since the results are similar for the two measures, I only discuss those for the headcount rate.

A 10% increase in total instrumented manufacturing employment decreases the headcount rate by 0.08 points, or by 2.2% from the 1987 mean (column c). The skill composition of manufacturing employment alters the effect: unskilled employment reduces the proportion of the population living in poverty by 0.13 points, or by 3.3%, whereas an increase in skilled employment has no impact. Agricultural productivity has a slightly larger impact than unskilled employment, mirroring the unskilled wage results - poverty is reduced by 0.14 points or by 3.8%.

7.3.3. *Inequality.* In panel C of table 10, I examine the impact of a change in manufacturing employment on the gini coefficient of inequality. The model predicts that agricultural productivity growth will raise inequality, while a shift out in unskilled manufacturing labor demand will decrease it. The estimated coefficients support the predictions of the model: a 10% increase in agricultural productivity raises the gini by 0.014 points, or by 5.4%. The effect of manufacturing growth varies by skill: a 10% increase in unskilled employment reduces the gini by 0.013 points or 4.4%, while skilled employment reduces inequality by a much smaller margin - 0.003 points or 1%. The results therefore support the prediction that agricultural productivity increases inequality.

7.4. **Interpretation of Results.** I find that households in different asset groups display heterogeneous income responses to the different sources of growth and that the drivers of growth differ in their impact on the income distribution. Data from the NSS indicate that average real household consumption per capita increased by 22% between 1987 and 1999. The head-count rate declined by 13 percentage points from 36% to 23% between 1987 and 1999 while the gini coefficient of inequality has increased by 0.03 over the period.

Agricultural productivity growth accounts for a larger fraction of consumption growth, poverty reduction and rising inequality than manufacturing job growth. It has raised the gini coefficient of inequality by 0.045, explaining 150% of rural inequality and has reduced poverty by 7 percentage points, accounting for over 53% of the total decline.

Unskilled manufacturing growth dampened the rise in inequality by 0.01. Unskilled manufacturing growth reduced poverty by 3.6 percentage points, accounting for approximately 27% of the decline. On the other hand skilled manufacturing growth has had no impact on poverty reduction. If all the growth in manufacturing employment over the period had been unskilled, poverty would instead have been reduced by 5.5 percentage points. The skill-biased employment growth witnessed in India has implied that there has been relatively little effect of changes in the size of this sector on the wages of unskilled individuals (Kochhar, 2006). This is an observation that has been repeatedly asserted in the policy literature. To my knowledge, the estimates in this paper are the first to empirically validate it.

7.5. **Robustness Checks.** If households earn profits from manufacturing, household incomes will be affected by industrial policy through the labor market as well as self-employment channel, violating the exclusion restrictions. An additional concern which alters the interpretation of the results is that the drivers of growth may capture income responses of households working in the service sector. To ascertain whether these alternative mechanisms are driving the results, I separate households according to the occupation of the household head. Households are separated into two groups: those who work on the wage labor market and those who conduct their own non-farm business activities. To examine whether the results are driven by households in the non-farm sector, I remove all households working in the service sector or who are self-employed in the manufacturing sector. The magnitudes and statistical significance of the coefficient estimates remain broadly unchanged by this exclusion, as shown in panel A of Table A.4. In panel B of Table A.5, I additionally verify whether agricultural productivity or the instrumental variables are correlated with employment in the service sector in rural areas. I separate services into construction, transport and retail (including shops, restaurants and hotels). I find no statistically significant relationship between service sector employment and my drivers of growth, indicating that the results are unlikely to be driven by this channel.

8. Step 3: Education Regressions

In the preceding two sections, the drivers of wage and income growth vary in their effect on incomes as well as on the returns to education, as captured by the differences in their estimated coefficients on unskilled and skilled wages. I use these observations to disentangle the effect of an increase in income from changes in the return to education on educational investment.

8.1. **Empirical Specification.** Prediction 4 states that the probability that an individual is educated increases in current household income, the anticipated return to schooling in the labor market (the relative wages of skilled and unskilled workers when the child enters the labor market), the anticipated return to education within the agricultural sector, and decreases in the opportunity cost of schooling. I restate the condition under which households educate their children for convenience (equation (3)):

$$S_h^y = 1 \quad \text{if} \quad \left(\frac{1}{p_c}\right)^{\rho} [(y_h^1 - C_s)^{\rho} - (y_h^1 - \alpha)^{\rho} + (w_s^2 + \Pi^A(s_h^{a2} = 1))^{\rho} - (w_u^2 + \Pi^A(s_h^{a2} = 0))] > 0$$

Schooling decisions reflect expected returns to education. To empirically estimate the determinants of education, we need to make assumptions about the formation of expectations. I assume that parents have myopic expectations about the returns to education faced by their children. They therefore believe that the future returns faced by their children are the same as those faced today:

$$\frac{w_{sd,t+1}}{w_{ud,t+1}} = \frac{w_{sdt}}{w_{udt}} + \varsigma_{hdt}$$
$$\theta_{d,t+1} = \theta_{dt} + \sigma_{hdt}$$

where $E[\varsigma_{hdt}] = 0$ and $E[\sigma_{hdt}] = 0$. I use current wages as a proxy for the wages children receive upon entering the labor market. I test the robustness of the estimates to empirically relaxing this assumption.[48] I examine whether a child starts school of not. The return I consider is therefore from becoming literate.

I specify the following linear approximation to equation (3):

$$\mathrm{S}_{hdt} = 1 \quad if \quad \kappa_0 + \kappa_1 y_{hdt} + \kappa_2(w_{sdt} - w_{udt}) + \kappa_3 \theta_{hdt} * A_h + \kappa_4 OC_{dt} + \kappa_5 \mathrm{C}_{dt}^s + \mu_h + \mu_d + \mu_t + \epsilon_{hdt} > 0$$

where S_{hdt} is a dummy variable capturing whether the individual started primary school and OC_{dt} represents the opportunity cost of schooling. Educational investment is predicted to increase in household income ($\kappa_1 > 0$) and in the labor market return to literacy ($\kappa_2 > 0$). For cultivator households, enrollment is predicted to increase directly in agricultural productivity, due to its complementarity with education in cultivation activities ($\kappa_3 > 0$). Education is predicted to decrease with the opportunity cost of schooling ($\kappa_4 < 0$) and the direct monetary cost of schooling $\kappa_5 < 0$. Education may also vary with time invariant district characteristics, such as regional preferences for education.

As noted in section (7.3.1), landless households earn income in the wage labor market and landed households additionally earn income from cultivation activities. Substituting household income into the education equation (14) and combining parameters, we get an estimating equation that is a function of the wages of skilled and unskilled workers, the cost of schooling and cultivation revenues:

$$\mathrm{S}_{hdt} = 1 \quad if \quad \kappa_0 + (\kappa_1\lambda - \kappa_2 + \kappa_4) w_{udt} + (\kappa_1 \xi + \kappa_2) w_s + \kappa_1 1(A > 0)\Pi^A + \kappa_3 \theta_{dt} * A_h + k_5 \mathrm{C}_{dt}^s + \mu_h + \mu_d + \mu_t + \epsilon_{hdt} > 0 \tag{14}$$

[48] First, I include a linear time trend interacted with the initial return to education in the education specification. As an additional test, I examine whether education choices of the younger cohort respond to future observed wages.

where λ denotes the proportion of household time spent working in the unskilled labor market and ξ denotes the proportion of household time spent in the skilled labor market. Π^A denotes revenue from household cultivation, net of hired labor and other input expenses.

The coefficients on wages and profits capture combinations of the reduced form parameters and vary across households according to their skill and land holdings. Firstly the parameters on wages vary across educated and uneducated households, since they devote different amounts of time to skilled and unskilled activities. Secondly landed households additionally earn income from cultivation activities. Cultivation revenues are themselves decreasing functions of wages. In the absence of data on cultivation revenues, the estimated coefficient on wages will additionally capture education responses to changes in revenues driven by wage variation.

Directly estimating equation (14) using OLS is unlikely to capture causal effects. For example, the ratio of skilled to unskilled wages is likely to be correlated with local preferences for education and household incomes may reflect school quality. To overcome the endogeneity bias, I examine education responses to income and returns to education variation induced by changes to agricultural productivity and instrumented manufacturing employment. As discussed in section (5.2), the excluded variables are unlikely to be partially correlated with local education characteristics. I use two approaches to estimate the coefficients in this specification: (1) an indirect least squares strategy and (2) an instrumental variables strategy.

In the indirect least squares approach, I substitute the empirical specifications for cultivation revenues and wages into equation (14). The estimating equation therefore examines education responses to changes to agricultural productivity and predicted manufacturing employment.[49]

$$\mathrm{S}_{hdt} = a_0 + a_1\hat{\mathrm{E}}_{total,dt} + a_2\hat{\mathrm{E}}_{skilled,dt} + a_3\theta_{dt} + a_4\mathrm{C}^s_{dt} + a_5 HH_{ht} + a_6 A_{d,t} + u^a \tag{15}$$

The parameter combinations vary across groups of households according to their skill and land assets. For landless households, who only earn income in the wage labor market, the estimated coefficients are combinations of the α and β terms from the wage regressions and the κ terms from the structural education specification (equation 14). For example, for landless literate households the parameters on the first three terms are given by:

$$\begin{aligned} a_1 &= (\kappa_1\lambda + \kappa_4)\alpha_1 + (\kappa_1\xi + \kappa_2)(\beta_1 - \alpha_1) \\ a_2 &= (\kappa_1\lambda - \kappa_2 + \kappa_4)(\alpha_2 - \alpha_1) + [\kappa_1\xi + \kappa_2](\beta_2 - \beta_1) \\ a_3 &= (\kappa_1\lambda + \kappa_4)\alpha_3 + (\kappa_1\xi + \kappa_2)(\beta_3 - \alpha_3) \end{aligned}$$

For landed households, the parameters additionally capture terms from the cultivation profit equation. The parameters in equation (15) therefore capture a weighted combination of the κ terms, where the weights capture the relationship between incomes, wages and the source of growth.

I use a probit and linear probability model to estimate the coefficients. The structural parameters are estimated using the Optimal Minimum Distance estimator, where the weight matrix is the inverse of the

[49]The profit specifications are discussed in greater detail in appendix B.2.

variance covariance matrix of the reduced form coefficients. The weights are the estimated α and β terms from sections 6.3 and 7.3.1. Since I am unable to empirically estimate the cultivation profit terms, I am only able to estimate the structural parameters for landless households.

I first use the estimates of a_1, a_2 and a_3 to estimate κ_1 and κ_2, setting κ_4 equal to zero. I estimate the parameters separately for literate and illiterate households, as well as by sex. Secondy, I set κ_2 and κ_4 to be equal for literate and illiterate households, and test whether κ_4 is statistically different from zero. The opportunity cost of schooling reflects productive uses of child time such as engaging in domestic production or income generating activities. The descriptive statistics in section 2.3 suggest that few children work prior to age 10. Therefore the opportunity cost of young children attending school is likely to be small; I therefore test whether $\kappa_4 = 0$.

I include district fixed effects in all specifications to capture variation across districts in unobserved time invariant determinants of education that may be correlated with the average level of manufacturing employment and agricultural productivity in a district, such as education quality. In the absence of data on the costs of schooling facing households, I include state-year interactions which capture variations in the provision and cost of education at a state level. I additionally include region time trends to capture trends in education provision within a group of districts.

Using the IV strategy, I instrument wages using agricultural productivity and the interaction of effective tariffs and district resource endowment. In the absence of data on incomes or cultivation revenues, this approach can only be used for landless households. I assume that illiterate households work only in the unskilled labor market, an assumption consistent with the results presented in section 7.3.1. Literate households work in both the skilled and unskilled labor markets. The estimating equations are therefore:

Illiterate Landless:

$$\text{S=1}_{hdt} \quad if \quad \kappa_0 + (\kappa_1 - \kappa_2 + \kappa_4)w_{udt} + \kappa_2 w_{sdt} + k_5 \text{C}^s_{dt} + \mu_d + \mu_t + \epsilon_{hdt} > 0 \tag{16}$$

Literate Landless:

$$\text{S=1}_{hdt} \quad if \quad \kappa_0 + (\kappa_1\lambda - \kappa_2 + \kappa_4)w_{udt} + (\kappa_1\xi + \kappa_2)w_{sdt} + k_5 \text{C}^s_{dt} + \mu_d + \mu_t + \epsilon_{hdt} > 0 \tag{17}$$

The three structural parameters, κ_1, κ_2 and κ_4 are estimated using the optimal minimum.

8.1.1. *Data.* To capture education choices, I use 4 waves of NSSO data between 1987 and 2004 to examine ex-post education choices between 1983 and 1999. I examine how the wages and incomes faced when children are aged 5 to 9 alter the school entry decision. For example, children aged 5 to 9 in 1983 are aged 10 to 14 in 1987-88. The education measure is thus whether children aged between 10 and 14 in 1987 ever attended school.[50]

8.1.2. *Results.* The estimates from the education specifications are presented in tables 11 and 12. Panels A and B of table 11 present the estimates from equation (15), while panel C presents the estimates from

[50] Households are asked the level of educational attainment that their child has achieved (Government of India, 2001). A child who has never attended school is reported as either illiterate or literate with no formal schooling.

equations (??) and (17). The structural parameters estimated using the minimum distance estimator are presented in table 12.

In table 11, the estimated parameters on agricultural productivity and instrumented manufacturing vary across household asset groups. In illiterate landless households, a 10% increase in agricultural productivity raises the probability that a boy attends school by 0.03. Since growth in agricultural productivity raises wages but reduces the returns to schooling, the coefficient sign indicates that the positive income effect outweighs the negative effect from a reduction in the returns to schooling. Landed households also exhibit positive education responses to agricultural productivity - a the same increase in agricultural productivity raises the probability that a boy in these households attends school by 0.03. This is likely to reflect both income and returns to education effects since agricultural productivity raises the return to primary school through managerial activities in cultivation (Foster and Rosenzweig, 1996).

In column (c), an increase in unskilled manufacturing employment raises the probability that a child in a literate household acquires education, while it reduces it in large landless households. Since both sets of households experience similar reductions in the return to literacy, the difference in education responses is likely to reflect heterogeneous household incomes responses to unskilled growth. Both literate and illiterate landless households display positive responses to increases in predicted skilled manufacturing employment. Since illiterate landless households are unlikely to be working in the literate labor market and skilled manufacturing growth has been found to have a small positive effect on unskilled wages, the positive education response in these households is likely to reflect a positive returns to education response.

Panels A and B of table A.5 examine the sensitivity of the results to using a linear probability model.

8.1.3. *Robustness Checks.* If the supply side of education responds to growth in incomes or returns, the estimated coefficients may be capturing a this supply side response. For example, Foster and Rosenzweig (2004) find that new schools are allocated to areas in which agricultural technological change is expected to be greatest.

To ascertain whether these alternative mechanisms are driving the results, table A.6 displays the response of a number of indicators of educational supply are correlated with agricultural productivity and the excluded instruments. In the absence of information on school placement, I use data on the number of teachers in a district, the ratio of teachers to pupils and the average distance to the nearest primary school as a measure of educational infrastructure (panel A, columns a-c). To capture changes in costs, I examine primary school tuition fees within a district (panel A, columns d and e) as well as the proportion of primary school aged children receiving tuition subsidies, free midday meals, books and travel subsidies and scholarships.

The results indicate that there may be some case to be made for a supply side response to the changes in manufacturing employment driven by the excluded instruments, but the response to changes in agricultural productivity appears to be minimal. In order to ensure that my results are not driven by changes in the supply of education infrastructure or education costs, I include the number of teachers, average distance from schools, the proportion of children receiving a scholarship, midday meals and books as explantory variables in the education regressions. The magnitude and statistical significance of the coefficient estimates are robust to this inclusion and are displayed in panel C, table A.5.

8.1.4. *Results - Income and Returns to Education Effects.* The parameters estimated capture combinations of the structural parameters, which are estimated using optimal minimum distance. The results are presented in table 12; columns (a) through (d) present the results for boys, while columns (e) through (h) present those for girls. In (a) and (b) the opportunity costs are constrained to zero, in (c) and (d) I test whether they are indeed zero. The probability a child starts school increases in household incomes. A 10% increase in income raises the probability that children start school by 0.11 and by 0.09 in illiterate and literate households respectively. The difference between the two effects is not statistically significant. The positive income effect may represent household credit constraints or that education enters directly into household utility. I am unable to distinguish between these two channels.

Demand for education raises in the returns to education - a 10% rise in returns to education increases the probability of attending school by 0.05 and 0.04 in illiterate and literate households. In other words, it pulls 5 in 100 children into school. The relative magnitude of the estimated coefficients indicates that education response to rising income are greater than those to returns to education. In comparison, the returns response for girls is small and insignificant. This may reflect the absence of female labor demand in skilled occupations in rural areas. Females are also more likely to marry outside of the district (Rosenzweig and Stark, 1989), therefore females may also respond to a different geographic return to education than boys.

In columns (c) and (d) of table 12, I test whether the opportunity cost effect is different from zero. I find that the opportunity cost for both boys and girls is not statistically different from zero. This is likely to reflect the observation that children who aren't at school also aren't working.

8.2. **Interpretation of Results.** The proportion of boys starting school in illiterate landless households has risen from 0.62 in 1987 to 0.81 in 1999. Consumption among these households has increased by 21%, while the average ratio of skilled to unskilled wages in rural area has decreased by 14%. The estimated coefficients therefore suggest that the rise in income during this period raises the proportion of children in these households by 0.23, while the reduction in the returns to education would reduce it by 0.07. Overall, the estimates suggest a rise in the proportion of children starting school of 0.17 over the period. A large part of this effect can be attributed to rising agricultural productivity. The estimates suggest that educational investment was raised by 0.21 through this channel, while rising unskilled manufacturing employment has raised investment by 0.07 percentage points.

The education increase among girls in this household group is even more startling: the proportion of girls attending school has increased from 0.43 in 1987 to 0.71 in 1999. The estimates suggest that growth in incomes raised the proportion of girls starting school by 0.17. The majority of this effect comes through rising agricultural productivity.

9. Conclusion

This paper makes two contributions to our understanding of how economic growth alters education choices and affects the distribution of household incomes. The first contribution is to estimate the effect of rising income and returnt to schooling on educational investment between 1983 and 1999. Male educational investment responds positively to growth in incomes and returns to education. The estimates suggest that rural income growth has raised the demand for education among primary school aged boys by 23 percentage points over the period, while decreases in rural returns to education have worked the other way. In contrast to male educational investment, female investment only responds to rising incomes.

The second contribution of this paper is to show the response of the level and distribution of incomes in rural areas due to growth in manufacturing employment. Growth in unskilled manufacturing employment reduces inequality in rural India, while agricultural productivity growth has contributed to rising inequality. Furthermore, agricultural productivity growth explains approximately half of the reducation in poverty and rise in unskilled wages seen over the period. Manufacturing growth has played a small role in increasing the levels of incomes in rural India and has only accounted for 20% of the reduction in poverty. This small impact is partly attributable to growth in skilled manufacturing employment over the period.

The results suggest that, in terms of male educational investment, a rising tide may lift all boats. Regardless of the source, growth increases educational investment at the bottom end of the income spectrum in rural India. Sources of growth that raise the unskilled wage but have little impact on the skilled wage, such as agricultural technological change and unskilled manufacturing growth, increase educational investment substantially through the income channel. At the same time, they increase household welfare by raising income and reducing poverty. Skill biased manufacturing growth, while having a smaller impact on the levels of current incomes, raises the educational investment of landless households through rising returns to education.

Table 1a: Primary Occupations for Working Individuals of Working Age, 20 to 55

Panel A: Males

	1983				1999			
Occupation	**Percent in Population**	**Proportion Illiterate**	**Proportion Primary +**	**Land Owned (acres)**	**Percent in Population**	**Proportion Illiterate**	**Proportion Primary +**	**Land Owned (acres)**
Agricultural Sector	74.72	0.57	0.29	4.99	69.8	0.44	0.43	2.77
of which…								
Hired Labor Market	30.24	0.69	0.17	1.15	33.04	0.58	0.27	0.64
Farm manager	0.02	0.35	0.47	3.62	0.02	0.49	0.40	1.86
Manual Laborer	30.22	0.70	0.16	1.17	33.02	0.56	0.28	0.64
Household Farm	44.48	0.49	0.37	7.61	36.76	0.33	0.55	4.69
Head	26.91	0.55	0.30	6.34	21.03	0.40	0.46	3.78
Other family member	17.57	0.41	0.47	9.54	15.73	0.23	0.67	5.91
Non-Agricultural (excl gov)	20.81	0.38	0.46	1.45	27.52	0.27	0.59	0.91
of which…								
Manufacturing Worker	7.07	0.37	0.45	1.05	8.00	0.27	0.58	0.89
Other	13.73	0.40	0.46	1.67	19.52	0.28	0.59	0.92
TOTAL - RURAL (excl gov)	**95.53**	**0.5**	**0.36**	**4.3**	**97.32**	**0.37**	**0.5**	**2.25**

Panel B: Females

	1983				1999			
Occupation	**Percent in Population**	**Proportion Illiterate**	**Proportion Primary +**	**Land Owned (acres)**	**Percent in Population**	**Proportion Illiterate**	**Proportion Primary +**	**Land Owned (acres)**
Agricultural Sector	86.08	0.90	0.06	4.27	83.75	0.79	0.13	2.43
of which…								
Hired Labor Market	41.21	0.93	0.04	1.18	47.37	0.83	0.09	0.73
Farm manager	0.04	0.81	0.19	0.51	0.1	0.84	0.09	0.74
Manual Laborer	99.96	0.93	0.04	1.20	99.9	0.76	0.00	2.09
Household Farm	44.87	0.87	0.08	7.10	36.38	0.74	0.19	4.64
Head	8.92	0.88	0.07	7.45	92.06	0.74	0.19	4.84
Other family member	91.08	0.81	0.12	3.54	7.94	0.75	0.16	2.32
Non-Agricultural (excl gov)	6.26	0.82	0.11	0.96	13.64	0.65	0.23	0.61
of which…								
Manufacturing Worker	6.14	0.77	0.14	0.76	7.25	0.61	0.26	0.57
Other	0.12	0.86	0.08	1.16	6.39	0.70	0.20	0.66
TOTAL - RURAL (excl gov)	92.34	0.81	0.13	4.31	97.39	0.66	0.25	2.15

TABLE 1b: Literacy in the Rural Labor Force by Industry and Occupation						
Dependent Variable: Takes a value of 1 if an individual is literate in 1987						
	Panel A: Males			**Panel B: Females**		
	(a)	**(b)**	**(c)**	**(d)**	**(e)**	**(f)**
Manufacturing	0.122***	0.120***	0.185***	0.017	0.017	0.061**
	(0.021)	(0.021)	(0.021)	(0.026)	(0.026)	(0.027)
Unskilled Manufacturing	-	-	-0.188***	-	-	-0.188***
	-	-	(0.011)	-	-	(0.011)
Own-Cultivation	0.101***	0.093***	0.093***	-0.136***	-0.136***	-0.136***
	(0.013)	(0.013)	(0.013)	(0.018)	(0.018)	(0.018)
Head*Own Cultivation	-	0.024***	0.024***	-	0.011	0.011
	-	(0.008)	(0.008)	-	(0.011)	(0.011)
Agriculutral Laborer	-0.230***	-0.226***	-0.226***	-0.254***	-0.253***	-0.253***
	(0.014)	(0.014)	(0.014)	(0.018)	(0.018)	(0.018)
Observations	91696	91696	91696	33076	33076	33076
R-squared	0.219	0.222	0.224	0.276	0.276	0.277
All columns include district identifiers and age polynomials.						

TABLE 1c: Wages by Industry and Occupation								
Dependent Variable: Log(Individual's Wage in 1987)								
	Panel A: Males				**Panel B: Females**			
	(a)	**(b)**	**(c)**	**(d)**	**(e)**	**(f)**	**(g)**	**(h)**
Literate	0.189***	0.171***	0.208***	0.167***	0.188***	0.185***	0.245***	0.154***
	-0.007	-0.007	-0.008	-0.007	(0.007)	(0.007)	(0.016)	(0.012)
Manufacturing	-	0.147**	0.112*	0.048	-	0.189***	0.245***	0.067
	-	(0.062)	(0.065)	(0.062)	-	-0.049	(0.072)	-0.068
Agricultural Sector	-	-0.042***	0.01	-	-	-0.056***	0.031	-
	-	-0.008	(0.010)	-	-	(0.008)	(0.010)	-
Literate*Manu	-	-	0.034	-	-	-	-0.267***	-
	-	-	(0.026)	-	-	-	(0.041)	-
Literate*Ag	-	-	-0.139***	-	-	-	-0.189***	-
	-	-	(0.014)	-	-	-	-0.025	-
Skilled*Manu	-	-	-	0.225***	-	-	-	0.106***
	-	-	-	(0.023)	-	-	-	(0.039)
Unskilled*Ag	-	-	-	-0.042***	-	-	-	0.006
	-	-	-	(0.008)	-	-	-	(0.010)
Skilled*Ag	-	-	-	0.325**	-	-	-	-
	-	-	-	(0.134)	-	-	-	-
Observations	23598	23598	23598	23598	28273	28273	28273	33076
R-squared	0.513	0.516	0.518	0.518	0.259	0.259	0.26	0.276
All columns include district identifiers and age polynomials.								

Table 2a: Descriptive Statistics on Individuals Working in Manufacuturing in 1987, by Occupation Type

Occupation Bracket	Proportion of		Literate		At Least Primary	
	Males	Females	Males	Females	Males	Females
White Collar	0.17	0.13	0.78	0.29	0.69	0.19
Blue-Collar, Skilled	0.48	0.32	0.72	0.33	0.56	0.26
Blue-Collar, Manual	0.35	0.54	0.49	0.23	0.35	0.15
Average In NSS	-	-	0.52	0.20	0.41	0.14

Table 2b: Tasks Done by Occupational Bracket in the Manufacturing Sector in 1987

	Non-Routine Analytic Tasks	Non-Routine Physical Tasks	Routine Physical Tasks	Non-Routine Interactive Tasks	Routine Cognitive Tasks
Panel A: Males					
White Collar	4.19	0.56	3.35	3.65	3.32
	(1.69)	(0.37)	(0.76)	(3.05)	(3.08)
Blue-Collar, Skilled	2.68	1.64	4.58	0.54	8.49
	(1.99)	(1.59)	(1.22)	(1.43)	(2.01)
Blue-Collar, Manual	2.10	1.63	4.23	0.39	6.81
	(1.26)	(1.26)	(0.99)	(0.39)	(2.56)
Average in DOT (1977)	3.84	1.45	4.02	2.01	4.94
Panel B: Females					
White Collar	4.50	0.16	3.31	3.75	1.56
	(1.08)	(0.21)	(0.49)	(3.45)	(1.68)
Blue-Collar, Skilled	1.48	0.82	4.77	0.04	8.49
	(1.86)	(0.75)	(1.21)	(0.22)	(1.83)
Blue-Collar, Manual	0.88	0.89	3.92	0.07	4.85
	(1.05)	(0.90)	(0.89)	(0.25)	(3.81)
Average in DOT (1977)	3.63	1.31	4.03	1.86	5.04

A crosswalk was created between Indian Occupational Codes (NCO-68) and the US Census Occupation Codes (1960). The Dictionay of Occupational Titles dataset was assembled by Autor, Levy, and Murnane (2003); they collected data on job task requirements from the US Department of Labor's Dictionary of Occupational Titles (DOT) and merged them with census occupation classifications. The classification of jobs into White Collar, Blue-Skilled and Blue Manual was conducted using the NCO-68. The average characteristics in the DOT is an unweighted mean across all occupations. Definitions: Non-Routine Analytical Tasks -General Educational Development, Maths; Non-Routine Physical Tasks - Eye-Hand Coordination; Routine Physical Tasks - Finger Dexterity; Non-Routine Interactive Tasks - Direction, Control and Planning; Routine Cognitive Tasks - Set Limits, Standards

TABLE 3: Decomposition of Variance in the Proportion of Educated Workers in a
Dependent Variable: Proportion of Region-Industry Workforce by education

	Literate	< Primary	< Middle	< Secondary
Panel A: 1987				
	(a)	**(b)**	**(c)**	**(d)**
Region	10.11	10.87	12.02	14.33
Industry	17.32	21.11	24.84	26.36
R-Squared	0.27	0.31	0.34	0.36
Panel B: 1993				
	(a)	**(b)**	**(c)**	**(d)**
Region	5.09	5.8	7.41	9.2
Industry	17.52	20.26	23.17	24.99
R-Squared	0.26	0.29	0.32	0.34
Obs	8372	8372	8372	8372
N Regions	52	52	52	52
N Industries	161	161	161	161

Table 4-a: Educational Enrollment and Child Labor, Childen Aged 5-9

	Boys			Girls		
	Work	Domestic	Started	Work	Domestic	Started
1983	0.03	0.00	0.71	0.03	0.04	0.48
1987	0.03	0.00	0.81	0.03	0.02	0.61
1993	0.01	0.00	0.84	0.02	0.02	0.73
1999	0.01	0.00	0.91	0.01	0.02	0.84

Table 4-b: Correlation between Educational Commencement, Wages and Returns

Panel A: Boys aged 5-9

	Proportion Started	Unskilled Wage	Skilled Wage (Literate)	Returns (Skilled/ Unskilled)
Fraction Started Primary School	1			
Unskilled Wage	0.589***	1		
Skilled Wage (Literate)	0.506***	0.498***	1	
Returns (Skilled/Unskilled Wage)	0.250***	-0.318***	0.428***	1

Panel A: Girls aged 5-9

	Proportion Started	Unskilled Wage	Skilled Wage (Literate)	Returns (Skilled/ Unskilled)
Fraction Started Primary School	1			
Unskilled Wage	0.683***	1		
Skilled Wage (Literate)	0.546***	0.498***	1	
Returns (Skilled/Unskilled Wage)	0.311***	-0.318***	0.428***	1

TABLE 5: First Stage Specification

Dependent Varible: Log(Region-Industry Total Employment)

Panel A: No Policy (column a) and Deregulation (b-e)

	Total	**Total**	**Literate**	**Primary+**	**Total**
	No Policy		**Dummy**		**Quantile**
	(a)	**(b)**	**(c)**	**(d)**	**(e)**
Policy	0.050***	-0.384	-0.954***	-1.177***	0.271
	(0.015)	(0.318)	(0.284)	(0.303)	(0.325)
Policy*Forest Cover	0.016***	0.487***	0.384***	0.338***	0.103***
*Wood Proportion	(0.005)	(0.042)	(0.039)	(0.037)	(0.013)
Policy*Ceramics Minerals	0.027***	0.026***	0.018*	0.017*	-0.001
*Ceramics Proportion	(0.004)	(0.009)	(0.010)	(0.009)	(0.002)
Policy*Construction	-0.030***	0.030***	0.031***	0.030***	0.008***
*Construction Proportion	(0.006)	(0.004)	(0.004)	(0.004)	(0.002)
Policy*Chemicals	-0.051***	-0.042***	-0.037***	-0.033***	-0.008***
*Chemicals Proportion	(0.005)	(0.006)	(0.005)	(0.005)	(0.002)
Policy*Metals	-0.001	-0.016***	-0.009*	-0.009*	-0.003*
*Metal Proportion	(0.009)	(0.005)	(0.005)	(0.005)	(0.002)
Policy*Coal	-0.315*	-0.059***	-0.047***	-0.033***	-0.016***
*Energy Proportion	(0.166)	(0.016)	(0.015)	(0.012)	(0.005)
Policy*Electricity	1.978***	-0.214***	-0.204***	-0.175***	-0.054***
*Energy Proportion	(0.071)	(0.024)	(0.024)	(0.021)	(0.009)
Constant	1.978***	13.247	12.723	14.582	11.976
	(0.071)	(15.636)	(13.916)	(14.287)	(15.795)
Observations	5856	23450	23450	23450	23450
F-Statistic - Triple Interactions	-	34.49	32.35	21.98	22.93

Panel B: Policy - Import Tariffs

	Total	**Literate**	**Primary+**	**Total**
		Dummy		**Quantile**
	(a)	**(b)**	**(c)**	**(d)**
Policy	1.039**	0.848**	0.710**	1.417**
	(0.420)	(0.333)	(0.297)	(0.636)
Policy*Forest Cover	0.336***	0.296***	0.288***	0.115***
*Wood Proportion	(0.044)	(0.043)	(0.044)	(0.020)
Policy*Ceramics Minerals	0.054***	0.034***	0.030***	0.014***
*Ceramics Proportion	(0.008)	(0.009)	(0.009)	(0.002)
Policy*Construction	0.016***	0.013***	0.011**	0.016***
*Construction Proportion	(0.004)	(0.005)	(0.004)	(0.002)
Policy*Chemicals	-0.007	-0.005	-0.002	-0.005*
*Chemicals Proportion	(0.005)	(0.005)	(0.005)	(0.003)
Policy*Metals	0.017***	0.019***	0.017***	-0.002
*Metal Proportion	(0.006)	(0.005)	(0.005)	(0.002)
Policy*Coal	0.060***	0.040***	0.033***	0.021***
*Energy Proportion	(0.012)	(0.012)	(0.009)	(0.004)
Policy*Electricity	-0.331***	-0.271***	-0.249***	-0.075***
*Energy Proportion	(0.062)	(0.049)	(0.056)	(0.026)
Constant	-0.881	-1.246	-1.008	-0.863
	(16.383)	(15.217)	(14.819)	(16.108)
Observations	22539	22539	22539	22539
F-Statistic - Triple Interactions	23.76	13.16	11.39	28.52

* significant at 10%; ** significant at 5%; *** significant at 1%

All specifications include NSS region fixed effects and lower order terms. Standard errors are clustered at the region level. Columns (b)-(d) in panel A and (a)-(c) in panel B measure the 0/1 use of a factor in an industry, where an industry is classified as using the input if the measured factor intensity lies above the median for all industries. In Column (e) in A and (d) in B, the intensity with which an industry uses a given factor is measured using quantiles.

TABLE 6: Unskilled Wage Response to InstrumentedManufacturing Employment and Agricultural Productivity							
Panel A: Dependent Variable: Log Unskilled Wages							
	FE	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV
	(a)	(b)	(c)	(d)	(e)	(f)	(g)
Log(Manufacturing Employment)	0.025**	0.144*	0.175**	0.159*	0.146**	0.118*	0.132
	(0.012)	(0.085)	(0.076)	(0.086)	(0.074)	(0.062)	(0.145)
Log(Agricultural Productivity)	0.414***	0.428**	0.443*	0.214	0.439**	0.456	0.365**
	(0.187)	(0.206)	(0.239)	(0.246)	(0.214)	(0.244)	(0.167)
Constant	-9.993***	-11.678***	-10.238***	-8.799***	-2.451***	-8.369***	-7.346***
	(2.773)	(3.056)	(3.032)	(3.694)	(0.879)	(3.107)	(2.566)
Adjusted R-Squared	0.800	0.806	0.756	0.842	0.888	0.674	0.747
District Fixed Effects	Y	Y	Y	Y	Y	N	Y
Region Time Trend	N	N	N	Y	N	N	N
First Stage F-Stat	-	34.49***	13.087***	23.57***	45.48***	34.49***	34.49***
Observations	930	930	678	930	930	930	1034

Panel B: Dependent Variable:Log Unskilled Wages, Instrumented Manufacturing Broken Apart by Skill							
	FE	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV
	(a)	(b)	(c)	(d)	(e)	(f)	(g)
Log(Manufacturing Employment)	0.029**	0.296**	0.309***	0.288*	0.299**	0.217	0.318*
	(0.013)	(0.135)	(0.138)	(0.129)	(0.120)	(0.142)	(0.165)
Log(Literate Manufacturing)	-0.022	-0.239*	-	-	-0.251**	-0.155	-0.379*
	(0.029)	(0.142)	-	-	(0.118)	(0.145)	(0.206)
Log(At Least Primary)	-	-	-0.259**	-	-	-	
	-	-	(0.151)	-	-	-	
Log(Skilled Manufacturing)	-	-	-	-0.226	-	-	
	-	-	-	(0.222)	-	-	
Log(Agricultural Productivity)	0.415*	0.354**	0.371*	0.386*	0.258	0.322	0.321*
	(0.188)	(0.206)	(0.215)	(0.216)	(0.221)	(0.250)	(0.173)
Constant	-10.947***	-9.999***	-10.667***	-10.433***	-7.319**	-6.553**	-6.900***
	(3.157)	(3.188)	(3.217)	(3.253)	(3.462)	(3.179)	(2.668)
Adjusted R-Squared	0.801	0.801	0.801	0.801	0.842	0.683	0.743
District Fixed Effects	Y	Y	Y	Y	Y	N	Y
Region Time Trend	N	N	N	N	Y	N	N
First Stage F-Stat	-	32.35***	21.98***	28.52**	11.82**	32.35***	32.35***
Observations	930	930	930	930	678	930	1034

The dependent variable is log male agrarian wages, from the Agricultural Wages of India. In column (g) of panel A, the dependent variable is the median wage of illiterates in all sectors other than manufacturing, imputed from the Employment-Unemployment Rounds collected by the NSSO. All specifications include district fixed effects (apart from (f) in both panels which includes region fixed effects), year dummies, rainfall variables (total rainfall between june and september, rainfall squared and a shock measure taking the value 1 if there is a positive rainfall shock, 0 if no shock and -1 for negative shocks), the log of male and female population for the landless and by land quantiles. Columns (b) onwards additionally control for the interaction of district level resources with average tariff and delicensing reforms. The first stage instruments are the average interaction between between district level resources, industry resource usage and industry tariffs and regulations. The number of observations in column (c) is lower than the rest because district identifiers are not available for the 1983 NSS. Column (e) in panel A uses district-industry representative data from the Economic Census for 1991 and 1998. Column (f) in both panels includes region fixed effects.

TABLE 7: Skilled Wage Response to Instrumented Manufacturing Employment and Agricultural Productivity

Dependent Variable: Log Male Skilled Wages

	FE	FE-IV	FE-IV	FE-IV	FE-IV
	(a)	(b)	(c)	(d)	(e)
Log(Manufacturing Employment)	-0.032	0.020	0.015	0.017	0.082*
	(0.021)	(0.040)	(0.042)	(0.038)	(0.041)
Log(Literate Manufacturing)	0.118***	0.203**	-	-	0.186*
	(0.041)	(0.101)	-	-	(0.112)
Log(At Least Primary Manufacturing)	-	-	0.231**	-	-
	-	-	(0.106)	-	-
Log(Skilled Manufacturing)	-	-	-	0.259*	-
	-	-	-	(0.152)	-
Log(Agricultural Productivity)	0.153	0.140	0.162	0.114	0.059
	(0.228)	(0.208)	(0.209)	(0.205)	(0.222)
Constant	-6.044*	-4.598***	-5.3061	-5.0383	-8.612
	(3.434)	(5.241)	(5.224)	(5.199)	(3.732)
Adjusted R-Squared	0.523	0.582	0.553	0.553	0.559
District Fixed Effects	Y	Y	Y	Y	Y
Region Time Trend	N	N	N	N	Y
First Stage F-Statistic	-	32.35***	21.98***	28.52***	10.46**
32.35***	1034	1034	1034	1034	1034

The dependent variable is log median male wages of literate workers in the non-agricultural sector, excluding manufacturing. This is imputed from the Employment-Unemployment Rounds collected by the NSSO. All specifications include district fixed effects, year dummies, rainfall variables (total rainfall between june and september, rainfall squared and a shock measure taking the value 1 if there is a positive rainfall shock, 0 if no shock and -1 for negative shocks), the log of male and female population for the landless and by land quantiles. Columns (b) onwards additionally control for the interaction of district level resources with average tariff and delicensing reforms. The first stage instruments are the average interaction between between district level resources, industry resource usage and industry tariffs and regulations. .

TABLE 8: Log(Consumption per capita), Instrumented Manufacturing Employment and Agricultural Productivity

Dependent Variable: Log Consumption Per Capita

	FE				**FE-IV**			
	By Land Status		**By Education**		**By Land Status**		**By Education**	
	Landless	**Landed**	**Illiterate**	**Literate**	**Landless**	**Landed**	**Illiterate**	**Literate**
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Log(Manufacturing Emp)	0.027**	0.014*	0.018**	0.014	0.302***	0.029	0.137***	0.059
	(0.011)	(0.008)	(0.009)	(0.009)	(0.104)	(0.056)	(0.047)	(0.053)
Log(Literate Manufacturing)	0.022	0.019	0.012	0.028	0.035	0.056	-0.117	0.036
	(0.020)	(0.017)	(0.018)	(0.019)	(0.086)	(0.045)	(0.095)	(0.042)
Log(Agricultural Productivity)	0.209***	0.124*	0.215**	0.159**	0.184***	0.165**	0.204**	0.206***
	(0.066)	(0.067)	(0.101)	(0.063)	(0.066)	(0.082)	(0.099)	(0.071)
Constant	1.167	2.761	2.147	1.159	0.484	1.787	2.932	-1.816
	(2.791)	(1.264)	(1.368)	(1.126)	(3.003)	(1.615)	(1.818)	(1.365)
Adjusted R-Squared	0.377	0.401	0.269	0.238	0.381	0.360	0.297	0.232
Observations	64855	127860	94650	98065	64855	127860	94650	98065

All specifications include district and year fixed effects, rainfall variables (total rainfall between june and september, rainfall squared and a shock measure taking the value 1 if there is a positive rainfall shock, 0 if no shock and -1 for negative shocks), the log of male and female population by land quantiles. Column (e) onwards additionally controls for the interaction of district level resources with average tariff and delicensing reforms. Household controls are the sex of the household head, his age, age squared and social group. The first stage instruments are the interaction between between district level resources, industry resource usage and industry tariffs and regulations. The F-Stats on the first stage are betwen 21.98 and 32.35.

TABLE 9: Consumption, Instrumented Manufacturing Employment and Agricultural Productivity

PANEL A: Landless Threshold is <0.25 acres, Small Landed <2, Large Landed >2

Dependent Variable: Log Consumption Per Capita

	FE-IV							
	Landless		**Landed**		**Small Landed**		**Large Landed**	
	Illiterate	**Literate**	**Small**	**Large**	**Illiterate**	**Literate**	**Illiterate**	**Literate**
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Log(Manufacturing Employment)	0.370**	0.284**	0.127***	0.022	0.149***	0.097**	0.136	0.014
	(0.164)	(0.126)	(0.038)	(0.098)	(0.043)	(0.042)	(0.102)	(0.091)
Log(Literate Manufacturing)	-0.288*	0.185*	-0.154**	0.065	-0.132	-0.139*	-0.056	0.173*
	(0.171)	(0.099)	(0.078)	(0.064)	(0.097)	(0.083)	(0.070)	(0.071)
Log(Agricultural Productivity)	0.202***	0.156*	0.165*	0.201**	0.132	0.171*	0.215**	0.197*
	(0.076)	(0.086)	(0.099)	(0.088)	(0.111)	(0.091)	(0.101)	(0.108)
Constant	-1.634	4.302	0.946	2.927*	3.325*	-1.472	3.604	1.677
	(1.265)	(2.036)	(1.712)	(1.722)	(1.839)	(1.662)	(2.361)	(1.804)
Adjusted R-Squared	0.318	0.327	0.338	0.365	0.322	0.336	0.378	0.304
District Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Observations	36042	28813	52787	75073	27026	25761	31582	43491

PANEL B: Examining Sensitivity to Land Thresholds

Dependent Variable: Log Consumption Per Capita

	Land Thresholds							
	<2.5 acres	**<3 acres**	**<3.5 acres**	**<4 acres**	**>2.5 acres**	**>3 acres**	**>3.5 acres**	**>4 acres**
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Log(Manufacturing Employment)	0.131***	0.126***	0.128***	0.133***	0.018	0.026	0.062	0.080
	(0.038)	(0.036)	(0.037)	(0.037)	(0.095)	(0.103)	(0.100)	(0.095)
Log(Literate Employment)	-0.159**	-0.134**	-0.124*	-0.124*	0.064	0.073	0.036	-0.021
	(0.077)	(0.074)	(0.074)	(0.075)	(0.069)	(0.076)	(0.073)	(0.064)
Log(Agricultural Productivity)	0.171*	0.192*	0.215**	0.238**	0.169*	0.212**	0.176*	0.168*
	(0.099)	(0.100)	(0.103)	(0.104)	(0.091)	(0.093)	(0.095)	(0.100)
Constant	1.489	1.454	1.165	0.996	3.344*	2.740	3.037*	4.510**
	(1.812)	(1.829)	(1.887)	(1.899)	(1.738)	(1.808)	(1.821)	(1.929)
Adjusted R-Squared	0.319	0.318	0.313	0.309	0.362	0.360	0.360	0.361
District Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Observations	64605	73213	78709	84354	63812	55204	49708	44063

All specifications include district and year fixed effects, rainfall variables (total rainfall between june and september, rainfall squared and a shock measure taking the value 1 if there is a positive rainfall shock, 0 if no shock and -1 for negative shocks), the log of male and female population by land quantiles. In addition, they control for the interaction of district level resources with average tariff and delicensing reforms. Household controls are the sex of the household head, his age, age squared and social group. The first stage instruments are the interaction between between district level resources, industry resource usage and industry tariffs and regulations. The F-Stats on the first stage are betwen 21.98 and 32.35.

TABLE 10: Poverty and Inequality estimates

Panel A: Dependent Variable: Headcount Rate

	FE			FE-IV		
		Rural			Urban	
	(a)	(b)	(c)	(d)	(e)	(f)
Log(Manufacturing Employment)	-0.002	-0.002	-0.088*	-0.127**	0.011	-0.012
	(0.008)	(0.003)	(0.045)	0.059	(0.012)	(0.021)
Log(Literate Manufacturing)	-	0.003	-	0.131*	-	0.027
	-	(0.003)	-	(0.071)	-	(0.023)
Log(Agricultural Productivity)	-0.136*	-0.136*	-0.153*	-0.139*	-0.092	-0.087
	(0.082)	(0.081)	(0.080)	(0.082)	(0.061)	(0.061)
Constant	2.360**	2.374***	2.737**	2.337*	1.353	1.203
	(1.259)	(1.259)	(1.252)	(1.301)	(1.058)	(1.041)

Panel B: Dependent Variable: Poverty Gap

	FE			FE-IV		
		Rural			Urban	
	(a)	(b)	(c)	(d)	(e)	(f)
Log(Manufacturing Employment)	-0.003	-0.004	-0.0274	-0.069***	0.011	-0.012
	(0.002)	(0.003)	(0.021)	(0.020)	(0.012)	(0.021)
Log(Literate Manufacturing)	-	0.001		0.077***		0.027
	-	(0.005)		(0.026)		(0.023)
Log(Agricultural Productivity)	-0.042*	-0.041*	-0.052**	-0.059**	-0.092	-0.087
	(0.023)	(0.023)	(0.024)	(0.025)	(0.061)	(0.061)
Constant	0.629	0.608	0.606**	0.715*	0.667	0.609
	(0.396)	(0.408)	(0.285)	(0.395)	(0.720)	(0.718)
Observations	967	967	967	967	967	967

Panel C: Dependent Variable: Gini Measure of Inequality

	FE		FE-IV		
	(a)	(b)	(c)	(d)	(e)
Log(Manufacturing Employment)	0.002	0.002	-0.037**	-0.123***	-0.127***
	(0.002)	(0.002)	(0.018)	(0.027)	(0.025)
Log(Literate Manufacturing)	-	0.003	-	0.096***	-
	-	(0.007)	-	(0.021)	-
Log(Primary and above Manufacturing)	-	-	-	-	0.107***
	-	-	-	-	(0.021)
Log(Agricultural Productivity)	0.150*	0.153**	0.165**	0.159**	0.159**
	(0.078)	(0.078)	(0.083)	(0.080)	(0.079)
Constant	-1.769*	-1.827*	-1.939*	-1.774*	-1.722*
	(1.018)	(1.022)	(1.029)	(0.974)	(0.978)
District Fixed Effects	Y	Y	Y	Y	Y
	967	967	967	967	967

All specifications include district fixed effects, year dummies, region time trends, rainfall variables (total rainfall between june and september, rainfall squared and a shock measure taking the value 1 if there is a positive rainfall shock, 0 if no shock and -1 for negative shocks), and the log of male and female population for the landless and by land quantiles. Columns (b) through (f) additionally control for the interaction of district level resources with average tariff and delicensing reforms. The first stage instruments are the average interaction between between district level resources, industry resource usage and industry tariffs and regulations. The F-Stats on the first stage are betwen 21.98 and 32.35. * significant at 10%; ** significant at 5%; *** significant at 1%; Huber-White standard errors are reported in parentheses, standard errors are clustered at a district level in columns (a) and at a region-year level in columns (b) through (f).

Table 11: Started School Between the Age of 5 and 9

Specification: Probit

Panel A: Dependent Variable: 1 if boy reports having started school

	Landless		Landed	
	Illiterate	Literate	Net Exporters	Net Importers
	(a)	(b)	(c - L)	(c - H)
Log(Manufacturing Employment)	0.251***	0.073**	-0.178***	-0.028
	(0.095)	(0.031)	(0.051)	(0.047)
Log(Manufacturing Literate)	-0.146**	0.212***	0.050	0.076
	0.071	(0.049)	(0.032)	(0.046)
Log(Agricultural Productivity)	0.418**	0.124**	0.234***	0.251***
	(0.177)	(0.051)	(0.074)	(0.043)
District Fixed Effects	Y	Y	Y	Y
State*Year Dummies	Y	Y	Y	Y
Region Time Trends	Y	Y	Y	Y
Adjusted R-Squared	0.144	0.128	0.136	0.202
Observations	14207	12524	24421	41315

Panel B: Dependent Variable: 1 if girl reports having started school

	Landless		Landed	
	Illiterate	Literate	Net Exporters	Net Importers
	(a)	(b)	(c - L)	(c - H)
Log(Manufacturing Employment)	0.084*	0.237***	0.289***	-0.093**
	(0.051)	(0.048)	(0.082)	(0.041)
Log(Manufacturing Literate)	-0.029	-0.214***	-0.259***	0.085
	(0.047)	(0.045)	(0.055)	(0.041)
Log(Agricultural Productivity)	0.252*	0.143	0.379***	0.276***
	(0.152)	(0.091)	(0.127)	(0.078)
District Fixed Effects	Y	Y	Y	Y
State*Year Dummies	Y	Y	Y	Y
Region Time Trends	Y	Y	Y	Y
Adjusted R-Squared	0.223	0.223	0.257	0.267
Observations	13849	12865	20983	35845

Panel C: Specification - Instrumental Variables

Dependent Variable: 1 if individual reports having started school

	Boys Aged 5 to 9		**Girls Aged 5 to 9**	
	Illiterate	Literate	Illiterate	Literate
	(a)	(b)	(c)	(d)
Log(Unskilled Wages)	0.534*	0.429*	0.254	0.269
	(0.306)	(0.236)	(0.332)	(0.312)
Log(Skilled Wages)	0.263*	0.287**	0.126	0.536
	(0.143)	(0.133)	(0.270)	0.407
District Fixed Effects	Y	Y	Y	Y
State*Year Dummies	Y	Y	Y	Y
Region Time Trends	Y	Y	Y	Y

All specifications include district fixed effects, year dummies, region time trends, rainfall variables (total rainfall between june and september, rainfall squared and a shock measure taking the value 1 if there is a positive rainfall shock, 0 if no shock and -1 for negative shocks), and the log of male and female population for the landless and by land quantiles. Household controls are the sex of the household head, his age, age squared and social group. The first stage instruments are the average interaction between between district level resources, industry resource usage and industry tariffs and regulations. The F-Stats on the first stage are between 10.5 and 14.1. * significant at 10%; ** significant at 5%; *** significant at 1%; Huber-White standard errors are reported in parentheses, standard errors are clustered at a region-year level in panels A and B and at a district level in panel C.

Table 12: Structural Parameters of Education Decision Rule for children aged 5 to 9

Panel A: Setting Opportunity Cost Equal To Zero, Tau=0.4 (Proportion of Unskilled Time)

	Boys				**Girls**			
	Indirect Least Squares		**IV**		**Indirect Least Squares**		**IV**	
	Illiterate	Literate	Illiterate	Literate	Illiterate	Literate	Illiterate	Literate
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Income	1.079**	0.904***	0.797	0.501**	0.791**	1.076***	0.374	0.500
(Kappa 1)	(0.416)	(0.202)	(0.539)	(0.255)	(0.341)	(0.338)	(0.572)	(0.446)
Returns to Education	0.497**	0.442**	0.263	0.279*	0.220	-0.016	0.120	0.348
(Kappa 2)	(0.213)	(0.134)	(0.182)	(0.143)	(0.175)	(0.199)	(0.408)	(0.216)
Xi-Squared	1.18	2.059	-	-	3.114	13.827	-	-
Degrees of Freedom	1	1	0	0	1	1	0	0
Tau	-	0.4	-	0.4	-	0.4	-	0.4

Panel B: Tau=0.6/0.8; Testing if Opportunity Cost=Zero,

	Boys		**Girls**		**Boys**		**Girls**	
	Literate	Literate	Literate	Literate	Literate	Literate	Literate	Literate
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Income	0.919***	0.920***	1.079***	1.079***	0.935***	0.879*	0.791**	0.534
(Kappa 1)	(0.203)	(0.204)	(0.346)	(0.346)	(0.197)	(0.495)	(0.341)	(0.456)
Returns to Education	0.266***	0.634***	0.226	0.442	0.497***		0.2202	
(Kappa 2)	(0.100)	(0.168)	(0.265)	(0.321)	(0.213)		(0.178)	
Opportunity Cost	-	-	-	-	-0.225		0.2572	
(Kappa 4)	-	-	-	-	(0.394)		(0.533)	
Xi-Squared	2.061	2.075	0.583	0.583	2.896		7.272	
Degrees of Freedom	1	1	1	1	2	2	2	2
Tau	0.6	0.8	0.6	0.8	0.4	0.4	0.4	0.4

Table contains minimum distance estimates of the structural parameters, based on auxiliary parameters are found in table 10. In columns (e), (g), (f) and (h) of panel B, the opportunity cost of education has been set to zero. Tau indicates the proportion of time spent by skilled households in the unskilled labor market.

Appendix A. Data

A.0.1. *Employment and Education Data:* The Employment-Unemployment surveys conducted by the National Sample Survey Organization (NSSO) are the main data source for all employment and education data. I use the "thick"employment-rounds (round 10) conducted in 1983, 1987-88, 1993-94 and 1999-2000. The surveys collect information on approximately 75,000 rural and 45,000 urban households and usually cover all states in India. Employment by industry is constructed at a region level, where a regions consist of 4 to 6 neighboring districts within the same agro-climatic zone.

I use two different education thresholds to define a skilled group. These thresholds are determined by the data: in 1983, approximately 40% of manufacturing employees were classified as illiterate, and 59% had less than primary education. The employment surveys ask households and individuals within household to list their primary and subsidiary usual occupations as well as their primary and subsidiary current occupations. I use primary occupation status to define an individual's sector and restrict the sample to individuals of working age, between 25 and 55.[51] A weekly time-use recall allows me to capture the days of work devoted to agricultural labor market, own farm and manufacturing activities over the course of the preceding week.[52]

I use employment data from two waves of the Economic Census, a country-wide census of all economic activities other than those related to crop production and plantation. It contains basic data on the principal activity conducted by the firm and the number of family and hired male and female labor employed.[53] Since it is a census it captures district-industry level data on employment for both the formal and informal sector.

My measure of education is whether a child aged 10-14 is reporter to be illiterate. The NSSO classifies the education of an individual according to the highest standard that they have obtained. Therefore in the case of an individual who has attended school but cannot read or write, they should be classified as having some formal education (Government of India, 2001). To verify whether this is the case, the 55th round of the NSSO survey conducted in 1999 asks children to report whether they have ever attended school. Of the 33% of illiterate 9 to 14 year olds, 92% had never attended school while 3% attended school but dropped out. This provides some reassurance that this measure is capturing individuals who decided not to enroll in school.

A.0.2. *Wages.* Wage data on agricultural wages come from the Agricultural Wages in India (AWI) collected by the Directorate of Economics and Statist at the Ministry of Agriculture. Data are collected on a monthly basis for various agricultural operations, such as ploughing and sowing. Where possible, separate wages are collected by sex and for children. Data are not collected for all districts within a state and data reporting is uneven across time and districts. The missing observations for male wages

[51] I test the sensitivity of my results to including all working aged individuals over the age of 25. The magnitude and sign of the coefficient estimates in the wage equations are largely unchanged.

[52] It is not necessary to account for seasonal variation when using these measuring since surveying is spread out uniformly over the calendar year.

[53] The Economic Census is conceived as a frame of non-farm firms for conducting detailed follow-up surveys on the unorganized sector.

however do not appear to be related to any characteristic of local labor markets but rather to failures in the data collection system.[54] Wages are deflated using the State level Rural Agricultural Labor Price Index published by the Indian Labor Bureau. I additionally impute literate non-farm wages from weekly wage labor income in the NSSO Employment surveys. This data is still being processed and will be used in the next draft of the paper.

A.0.3. *Natural Resources.* Raw material endowments are grouped according to the main categories of industrial usage following the Mineral Atlas of India and various NCAER Economic Plans. They are listed below.

- Raw Material Groups using definitions from Mineral Atlas of India/NCAER Economic Plans (1960s-1990s)
  - *Forestry* : Proportion of District Covered in Forests
  - *Metal*: Alumnium, Chromium, Copper, Iron Ore, Lead, Manganese, Zinc.
  - *Ceramics:* Kaolin, Feldspar, Glass and Foundry Sand.
  - *Construction Sector:* Calcite, China Clay, Limestone, Sandstone.
  - *"Strategic"* Chemicals - Asbestos, Baryte, Dolomite, Fluorite and Limonite.
  - *Energy:* Coking and Non-Coking Coal and Electricity Prices

To measure the mineral and metal endowments, I have geocoded the National Mineral Atlas and Geological Map of India published by the Geological Survey of India (GSI). Figure A.4 displays the data in its raw form. I define a district's raw material endowment using all resources found within the district boundaries.

Region level data on soil conditions and agricultural yields comes from the India Agriculture and Climate Dataset compiled by Robert E. Evenson and James W. McKinsey, Jr., using data from the Directorate of Economics and Statistics within the Indian Ministry of Agriculture. Wood endowments are captured using the proportion of the district covered by forest, from the Forest Survey of India (1993).

A.0.4. *Factor Intensity.* Data from the Input-Output Matrix is used to build measures of the intensity of use of different inputs by industry. The measures reflect the share of costs accounted for by a given category of inputs. The definition of factor intensity used is given by:

$$\text{Factor Intensity}_{k,l} = \frac{\text{Cost of Input k}}{\text{Total Cost}}$$

I use discrete measures which are designed to capture the broader technological requirements of an industry. The quantile measure separates all non-zero factor intensities into quantiles, while the dummy measure codes an industry as using a factor if it lies above the median factor intensity for that raw material among all non-zero factor intensities. Table A.3 lists the industries which lie in the top quintile of the factor intensity measure for wood, ceramics, chemicals and metal inputs. While there are a few surprises, the list largely conforms to expectations: the iron producing and metal extracting industries

[54]The actual data collection is left to individual states, who assign village level officials to collect the locally common current wage on a monthly basis (Himanshu, 2004). Since there appears to be no check or enforcement of data collection at a state level by the Directorate of Economics and Statistics, data is regularly missing at a district or even state level.

are metal intensive; the glass, cement and refractory industries are ceramics minerals intensive while the aluminium processing, chemicals and coke oven industries are energy intensive.

A.0.5. *Industrial Policy and Regulations.* I use tariff measures and delicensing reforms compiled by Aghion, Burgess, Redding and Zillibotti (2008). Post Independence Indian Industrial Policy consisted of state directed, centrally planned development strategies. The "Licence Raj" entailed a series of rigid controls over the establishment, capacity, investment and production of the industrial sector (Srinivasan, 2000). Import-substitution policies, aimed at stimulating growth, gave rise to a restrictive trade regime with high average and peak nominal tariffs and non-tariff barriers. Reforms during the 1980s and 1990s represented a major break from the previous approach (Srinivasan, 2000). The first phase of reforms occurred in 1985, after a political crisis sparked by the assassination of the incumbent prime minister. A subset of manufacturing industries was removed from the jurisdiction of the license regime whilst others were allowed more flexibility in their functioning. The government adopted an expansionary fiscal stance leading to large fiscal deficits, high levels of external debt and falling foreign reserves by 1991. A serious macroeconomic and balance of payments crisis ensued; in the wake of this crisis, systemic structural reforms were introduced (Krueger and Chiony, 2002). Reforms to industrial policy and excise tariffs altered the institutional framework and profitability of private sector firms, irrespective of location. Fiscal reforms initiated by the central government resulted in substantial reductions in customs duties. Both average and peak tariffs have been drastically reduced since 1990-91, from an average nominal tariff of 125% to 25% in 1997-98. Excise duties were simplified and harmonized across products.

Appendix B. Derivations

B.1. **Obtaining the Wage Equations.** Linearizing labor demand and supply for manual and skilled workers, we get:

$$\begin{aligned}
\text{Labor Demand}^{Manuacturing}_{unskilled} &= a_0 + a_1 w_{udt} + a_2 w_{sdt} + a_3 X_{dt} + a_4 Z^M_{dt} + u^a_{dt} \\
\text{Labor Demand}^{Manuacturing}_{skilled} &= b_0 + b_1 w_{udt} + b_2 w_{sd,t} + b_3 X_{dt} + b_4 Z^M_{dt} + u^b_{dt} \\
\text{Labor Demand}^{Agriculture}_{unskilled} &= c_0 + c_1 w_{udt} + c_2 w_{sdt} + c_3 X_{dt} + c_4 Z^A_{dt} + u^c_{dt} \\
\text{Labor Demand}^{Agriculture}_{skilled} &= d_0 + d_1 w_{udt} + d_2 w_{sdt} + d_3 X_{dt} + d_4 Z^A_{dt} + u^d_{dt} \\
\text{Labor Supply}_{unskilled} &= f_0 + f_1 w_{udt} + f_2 w_{sdt} + f_3 X_{dt} + f_4 Z^S_{dt} + u^d_{dt} \\
\text{Labor Supply}_{skilled} &= g_0 + g_1 w_{udt} + g_2 w_{sdt} + g_3 X_{dt} + g_4 Z^S_{dt} + u^d_{dt}
\end{aligned}$$

Setting labor demand equal to labor supply for both manual and skilled labor, but keeping manufacturing labor demand in it's raw form, we get wages for skilled and manual labor:

$$\begin{aligned}
w_{udt} &= \left(\frac{1}{f_1 - c_1}\right)[(c_0 - f_0) + (c_2 - f_2) w_{sd,t} + (c_3 - f_3) X_{dt} + c_4 Z^A_{dt} - f_4 Z^S_{dt} + \mathrm{L}^{Manufacturing}_{manual,dt} + u^c_{dt} - u^f_{dt}] \\
w_{sdt} &= \left(\frac{1}{g_2 - d_2}\right)[(d_0 - g_0) + (d_1 - g_1) w_{md,t} + (d_3 - g_3) X_{dt} + d_4 Z^A_{dt} - g_4 Z^S_{dt} + \mathrm{L}^{Manufacturing}_{skilled,dt} + u^f_{dt} - u^c_{dt}]
\end{aligned}$$

Solving the system of simultaneous equations, we get:

$$\begin{aligned} w_{udt} &= \alpha_0 + \alpha_1 \mathrm{E}^M_{unskilled,dt} + \alpha_2 \mathrm{E}^M_{skilled} + \alpha_3 \text{Ag Productivity}_{dt} + \alpha_4 Z^A_{dt} + \alpha_5 Z^S_{dt} + \alpha_6 X_{dt} + \epsilon_{dt} \\ w_{sdt} &= \beta_0 + \beta_1 \mathrm{E}^M_{unskilled,dt} + \beta_2 \mathrm{E}^M_{skilled} + \beta_3 \text{Ag Productivity}_{dt} + \beta_4 Z^A_{dt} + \beta_5 Z^S_{dt} + \beta_6 X_{dt} + \varepsilon_{dt} \end{aligned}$$

where $\xi = \left(\frac{(g_2-d_2)(f_1-c_1)}{(g_2-d_2)(f_1-c_1)-(c_2-f_2)(d_1-g_1)}\right)$, $\alpha_1 = \xi * \frac{1}{f_1-c_1}$, $\alpha_2 = \alpha_1 * \left(\frac{c_2-f_2}{g_2-d_2}\right)$. $\beta_2 = \xi * \frac{1}{g_2-d_2}$ and $\beta_1 = \beta_1 * \left(\frac{d_1-g_1}{f_1-c_1}\right)$.

Under the assumption that unskilled and skilled labor demand are complements in agricultural and manufacturing, $c_1 < 0$, $c_2 < 0$, $d_1 < 0$ and $d_2 < 0$. If total labor supply is constant, it follows that $f_1 + g_1 + f_2 + g_2 = 0$ and that $f_1 > 0$, $g_1 < 0$, $f_2 < 0$ and $g_2 > 0$. $\xi > 0$ if the own-price labor supply and demand response is greater than the cross-price labor demand and supply response for both unskilled and skilled labor. Therefore $\alpha_1 > 0$ and $\beta_2 > 0$. Under the same conditions, $\alpha_2 < \alpha_1$ since $f_1 + f_2 > 0$ and $c_1 - c_2 < 0$ therefore $f_1 - c_1 > -c_2 - f_2 > c_2 - f_2$. Similarly $\beta_1 < \beta_2$.

α_1 decreases in the wage elasticity of agricultural unskilled labor demand - intuitively, as agricultural labor demand becomes less responsive to changes in wages, a larger increase in agrarian wages is required to "release" a given quantity of labor from agriculture. α_1 increases as the unskilled wage elasticity of unskilled labor supply decreases - the intuition is that if manual labor supply is unresponsive to manual labor wages, an increase in manual labor demand in the manufacturing sector will have a greater impact on wages than if the labor supply of manual laborers is responsive to changes in the wage. Similarly, α_2 decreases in the wage elasticity of skilled labor supply and in the elasticity of skilled agricultural labor demand.

B.2. **Obtaining the Income equation.** Household income is given by:

$$y_h = w_u \lambda_h + w_s * \xi_h + \Pi^{Ag}(w_u, w_s, Land_h, \theta, s^a_h, X_h, p, Rain) + m_h$$

where y_h denotes household income, λ denotes the proportion of household time spent working in the manual labor market, ξ denotes the proportion of household time spent in the skilled labor market. Π^{Ag} is the household cultivation profit function, which is a function of household land endowments, $Land_h$, the agricultural technology θ, the education of the household head, farm assets (X_h) and p denotes local prices of inputs (for example, the price of agricultural seeds). m_h denotes all other sources of household income, for example income earned through the bullock rental market or through local non-agricultural enterprizes.

A linear approximation to the cultivation profit function is given by:

$$\begin{aligned} \Pi^{Ag}_{hdt} &= \zeta_0 + \zeta_1 w_{mdt} + \zeta_2 w_{sdt} + \zeta_3 \theta_{dt} + \zeta_4 Land_{hdt} + \zeta_5 s_{hdt} + \zeta_6 \theta_{dt} * A_{hdt} \\ &\quad + \zeta_7 \theta_{dt} * s_{hdt} + \zeta_8 Land_{hdt} * w_{mdt} + \zeta_9 Land_{hdt} * w_{sdt} + \zeta_{10} Rain_{dt} \\ &\quad + \zeta_{11} \theta_{dt} * w_{mdt} + \zeta_{12} \theta_{dt} * w_{sdt} + \mu_h + \mu_d + \nu_{hdt} \end{aligned}$$

where the error term contains household level farm assets, local prices and potentially the interaction of these terms with agricultural productivity. The interaction between agricultural productivity and education arises from the complementarity between these two variables (Foster and Rosenzweig, 1996). In addition, if large landowners hire more labor on the labor market, the relationship between wages and agricultural profits will vary according to land holdings.

Substituting the empirical specifications for the profit function (18) and wages (10 and 11) into incomes and assuming (at first) that there is no variation across households in unskilled and skilled labor market time, we get:

$$\begin{aligned} y_{mdt} &= g_0 + g_1 \mathrm{E}^M_{total,dt} + g_2 \mathrm{E}^M_{sdt} + g_3\theta_{dt} + g_4 A_{dt} + g_5 Land_{hdt} + g_6 s^a_{hdt} + g_7 \mathrm{E}^M_{total,dt} * Land_{hdt} \\ &\quad + g_8 \mathrm{E}^M_{sdt} * Land_{hdt} + g_9\theta_{dt} * Land_{hdt} + g_{10} s_{hdt} * Land_{hdt} + g_{11} \mathrm{E}^M_{mdt} * \theta_{dt} + g_{12} \mathrm{E}^M_{sdt} * \theta_{dt} \\ &\quad + g_{13} s * \theta + g_{14} A_{dt} * Land_{hdt} + g_{15} A_{dt} * \theta + g_{16} Rain + g_{17}\theta^2 + \mu_h + \mu_d + \nu_{hdt} \end{aligned}$$

where the g terms are linear functions of skilled and unskilled labor market time:[55] The error term includes other sources of household income, household level farm assets, local prices and potentially the interaction of these term with agricultural productivity.

I estimate a compressed version of this regression. I separate households into groups according to their landholding status and the education of the household head. The coefficients on agricultural productivity and predicted manufacturing employment capture the average effect of a change in these terms on households within that endowment category:

$$y_{mdt} = \gamma_0 + \gamma_1 \mathrm{E}^M_{total,dt} + \gamma_2 \mathrm{E}^M_{sdt} + \gamma_3\theta_{dt} + \gamma_4 A_{dt} + \gamma_5 Land_{hdt} + \gamma_6 s^a_{hdt} + \mu_h + \mu_d + \nu_{hdt}$$

where: $\gamma_1 = \alpha_1(\zeta_1+\lambda)+\beta_1(\zeta_2+\xi)+Land_h*(\zeta_8\alpha_1+\zeta_9\beta_1)$ $\gamma_2 = \alpha_2(\zeta_1+\lambda)+\beta_2(\zeta_2+\xi)+Land_h*(\zeta_8\alpha_2+\zeta_9\beta_2)$ $\gamma_3 = \alpha_3(\zeta_1+\lambda)+\beta_3(\zeta_2+\xi)+\zeta_{11}\alpha_0+\zeta_{12}\beta_0+Land_h\zeta_6+s_h\zeta_7$

B.3. **Alternative specification for the wage equation.** In the first stage specification, employment in industry i is written as a function of industry i's policy changes, industry i's resource usage and regional endowments. The error term in the first stage specification includes the local wage, all other industries' policy changes and potentially the interaction of industry i's policies with industry j's policies.

To violate my identification strategy, the omitted variables in the first-stage error term must be correlated with the triple interaction terms *as well as* with unobserved determinants of the wage regression. If the instrumental variables are valid then the *partial* correlation between the second-stage error term and each instrumental variable should be zero. If policy changes are correlated across industries over time, the covariance between the interaction between own-industry tariffs, own-industry resource use and regional raw material endowments and any interaction involving industry j's tariffs may be non-zero.

In the case that the manufacturing and agriculture only overlap in the labor market, these additional terms are highly unlikely to be partially correlated with the wage error term. If there is another plausible channel through which the additional terms directly enter into the second stage regressions, these effects should in any case be largely absorbed by the interaction of average industry tariffs and local natural

[55] $g_0 = \zeta_0+\zeta_1\alpha_0+\zeta_2\beta_0+\lambda\alpha_0+\xi\beta_0$ $g_1 = \alpha_1(\zeta_1+\lambda)+\beta_1(\zeta_2+\xi)$, $g_2 = \alpha_2(\zeta_1+\lambda)+\beta_2(\zeta_2+\xi)$, $g_3 = \alpha_3(\zeta_1+\lambda)+\beta_3(\zeta_2+\xi)+\zeta_{11}\alpha_0+\zeta_{12}\beta_0$, $g_4 = \alpha_4(\zeta_1+\lambda)+\beta_4(\zeta_2+\xi)$, $g_5 = \zeta_4$, $g_6 = \zeta_5$, $g_7 = \zeta_8\alpha_1+\zeta_9\beta_1$, $g_8 = \zeta_8\alpha_2+\zeta_9\beta_2$, $g_9 = \zeta_6$, $g_{10} = \zeta_6$, $g_{11} = \zeta_{11}\alpha_1+\zeta_{12}\beta_1$, $g_{12} = \zeta_{11}\alpha_2+\zeta_{12}\beta_2$, $g_{13} = \zeta_7$, $g_{14} = \zeta_8\alpha_4+\zeta_9\beta_4$, $g_{15} = \zeta_{11}\alpha_4+\zeta_{12}\beta_4$, $g_{16} = \zeta_{10}$,$g_{17} = \zeta_{11}\alpha_3+\zeta_{12}\beta_3$

resource characteristics included in the second stage. Finally the nature of the policy reforms conducted greatly reduces the possibility of concern about restrictions on the first stage specification.[56]

I estimate two alternative specification to show that my results are robust to other approaches. In the first, I allow industry i to be affected by industry j's tariff change through the labor market. In the second more general approach, I allow industry i's employment to vary with policy changes across all industries. With a long panel of industries, this alternative specification could in principal be conducted at a 3-digit level. With only four data points in time, for reasons of parsimony I use 2-digit industry categories.

In the following, I allow shocks in industry j to enter into industry i's employment through the labor market, notably through the wage. Writing manufacturing employment in industry i, district d, time t as:

$$\begin{aligned} l_{idt} &= \beta_0 + \beta_1 \tau_{it} + \beta_2 \tau_{it} * r_i + \beta_3 \tau_{it} * r_i * n_d + \beta_4 \tau_{it} * n_d + \beta_5 r_i * n_d \\ &\quad + \beta_6 w_{dt} + \beta_7 w_{dt} * s_i + \beta_8 \tau_{it} * w_{dt} + \beta_9 \tau_{it} * w_{dt} * s_i + \beta_{10} \tau_{it} s_i + \beta_{11} X_{dt} + \delta_d + \delta_t + u_{idt} \end{aligned}$$

Where τ_{it} is the import tariff covering products in an industry at a given moment in time, n_d is the natural resource stock in district d, r_i is a measure of natural resource use, w_{dt} are equilibrium district wages at time t, s_i is a measure of labor usage, X_{dt} are other explanatory variables. δ_d are district level fixed effects and δ_t are time dummies. The only difference between equation 18 and the first stage equation introduced in the main body of the text are terms w_{dt} through $\tau_{it} * s_i$. The presence of w_{dt} in equation (5) highlights the simultaneity problem since in a structural equation equilibrium employment is clearly a function of equilibrium wages.

Aggregating equation this equation over J industries to obtain employment in the manufacturing sector at a district level:

$$\begin{aligned} l_{dt} &= J\beta_0 + J\beta_1 \overline{\tau}_t + J\beta_2 \overline{\tau r}_t + J\beta_3 \overline{\tau r}_t * n_d + J\beta_4 \overline{\tau}_t * n_d + J\beta_5 \overline{r} * n_d \\ &\quad + J\beta_6 w_{dt} + J\beta_7 \overline{s} w_{dt} + J\beta_8 \overline{\tau}_t * w_{dt} + J\beta_9 \overline{\tau s}_t * w_{dt} + J\beta_{10} \overline{\tau s}_t + J\beta_{11} X_{dt} + J\delta_d + J\overline{u}_{dt} \end{aligned}$$

where $\overline{x} = \frac{1}{J}\sum_i^J x_i$. Since district and time dummies are included in the specification, only variables that vary at a district time level are included leaving:

$$\begin{aligned} l_{dt} &= J\beta_0 + J\beta_3 \overline{\tau r}_t * n_d + J\beta_4 \overline{\tau}_t * n_d + J\beta_6 w_{dt} + J\beta_7 w_{dt} * \overline{s} + J\beta_8 w_{dt} * \overline{\tau}_t \\ &\quad + J\beta_9 w_{dt} * \overline{\tau s}_t + J\beta_{11} X_{dt} + J\delta_d + J\overline{u}_{dt} \end{aligned}$$

[56]Topalova (2004) shows that the differential changes in import tariffs across industries during the trade policy reforms conducted between 1991 and 1997 were unrelated to the state of the industries at the beginning of the reform.

Inserting (33) into (31) we get:[57]

$$
\begin{aligned}
l_{dt} &= b_0 + b_1\overline{\tau r}_t * n_d + b_2\overline{\tau}_t * n_d + b_{3t}l_{dt} + b_{4,t}X_{dt} + D_{dt} + \xi_{dt} \\
b_3 &= \alpha_2(\beta_6 - \beta_7\overline{s} - \beta_8\overline{\tau}_t - \beta_9\overline{\tau s}_t) \\
l_{dt} &= \frac{1}{(b_{3t})}[b_0 + b_1\overline{\tau r}_t * n_d + b_2\overline{\tau}_t * n_d + b_{4t}X_{dt} + D_{dt} + \xi_{dt}]
\end{aligned}
$$

Inserting district level wages into the industry-district level wage regression, and ignoring interactions of industry level variables with district and time dummies:[58]

$$
\begin{aligned}
l_{idt} &= a_0 + a_1\tau_{it} + a_2\tau_{it} * r_i + a_3\tau_{it} * r_i * n_d + a_4\tau_{it} * n_d + a_5 r_i * n_d + a_6 X_{dt} \\
&\quad + a_7 X_{dt} * s_i + a_8 X_{dt} * \tau_{it} + a_9 X_{dt} * \tau_{it} * s_i + a_{10}l_{dt} + a_{11}l_{dt} * s_i + a_{12}l_{dt} * \tau_{it} \\
&\quad + a_{13}l_{dt} * \tau_{it} * s_i + a_{14}\tau_{it} * s_i + D_d + D_t + \epsilon_{idt}
\end{aligned}
$$

Substituting equation (34) into industry-level employment above:[59]

$$
\begin{aligned}
l_{idt} &= c_0 + c_1\tau_{it} + c_2\tau_{it} * r_i + c_3\tau_{it} * r_i * n_d + c_4\tau_{it} * n_d + c_5 r_i * n_d + c_6 X_{dt} \\
&+ c_7 X_{dt} * s_i + c_8 X_{dt} * \tau_{it} + c_9 X_{dt} * \tau_{it} * s_i + c_{10t}X_{dt} + c_{11t}X_{dt} * s_i + c_{12t}X_{dt} * \tau_{it} \\
&+ c_{13t}X_{dt} * s_i * \tau_{it} + c_{14t}\overline{\tau r}_t * n_d + c_{15t}\overline{\tau r}_t * n_d * s_i + c_{16t}\overline{\tau r}_t * n_d * \tau_{it} \\
&+ c_{17t}\overline{\tau r}_t * n_d * s_i * \tau_{it} + c_{18t}\overline{\tau}_t * n_d + c_{19t}\overline{\tau}_t * n_d * s_i \\
&+ c_{20t}\overline{\tau}_t * n_d * \tau_{it} + c_{21t}\overline{\tau}_t * n_d * s_i * \tau_{it} + D_d + D_t + \varepsilon_{idt}
\end{aligned}
$$

$c_{10t}X_{dt}$,$c_{11,t}X_{dt} * s_i$,$c_{12,t}X_{dt} * \tau_{it}$ and $c_{13t}X_{dt} * s_i * \tau_{it}$ capture third-order effects of employment responses to changes in aggregate employment driven by changes in average tariffs vary with X_{dt} and interactions of X_{dt} with industry level labor usage and tariffs. Since these third-order effects are likely to be small, I omit these variables from the analysis. A similar case can be made for excluding $c_{16t}\overline{\tau r}_t * n_d * \tau_{it}$, $c_{17t}\overline{\tau r}_t * n_d * s_i * \tau_{it}$, $c_{20t}\overline{\tau}_t * n_d * \tau_{it}$, $c_{21t}\overline{\tau}_t * n_d * s_i * \tau_{it}$. In addition, under the assumption that $\beta_8 - \beta_{10} = 0$, c_{13t} through c_{19t} don't vary over time.[60] This leaves:

$$
\begin{aligned}
l_{idt} &= c_0 + c_1\tau_{it} + c_2\tau_{it} * r_i + c_3\tau_{it} * r_i * n_d + c_4\tau_{it} * n_d + c_5 r_i * n_d + c_6 X_{dt} \\
&+ c_7 X_{dt} * s_i + c_8 X_{dt} * \tau_{it} + c_9 X_{dt} * \tau_{it} * s_i + c_{14}\overline{\tau r}_t * n_d + c_{15}\overline{\tau r}_t * n_d * s_i + \\
&+ c_{18}\overline{\tau}_t * n_d + c_{19}\overline{\tau}_t * n_d * s_i + D_d + D_t + \varepsilon_{idt}
\end{aligned}
$$

---

[57] where $b_0 = (\tilde{\beta}_0 + \alpha_0(\beta_6 + \beta_7\overline{s}))$, $b_1 = \beta_3$, $b_2 = \beta_4$, $b_{3,t} = \alpha_2(\beta_6 - \beta_7\overline{s} - \beta_8\overline{\tau}_t - \beta_9\overline{\tau s}_t)$, $b_{4,t} = (\beta_{11} + \alpha_1(\beta_6 + \beta_7\overline{s} + \beta_8\overline{\tau}_t + \beta_9\overline{\tau s}_t))$, $D_{d,t} = (\delta_d + \alpha_0(\beta_8\overline{\tau}_t + \beta_9\overline{\tau s}_t) + d_d(\beta_6 + \beta_7\overline{s} + \beta_8\overline{\tau}_t + \beta_9\overline{\tau s}_t)) + (\delta_t + d_t(\beta_6 + \beta_7\overline{s} + \beta_8\overline{\tau}_t + \beta_9\overline{\tau s}_t))$ $\xi_{d,t} = (u_{i,d,t} + v_{d,t}(\beta_6 + \beta_7\overline{s} + \beta_8\overline{\tau}_t + \beta_9\overline{\tau s}_t))$

[58] $a_0 = (\tilde{\beta}_0 + \alpha_0(\beta_6 + \beta_7 s_i))$, $a_1 = \beta_1 + \beta_8\alpha_0$, $a_2 = \beta_2$ $a_3 = \beta_3$, $a_4 = \beta_4$ $a_5 = \beta_5$ $a_6 = \beta_{11} + \alpha_1\beta_6$, $a_7 = \alpha_1\beta_7$, $a_8 = \alpha_1\beta_8$, $a_9 = \alpha_1\beta_9$, $a_{10} = \alpha_2\beta_6$, $a_{11} = \alpha_2\beta_7$, $a_{12} = \alpha_2\beta_8$, $a_{13} = \alpha_2\beta_9$, $a_{14} = \beta_{10}$ $\epsilon_{idt} = u_{idt} + v_{dt}(\beta_6 + \beta_7 * s_i + \beta_8\tau_{it} + \beta_9\tau_{it} * s_i)$

[59] $c_0 = (\tilde{\beta}_0 + \alpha_0(\beta_6 + \beta_7 s_i)) + b_0 a_{10}$, $c_1 = \beta_1 + \beta_8\alpha_0$, $c_2 = \beta_2 + b_0 a_{12}$ $c_3 = \beta_3$, $c_4 = \beta_4$ $c_5 = \beta_5$, $c_6 = \beta_{11} + \alpha_1\beta_6$, $a_7 = \alpha_1\beta_7$, $a_8 = \alpha_1\beta_8$, $a_9 = \alpha_1\beta_9$, $a_{10} = \alpha_2\beta_6$, $c_{10t} = a_{10}\frac{1}{b_{3t}}(\beta_{11} + \alpha_1(\beta_6 + \beta_7\overline{s} + \beta_8\overline{\tau}_t + \beta_9\overline{\tau s}_t))$, $c_{11t} = a_{11}\frac{1}{b_{3t}}(\beta_{11} + \alpha_1(\beta_6 + \beta_7\overline{s} + \beta_8\overline{\tau}_t + \beta_9\overline{\tau s}_t))$,$c_{12t} = a_{12}\frac{1}{b_{3t}}(\beta_{11} + \alpha_1(\beta_6 + \beta_7\overline{s} + \beta_8\overline{\tau}_t + \beta_9\overline{\tau s}_t))$, $c_{13t} = a_{13}\frac{1}{b_{3t}}(\beta_{11} + \alpha_1(\beta_6 + \beta_7\overline{s} + \beta_8\overline{\tau}_t + \beta_9\overline{\tau s}_t))$ $c_{14t} = a_{10} * \frac{1}{b_{3t}} * \beta_3$, $c_{15t} = a_{11} = a_{11} * \frac{1}{b_{3t}} * \beta_3$ $c_{16t} = a_{12} * \frac{1}{b_{3t}} * \beta_3$ $c_{17t} = a_{13} * \frac{1}{b_{3t}} * \beta_3$ $c_{18t} = a_{10} * \frac{1}{b_{3t}} * \beta_3$, $c_{19t} = a_{11} = a_{11} * \frac{1}{b_{3t}} * \beta_3$ $c_{20t} = a_{12} * \frac{1}{b_{3t}} * \beta_3$ $c_{21t} = a_{13} * \frac{1}{b_{3t}} * \beta_3$

[60] Including the interactions of these variables with time trends has little impact on the coefficient estimates in the second stage. These results are therefore not included in this robustness checks but are available upon request.

Table A.1: Physical Costs of Transportation of Raw Materials						
	Density		Specific Gravity	Domestic Price/Ton		Density/ Price
	Mean	SD	Mean	Mean	SD	
Tobacco	320	-	0.32	36000	-	0.01
Timber	400	-	0.4-0.86	-	-	-
Cotton (w Lint)	560	-	0.56	12900	-	0.04
Rice	580	-	0.58	18400	-	0.03
Sugar (Raw)	960	-	0.96	32000	-	0.03
Water	1000	-	1	-	-	-
Coal (Bitumious)	1200	-	1.2	390	-	3.08
Clay (Kaolin)	2160	-	2.16	180	-	12.00
Ceramics	2620	300	2.62	187	160.10	14.04
Construction	2660	120	2.66	203	40.41	13.08
Chemicals	3684	1100	3.68	509	297.59	7.24
Steel	4610	2410	4.61	1050	1257.98	4.39
Metals	8850	6400	8.85	45135	61042.80	0.20
Gold	19300	-	19.3	-	4.40E+09	4.39E-06

$$\boxed{Density = \frac{Mass}{Volume}}$$

Specific Gravity, Source: Compiled from Various Engineering Texts
Prices: a) Minerals - Domestic Ore Prices in 1990 Rs at Mine Head. b) Crops - Wholesale prices, from Agricultural Situation in India. Source Mineral Data: Mineral Yearbook, Indian Bureau of Mines.

Table A.2: Industries in top quintiles for different raw materials			
	Wood	**Ceramics**	**Metal**
1	Bidi Manufacture	Structural Clay Products	Fertilizers and Pesticides
2	Sawing of Wood	Glass and Glass Products	Refractory and Structural Clay
3	Veneer Manufacture	Earthen and Plaster Products	Semi Finished Iron
4	Structural Wooden Products	Non-Structural Ceramics	Ferro-Alloys
5	Wooden and Cane Boxes	Cement and Plaster	Copper Manufacturing
6	Wood Industrial Products	Mica Products	Brass Manufacturing
7	Cork Products	Structural Stone Goods	Aluminium Manufacturing
8	Wooden Furniture	Asbestos Cement	Zine Manufacturing
9	Bamboo Furniture	Misc Non-Metallic Mineral Products	Processing of Metal Scraps
10	Wooden Products nec	Radiographic Aparatus	Other Non-Ferous Metal Products
	Energy	**Chemicals**	**Proportion of Literate Labor**
1	Pulp, Paper and Paper Board	Tea Processing	Plastics
2	Containers and Boxes	Organic and Inorganic Chemicals	Drugs and Medecines
3	Paper n.e.c	Fertilizers and Pesticides	Coke Oven Products
4	Organic and Inorganic Chemica	Paints and Varnishes	Batteries
5	Coke Oven Products	Drugs and Medecines	Printing and Publishing of Books
6	Coal and Coal Tar Products n.e.	Perfumes and Cosmetics	Tyre and Tubes
7	Cement, Lime and Plaster	Refining of Sugar	Refined Petroleum Products
8	Ferro-Alloys	Tyre and Tubes	Metal Furniture and Fixtures
9	Brass Manufacturing	Footwear	Insulated Wires and Cables
10	Aluminium Manufacturing	Rubber Products	Electrical Equipment n.e.c.

Table A.3: Robustness Checks, Wages

Panel A: Agricultural Prices

Dependent Varible: Farm-Harvest Price of Crop, in 1987 Rs.

	Jowar	**Cotton**	**Rice**	**Wheat**	**Sugar**
	(a)	**(b)**	**(c)**	**(d)**	**(d)**
Deregulated*Forest Cover	0.621	-0.166	0.324	-0.631	0.105***
*Wood Proportion	(0.421)	(0.134)	(0.345)	(0.618)	(0.262)
Deregulated*Ceramics	0.044	-0.203	0.006	-0.003	-0.002
*Ceramics Proportion	(0.048)	(0.136)	(0.005)	(0.004)	(0.002)
Deregulated*Construction	-0.001	0.012	-0.002	0.002	-0.003
*Construction Proportion	(0.003)	(0.009)	(0.006)	(0.004)	(0.002)
Deregulated*Coal	-0.002	0.003	0.006	0.004	-0.003
*Energy Proportion	(0.005)	(0.014)	(0.005)	(0.006)	(0.004)
Deregulated*Electricity	-0.035***	0.018	-0.156	0.352	-0.115
*Energy Proportion	(0.011)	(1.422)	(0.379)	(0.385)	(0.187)
Constant	32.38**	1.777	11.8	-15.937	-4.747
	(13.690)	(33.863)	(11.118)	(12.764)	(7.789)
Observations	721	614	844	652	586
F-Statistic over Triple Interactions	2.35	1.29	0.95	0.88	2.73

Panel B: Groundwater Table Depth

Dependent Varible: Proportion of Villages in a District under a table threshold.

	<10m	**<20m**	**<30m**	**<40m**	**<50m**
	(a)	**(b)**	**(c)**	**(d)**	**(d)**
Deregulated*Forest Cover	-0.009	-0.005	-0.003	0.004*	-0.002
*Wood Proportion	(0.008)	(0.006)	(0.006)	(0.002)	(0.002)
Deregulated*Ceramics	0.003	0.004	0.005	0.002	0.004
*Ceramics Proportion	(0.004)	(0.003)	(0.003)	(0.004)	(0.002)
Deregulated*Construction	0.002	0.001	0.001	0.001	0.000
*Construction Proportion	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Deregulated*Coal	0.212**	0.094	0.058	0.046	0.022
*Energy Proportion	(0.090)	(0.062)	(0.046)	(0.031)	(0.018)
Deregulated*Electricity	0.004	0.004	0.003	0.002	0.001
*Energy Proportion	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
Constant	-1.859	-1.778	-3.501	-5.109**	-0.847
	(5.628)	(4.465)	(3.839)	(2.441)	(1.391)
Observations	313	313	313	313	313
F-Statistic over Triple Interactions	1.7	1.19	1.1	1.79	0.88

* significant at 10%; ** significant at 5%; *** significant at 1%

All specifications include district fixed effects, year dummies, rainfall variables (total rainfall between june and september, rainfall squared and a shock measure taking the value 1 if there is a positive rainfall shock, 0 if no shock and -1 for negative shocks), the log of male and female population for the landless and by land quantiles. In addition, they include the interaction of district level resources with average tariff and delicensing reforms.

TABLE A.4: Robustness Checks, Income

Panel A: Examining Sensitivity to Removing Self-Employment in the Service or Manufacturing Sector

Dependent Variable: Log Consumption Per Capita

	Landless		<2 acres		>2 acres	
	Illiterate	Literate	Illiterate	Literate	Illiterate	Literate
	(a)	(b)	(c)	(d)	(e)	(f)
Log(Manufacturing Employment)	0.362**	0.230*	0.145***	0.106**	0.134	0.012
	(0.162)	(0.132)	(0.043)	(0.042)	(0.099)	(0.112)
Log(Literate Manufacturing Employment)	-0.289*	0.194*	-0.119	-0.152*	-0.039	0.169*
	(0.173)	(0.106)	(0.098)	(0.082)	(0.069)	(0.094)
Log(Agricultural Productivity)	0.216***	0.146	0.176*	0.178**	0.201**	0.192*
	(0.081)	(0.125)	(0.105)	(0.090)	(0.102)	(0.109)
Constant	1.907	-2.153	2.456	-0.171	3.174	1.778
	(1.247)	(2.106)	(1.848)	(1.683)	(2.314)	(1.841)
Adjusted R-Squared	0.324	0.301	0.319	0.327	0.383	0.344
Observations	36042	26600	24382	21173	30478	40654

Panel B: Examining whether Service Sector Employment changes with instruments

Dependent Variable: Log of District Level Sectoral Employment

	Construction		Retail		Transport	
	(a)	(b)	(c)	(d)	(e)	(f)
Log(Agricultural Productivity)	-0.99	-0.637*	0.122	-0.219	-0.412	-0.269
	(0.723)	(0.373)	(0.442)	(0.542)	(0.485)	(0.745)
Deregulated*Forest Cover	30.978	4.868	-19.070	30.476	8.390	-22.166
*Wood Proportion	(25.202)	(13.542)	(24.896)	(25.986)	(13.965)	(25.820)
Deregulated*Ceramics	0.089	0.030	0.208	0.013	0.028	0.185
*Ceramics Proportion	(0.209)	(0.152)	(0.163)	(0.205)	(0.155)	(0.172)
Deregulated*Construction	0.008**	0.001	0.000	0.001	0.003	-0.001
*Construction Proportion	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
Deregulated*Coal	-40.156	24.035	45.085	-8.531	42.874	28.195
*Energy Proportion	(62.847)	(39.415)	(62.620)	(64.347)	(42.756)	(63.729)
Deregulated*Electricity	-0.848	0.414	-2.307*	-0.774	0.812	-2.379*
*Energy Proportion	(1.417)	(0.496)	(1.288)	(1.352)	(0.542)	(1.242)
Constant	15.200	3.297	(10.805)	7.581	0.518	(8.274)
	(12.148)	(6.966)	(11.756)	(13.328)	(7.973)	(13.958)
Observations	967	967	967	967	967	967
Region Time Trends	N	Y	N	Y	N	Y
F-Statistic over Triple Interactions	1.78	0.41	0.78	0.94	1.12	1.1

* significant at 10%; ** significant at 5%; *** significant at 1%

All specifications include district fixed effects, year dummies, rainfall variables (total rainfall between june and september, rainfall squared and a shock measure taking the value 1 if there is a positive rainfall shock, 0 if no shock and -1 for negative shocks), the log of male and female population for the landless and by land quantiles. In addition, they include the interaction of district level resources with average tariff and delicensing reforms.

Table A.5: Robustness Checks, Education

Specification: Linear Probability Model

Panel A: Dependent Variable: 1 if boy reports having started school

	Landless		Landed	
	Illiterate	Literate	Net Exporters	Net Importers
	(a)	(b)	(c - L)	(c - H)
Log(Manufacturing Employment)	0.167***	0.042**	-0.091***	-0.126*
	(0.057)	(0.021)	0.039	(0.075)
Log(Manufacturing Literate)	-0.067*	0.136***	0.095	0.028
	(0.037)	(0.049)	(0.067)	(0.040)
Log(Agricultural Productivity)	0.343	0.207**	0.261	0.271
	(0.553)	(0.100)	(0.231)	(0.486)
District Fixed Effects	Y	Y	Y	Y
State*Year Dummies	Y	Y	Y	Y
Region Time Trends	Y	Y	Y	Y
Adjusted R-Squared	0.178	0.132	0.179	0.2063
Observations	14207	12524	24421	41315

Panel B: Dependent Variable: 1 if girl reports having started school

	Landless		Landed	
	Illiterate	Literate	Net Exporters	Net Importers
	(a)	(b)	(c - L)	(c - H)
Log(Manufacturing Employment)	0.182**	0.189**	0.154*	-0.328***
	(0.080)	(0.092)	(0.089)	(0.086)
Log(Manufacturing Literate)	-0.108*	-0.019	-0.186***	0.162***
	(0.061)	(0.092)	(0.059)	(0.595)
Log(Agricultural Productivity)	0.236	0.201	0.589*	0.567*
	(0.173)	(0.134)	(0.329)	(0.318)
District Fixed Effects	Y	Y	Y	Y
State*Year Dummies	Y	Y	Y	Y
Region Time Trends	Y	Y	Y	Y
Adjusted R-Squared	0.242	0.194	0.273	0.292
Observations	13849	12865	20983	35845

Specification: Probit; Include Education Infrastructure Controls and Time Trends Interacted with Initial Characteristics

Panel C: Dependent Variable: 1 if boy or girl reports having started school

	Landless			
	Boys		**Girls**	
	Illiterate	**Literate**	**Illiterate**	**Literate**
	(a)	(b)	(c)	(d)
Log(Manufacturing Employment)	0.279***	0.057**	0.093*	0.258***
	(0.102)	(0.029)	(0.048)	(0.046)
Log(Manufacturing Literate)	-0.197***	0.172***	-0.036	-0.221***
	(0.076)	(0.043)	0.045	(0.043)
Log(Agricultural Productivity)	0.327*	0.125**	0.251*	0.127
	(0.175)	(0.063)	(0.134)	(0.088)
District Fixed Effects	Y	Y	Y	Y
State*Year Dummies	Y	Y	Y	Y
Region Time Trends	Y	Y	Y	Y
Time Trends*Initial Literacy/Primary	Y	Y	Y	Y
Adjusted R-Squared	0.137	0.127	0.194	0.211
Observations	14207	12524	13849	12865

* significant at 10%; ** significant at 5%; *** significant at 1%

All specifications include district fixed effects, year dummies, rainfall variables (total rainfall between june and september, rainfall squared and a shock measure taking the value 1 if there is a positive rainfall shock, 0 if no shock and -1 for negative shocks), the log of male and female population for the landless and by land quantiles. In addition, they include the interaction of district level resources with average tariff and delicensing reforms. Finally they include the sex of the household head, his age, age squared and social group.

Table A.6: Robustness Checks, Education

Panel A: Dependent Varibles: Measures of Education Cost and Infrastructure

	Number of Teachers	Average Distance from School	Cost of Transport	Average Total Cost Per Student	Average Tuition Per Student
	(a)	(b)	(c)	(d)	(e)
Tariff*Forest Cover	3.749	-0.016	0.004	-0.643	0.051
*Wood Proportion	(17.364)	(0.031)	(0.014)	(1.012)	(0.037)
Tariff*Ceramics	0.008	-0.032*	-0.010	-0.019	-0.026
*Ceramics Proportion	(0.009)	(0.018)	(0.008)	(0.030)	(0.019)
Tariff*Construction	0.004*	-0.013***	0.001	-0.001	0.002
*Construction Proportion	(0.002)	(0.004)	(0.002)	(0.008)	(0.004)
Tariff*Coal	0.017	-0.008	0.004	-0.010	0.000
*Energy Proportion	(0.012)	(0.010)	(0.003)	(0.021)	(0.010)
Tariff*Electricity	-31.547*	-0.658**	0.473***	0.394	0.689
*Energy Proportion	(17.722)	(0.272)	(0.176)	(6.813)	(0.519)
Agricultural Productivity	0.499	0.710	0.951	2.088	0.941
	(0.437)	(0.949)	(0.902)	(2.289)	(1.276)
Constant	-13.158**	6.169	-22.966*	-60.489*	-26.931
	(6.365)	(14.942)	(12.377)	(34.886)	(23.709)
Observations	923	653	653	653	653
F-Stat over instruments and agriculutural index	2.67**	3.31***	1.63	0.55	0.94

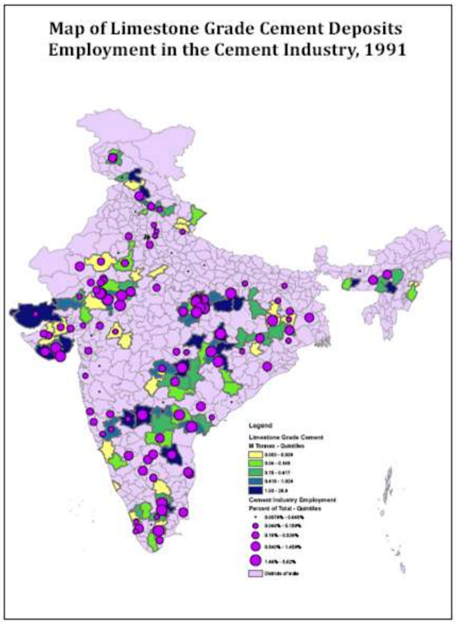
Panel B: Dependent Varibles: Proportion of Population Receiving Education Subsidies

	Free Tuition	Midday Meal	Books	Transport	Scholarship
	(a)	(b)	(c)	(d)	(e)
Tariff*Forest Cover	-0.010	-0.041***	-0.008	-0.029	0.013***
*Wood Proportion	(0.008)	(0.011)	(0.011)	(0.027)	(0.004)
Tariff*Ceramics	0.002	0.016	-0.040***	0.035	0.007
*Ceramics Proportion	(0.008)	(0.014)	(0.013)	(0.023)	(0.005)
Tariff*Construction	0.002*	0.000	-0.004**	-0.011*	0.000
*Construction Proportion	(0.001)	(0.003)	(0.002)	(0.006)	(0.001)
Tariff*Coal	-0.167	0.215	0.371***	-0.534*	-0.175***
*Energy Proportion	(0.115)	(0.176)	(0.136)	(0.309)	(0.056)
Tariff*Electricity	0.006	0.003	0.019***	-0.068**	0.007***
*Energy Proportion	(0.004)	(0.010)	(0.005)	(0.030)	(0.002)
Agricultural Index	-0.117	0.251	-0.206	-1.651	-0.204
	(0.343)	(0.490)	(0.565)	(1.060)	(0.244)
Constant	5.671	-8.13	-6.178	33.101*	5.774*
	(5.665)	(7.286)	(7.454)	(16.966)	(3.256)
Observations	653	653	653	653	653
F-Stat over instruments and agriculutural index	1.77	3.42***	6.73***	1.76	6.98***

* significant at 10%; ** significant at 5%; *** significant at 1%

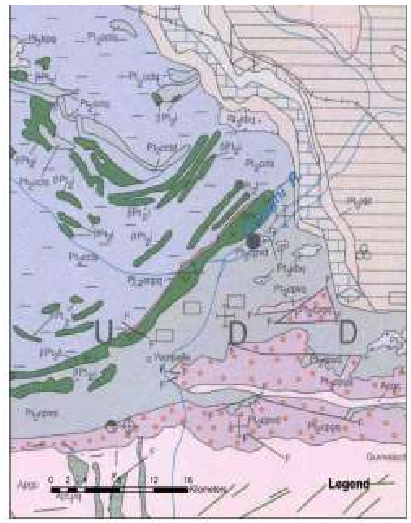
All specifications include district fixed effects, year dummies, rainfall variables (total rainfall between june and september, rainfall squared and a shock measure taking the value 1 if there is a positive rainfall shock, 0 if no shock and -1 for negative shocks), the log of male and female population for the landless and by land quantiles. In addition, they include the interaction of district level resources with average tariff and delicensing reforms.

Figure A1: Map of Employment in the Cement and Plywood industry and Limestone and Cement Stocks in 1990.



y

Figure A2: Exerpt from the National Mineral Atlas – Map showing Limestone, Barytes, and Asbestos Deposits in Cuddapah, Andhra Pradesh



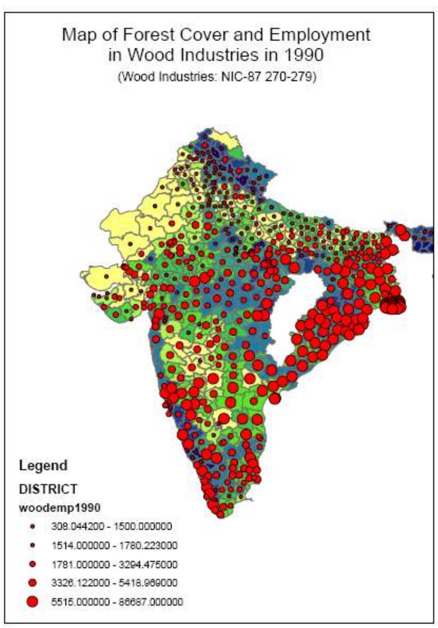
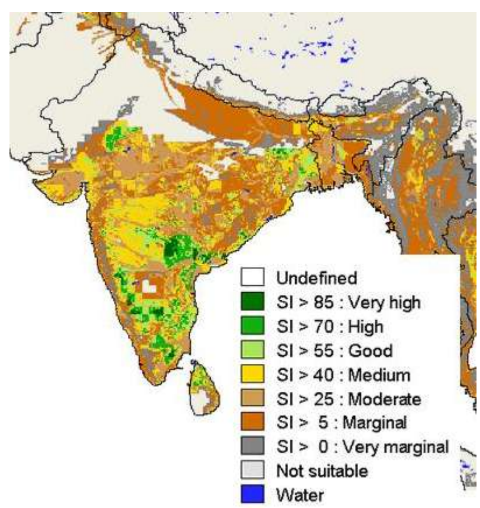


Figure A3: Map of Suitability for Rainfed Pulses in India



References

[1] Aghion, Philippe, Burgess, Robin, Redding, Stephen, and Zilibotti, Fabrizio, (2009) "The Unequal of Liberalization: Evidence from Dismantling the License Raj in India", *The American Economic Review*, 98(4), pp. 1397-141

[2] Ahluwalia, M.S. (1975) "Income Distribution and Development: Some Stylized Facts", *The American Economic Review*

[3] Ahluwalia, M.S. (2002) "State-Level Performance under Economic Reforms in India", in Anne O. Krueger (ED.) Economic Policy Reforms and the Indian Economy, New Delhi: Oxford University Press.

[4] Atkin, David (2009) "Trade, Tastes and Nutrition in India" *Working Paper*

[5] Autor, David, Levy, and Murnane (2003) "The Skill Content of Recent Technological Change: An Empirical Exploration", (2003) *Quarterly Journal of Economics*, 118(4), November 2003, 1279-1334.

[6] Banerjee, Abhijit (2004) "Education Policy and the Economics of the Family" *Journal of Development Economics*, 74(1), New Research on Education in Developing Economies, pp 3-32

[7] Banerjee, Abhijit and Esther Duflo (2005) "Growth Through the Lens of Development Economics" *Handbook of Development Economics*, 1(a) pp. 473-552

[8] Behrman, Jere B. and James C. Knowles "Household Income and Child Schooling in Vietnam" *The World Bank Economic Review*, 13(2) pp. 211-56

[9] Besley, Timothy and Robin Burgess, (2004) "Can Labour Regulation Hinder Economic Performance? Evidence from India," *The Quarterly Journal of Economics*, MIT Press, vol. 119(1), pages 91-134, February.

[10] Bharadwaj, Prashant (2009) "Fertility and Rural Labor Market Inefficiencies: Evidence from India" *Working Paper*

[11] Chari, A.V. (2009) "The aggregate productivity effects of entry and output restrictions: an analysis of license reform in India" *Working Paper*

[12] Datt, Gaurav and Martin Ravallion (2002) "Is India's Economic Growth Leaving the Poor Behind?," *Journal of Economic Perspectives*, 16(3), pp 89-108

[13] Deaton, Angus (2003) "Adjusted Indian Poverty Estimates for 1999-2000", *Economic and Political Weekly*, 38(4) pp. 322-326

[14] Deaton, Angus and Jean Dreze and "Poverty and Inequality in India: A Re-Examination" *Economic and Political Weekly*, 37(36), pp. 3729-3748

[15] Dreze, Jean and Geeta Kingdon (2001) "School Participation in Rural India" *Review of Development Economics* 5(1), pp 1-24

[16] Drze Jean and Amartya K. Sen, (1995) *India: Economic Development and Social Opportunity.* Oxford University Press.

[17] Duflo, Esther "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment" *American Economic Review* 91(4) pp795

[18] Edmonds, Eric, Nina Pavnick and Petia Topalova (2007) "Trade Adjustment and Human Capital Investments: Evidence from Indian Tariff Reform," NBER Working Papers 12884

[19] Ellison, Glen and Edward Glaeser (1997) "Geographic Concentrateion in U.S. manufacturing industries: A dartboard approach" *Journal of Political Economy* 105(5), 889-927

[20] Ellison, Glenn and Edward L. Glaeser (1999) "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?," *American Economic Review*, American Economic Association, vol. 89(2), pages 311-316, May.

[21] Foster, Andrew D. and Mark Rosenzweig (1993) "Information, Learning and Wage Rates in Low-Income Rural Areas" *Journal of Human Resources*, Vol 28(4) pp759-790

[22] Foster, Andrew D. and Mark Rosenzweig (1995) "Learning by doing and learning from others: Human capital and technical change in agriculture" *Journal of Political Economy* 103(6), 1176-1209

[23] Foster, Andrew D. and Mark Rosenzweig (1996) "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution," *American Economic Review*, American Economic Association, vol. 86(4), pages 931-53, September.

[24] Foster, Andrew D. and Mark Rosenzweig (2004) "Technological change and the distribution of schooling: evidence from the green revolution in India" *Journal of Dvelopment Economics*, 74(1), New Research on Education in Developing Economies, pp.87-111

[25] Foster, Andrew D. and Mark Rosenzweig (2004) “Agricultural Development, Industrialization and Rural Inequality” *Working Paper*

[26] Goldberg, Penny, Amit Khandewal, Nina Pavnick and Petia Topalova (2009) “Trade Liberalization and New Imported Inputs,” *forthcoming, American Economic Review Papers and Proceedings*

[27] Government of India, 1968, *National Classification of Occupations.* Government of India, Ministry of Labor.

[28] Government of India, 2001, Concepts and Definitions Used in NSS: Golden Jubilee Publication. Government of India, Ministry of Statistics and Program Implementation, National Sample Survey Organisation.

[29] Griliches, Zvi, *Technology, Education and Productivity: Essays in Applied Econometrics*, New York: Basil Blackwell

[30] Gulati, Ashok and Sudha Narayanan (2003) *The subsidy syndrome in Indian agriculture”* Oxford University Press: New Delhi

[31] Hasan, Rana, Devashish Mitra and K.V Ramaswamy (2007) “Trade Reforms, Labor Regulations, and Labor-Demand Elasticities: Empirical Evidence from India,” *The Review of Economics and Statistics*, MIT Press, vol. 89(3), pages 466-481, 02

[32] Jacboy, Hanan G (1994) “Borrowing Constraints and Progress Through School: Evidence from Peru,” *Review of Economics and Statistics,* 76(1) pp 151-60

[33] Jacoby, Hanan G and Emmanuel Skoufias (1997) “Risk, Financial Markets, and Human Capital in a Developing Country,” *Review of Economic Studies,* 64(3), pp 311-35

[34] Jayachandra, Seema (2006) “Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries” *Journal of Political Economy*, 114(3), pp. 538-575.

[35] Jessoe, Katrina (2009) “Improved Source, Improved Quality? Demand for Drinking Water Quality in Rural India” *Working Paper*

[36] Keskin, Pinar (2009) “Thirsty Factories, Hungry Farmers: Intersectoral Impacts of Industrial Water Demand”

[37] Kingdon, Geeta (2007) “The progress of school education in India” *Oxford Review of Economic Policy* 23(2) pp 168-195

[38] Kim, Sukkoo (1995), “Expansion of Markets and the Geographic Distribution of Economic Activity: The trends in U.S. Regional Manufacturing Structure, 1860-1987” *Quarterly Journal of Economics* 110, 881-908

[39] Kochar, Anjini (2002) “Emerging Challenges for Indian Education Policy” in Anne O. Krueger (ED.) Economic Policy Reforms and the Indian Economy, New Delhi: Oxford University Press.

[40] Kochar, Anjini “Urban influences on rural schooling in India”, *Journal of Development Economics*, 74(1), New Research on Education in Developing Economies, pp 113-136

[41] Kochhar, Kalpana, Utsav Kumar, Raghuram Rajan, Arvind Subramanian and Ioannis Tokatlidis (2006), “Indias Pattern of Development: What Happened, What Follows?” IMF Working Paper, WP/06/22

[42] Kuznets, Simon (1955) “Economic Growth and Income Inequality,” *American Economic Review* 65(1), pp. 1-28. See Kuznets Curve.

[43] Krueger, Anne, and Sajjid Chiony (2002), “The Indian Economy in Global Context” in Anne O. Krueger (ED.) Economic Policy Reforms and the Indian Economy, New Delhi: Oxford University Press.

[44] Krugman, Paul, (1991), *Geography and Trade*, Cambridge, Mass: MIT Press.

[45] Lanjouw, Peter and Rinku Murgai (2009), “Poverty Decline, Agricultural Wages and Non-Farm employment” *Agricultural Economics* Vol 40(2) pp243-263

[46] Mehta B. (1996) “’Free’ Primary Education”’ *Economic and Political Weekly* 31(16), pp 1035-1036

[47] Mehta, R. (1999) “Tariff and Non-tariff Barriers of Indian Economy: A Profile,” Research and Information System for the non-aligned and Other Developing Countries (RIS), New Delhi.

[48] Munshi, Kaivan and Mark Rosenzweig (2005), “Why is Mobility in India so Low? Social Insurance, Inequality, and Growth,” BREAD Working Paper No. 092.

[49] NCAER Publication, (1969) “Techno-Economic Survey of Andhra Pradesh” and other states; National Council of Applied Economic Research: New Delhi

[50] Nelson, R. and Phelps, E., 1966. “Investment in humans, technological diffusion, andeconomic growth” *American Economic Review: Papers and Proceedings* 51(2), pp. 69-75

[51] Panagariya, Arvind (2008) *India: The Emerging Giant* Oxford University Press: USA

[52] Ranis, Gustav and John Fei "A Theory of Economic Development" *The American Economic Review* 51(4), pp. 533-565

[53] Rosenzweig, Mark (1995), "Why Are There Returns to Schooling?", *American Economic Review (Papers and Proceedings)*, Vol 85(2), pp. 153-158

[54] Schultz, T.W. (1975) "The Value of the Ability to Deal with Disequilbria" *Journal of Economic Literature* 13(1) pp.827-46

[55] Schultz, Paul (2004) "School subsidies for the poor: evaluating the Mexican Progresa poverty program," *Journal of Development Economics*, 74(1), pp 199-250,

[56] Srinivasan, T.N. (2000), "Eight Lectures of India's Economic Reforms", Oxford India Paperbacks, New Delhi: Oxford University Press.

[57] Srinivasan, T.N. (2002), "The Indian Economy in Global Context" in Anne O. Krueger (ED.) Economic Policy Reforms and the Indian Economy, New Delhi: Oxford University Press.

[58] Topalova, Petia (2005) "Trade Liberalization, Poverty, and Inequality: Evidence from Indian Districts," NBER Working Papers 11614, National Bureau of Economic Research, Inc.

[59] Wooldridge, Jeffrey (2002) *Econometric Analysis of Cross Section and Panel Data* MIT Economic Press.