Health, Technical Efficiency And Agricultural Production In Indian Districts

Atheendar S. Venkataramani
Department of Health Policy and Administration
Yale University School of Public Health
New Haven, CT 06520 (USA)
Email: atheendar.venkataramani@yale.edu

K.R. Shanmugam (Corresponding Author)
Madras School of Economics
Chennai 600 025 (India)
Email: shanmugam@mse.ac.in

and

Jennifer Prah Ruger
Division of Global Health, Yale University School of Public Health
New Haven, CT 06520 (USA)
Email: jennifer.ruger@yale.edu

WORKING PAPER 11/2006

August 2006

Price: Rs.35

MADRAS SCHOOL OF ECONOMICS
Gandhi Mandapam Road
Chennai 600 025
India

Phone: 2230 0304/ 2230 0307/2235 2157
Fax : 2235 4847 /2235 2155
Email : info@mse.ac.in
Website: www.mse.ac.in

Health, Technical Efficiency and Agricultural Production in Indian Districts

Atheendar S. Venkataramani, K.R. Shanmugam (Corresponding Author) and Jennifer Prah Ruger

Abstract

In this study, we attempt to quantify the effect of improved population health on technical efficiency in agricultural production. Using data for over 260 districts in 15 Indian states, we employ the random-coefficients technique to estimate a Cobb-Douglas production function, computing overall and input specific technical efficiencies for each district. We then model health (the district infant mortality rate) as a determinant of (in) efficiency in a second stage, controlling for a range of other socioeconomic variables. We find that decreases in the infant mortality rate, as well as increases in the literacy rate and level of irrigation, are associated with significant increases in overall technical efficiency, and that a good portion of health's effect is probably due to improvements in the efficiency of labor use. While efficiency increases from improvements in irrigation and literacy are larger, the potential gains from health are still fairly substantial.

Keywords: Technical efficiency, Random coefficients model, the frontier production function
JEL classification code: *C33; D20*

1. Introduction

The process of economic development in poor countries is multi-faceted and likely requires the confluence of a number of factors – investment, technological change, public policy and institutional change, human capital generation, among many others – to succeed. Recent research has illustrated the role of health human capital in augmenting the processes of development and wealth accumulation. In fact, from the literature on cross-country economic growth, this role may be quite large (see Bhargava et al., (2001); Gallup and Sachs (2001); Bloom et al., (2004); Barro and Xala-i-Martin (2004); see Acemoglu and Smith (2006) for contrary findings). This macro evidence is supported by various micro-level studies, such as those which seek to quantify wage returns to improved body mass index, height, and freedom from illness (see, among others, Deolalikar (1988); Alderman et al., (1996); Strauss and Thomas (1998); Schultz and Tanzel (1997); and Schultz (2002, 2003)).

Bloom and Canning (2000) and Ruger et al., (2001) suggest several possible avenues through which health can exert a positive influence on economic performance. As implied by the results of studies looking at wage returns, better health human capital can improve (labor) productivity. Additionally, healthier individuals are more likely to increase investment in, and derive greater returns from, schooling and education (see, for example, Alderman et al., (2001), Glewwe et al., (2001), and Behrmann and Rosenzweig (2004), and Miguel and Kramer (2004), which all look at the impact of child health status on schooling and other indicators of present and future economic performance). Increased savings, and therefore, investment in physical capital, increased foreign direct investment, and demographic benefits, thanks to decreased fertility and lower dependency ratios, may all follow with improvements in health human capital (see Bloom et al., (2003) and Alsan et al., (2004)).

In the present study, we approach the question of the economic returns to health from a different perspective; *in particular, we seek to explore and quantify the impact of health on technical efficiency in production*. Technical efficiency is a measure of how well the decision-making units use their inputs in generating output(s) and represents an interesting way to think about the role of health (and human capital in general) in the production process.[1] In specifying health as a determinant of technical efficiency, here we are essentially conceptualizing health and indeed other aspects of human capital as *factors that allow firms to use their physical inputs (such as labor or capital) in a more efficient manner*.

Indeed, our study is fundamentally different from other studies, which generally specify health as an input in the production function process. While these approaches allow for direct estimation of production returns to health, they say little about *how* exactly health does this. Also, specifying health as an input to production in the same manner as one would specify labor or capital, may not necessarily be desirable since human capital in general is perhaps better viewed as an input which *augments the production process* indirectly.[2]

In essence, our study can be grouped with the literature on the influence of health on educational outcomes, foreign direct investment, wages, demographic change, and so forth, as studies seeking to disaggregate the findings on overall returns to health into the pathways that generate these returns. To the best of our knowledge, this is the first study to model health as a determinant of technical efficiency.[3]

Using cross-sectional agricultural production data for over 260 Indian districts in the early 1990s, we employ the random coefficients technique to compute district-wise input specific and overall technical

efficiency values. We then model overall technical efficiency and input-specific efficiencies as functions of district health status (proxied for by infant mortality rate) and a set of socioeconomic and ecological controls. We find that (i) there is a great deal of heterogeneity across Indian districts with respect to efficiency in agricultural production, (ii) better health is associated with increased technical efficiency, and (iii) much of this effect may be due to more effective use of the existing labor inputs.

The structure of the remainder of the paper is as follows. In section 2, we develop the econometric framework and the modeling strategy. In section 3, we describe the data and variables. In section 4, we present the empirical results of the production frontier function estimates. In section 5, we present the patterns in technical efficiency and discuss the determinants of overall and input-specific technical efficiency. In section 6, we offer some conclusions, prospects for future work, and policy implications.

2. Modeling

In this section, we outline the methodology used to quantify the effect of improved population health on technical efficiency in agricultural production. We employ a two-step procedure: we first estimate overall and input specific technical efficiency values for each district in the sample and subsequently use these estimates as dependent variables, specifying health and other factors as independent variables.

2.1 Deriving Efficiency Using the Random Coefficients Model

The technical efficiency (TE) of production for a given decision-making unit (such as firm or region) can be defined as the ratio of its actual output to the output that could potentially be produced if all existing inputs/technologies are used in the best possible fashion. Farrell (1957) carried out the first empirical study to measure technical efficiency for a cross-section of firms by using a deterministic/non-parametric frontier approach. Consequently, frontier efficiency comparisons have become synonymous with the term *'Farrell efficiency measurement'*. Later, Aigner *et al.,* (1977) and Meeusen and Broeck (1977) independently developed a stochastic frontier approach, which was originally applied to cross-sectional data, to measure technical efficiency. Here, the error term was modeled as a composite variable, consisting of a random noise component and a one-sided residual component (which follows a half normal distribution). This approach has been extended in many ways, both in terms of the specification of the error term (through the use of truncated normal, exponential and gamma distributions), as well as in the consideration of panel data (see Bauer (1990); Battese (1992); and Greene (1993) for comprehensive reviews of this literature).[4]

The stochastic frontier methodology essentially introduces technical efficiency as a multiplicative (neutral) shift variable within a production function framework. This means that the input coefficients of the conventional production function remain constant; only the intercept term is modified. However, there is no economic logic that substantiates this strong assumption. Since different uses of the same set of inputs can yield drastically different results, it stands to reason that the (output) returns from a given input may vary from firm to firm. In this case, the stochastic function approach, which assumes that the input elasticities are identical from firm to firm, is at best too rigid and at worst devoid of any real economic meaning.

The idea of varying coefficients was first appreciated by Nerlove (1965), who proposed treating the output elasticity of each input as a stochastic variable, thus being allowed to differ from firm to firm. Swamy (1970, 1971) later popularized this random coefficients approach. Kalirajan and Obwona (1994) and Kalirajan and Shand (1994a,b) discuss the use of this approach to model a frontier production function with (cross-sectional) heterogeneity in slopes and intercepts.[5] The model expounded upon in the latter two studies facilitates the estimation of firm- and input-specific technical efficiency values in a cross-sectional sample.

To be sure, the random coefficients approach has many advantages over alternate methodologies. Kalirajan and Shand (1994) list these advantages as follows. First, this approach may be viewed as a stochastic counterpart of the deterministic (data envelopment) approach: it does not require *a priori* assumptions on the distribution of firm specific TE. Second, as alluded to early, the modeling of production technology in this approach is in conformity with production theory. Third, the approach facilitates the calculation of overall and input-specific technical efficiency values, without involving any significant additional calculations. Finally, heteroscedasticity does not pose a problem with this technique. However, a major criticism leveled against this approach is that it imposes the constant returns to scale assumption.

Following Kalirajan and Obwona (1994), we posit that the (production function) parameters describing the production technology are random. Using this, the random coefficients model can be written as:

$$\text{Ln } Q_i = \Sigma_j \beta_{ij} \ln x_{ij} + u_i \; ; \quad i = 1,2,....,n \text{ and } j = 1, 2,...., k \quad (1)$$

where Q_i represents the output of the i^{th} firm (or region); x_j's are inputs (and the intercept), β_{ij}'s are parameters to be estimated, and u_i, is the residual term.

Each firm's coefficient vector β_{ij} is allowed to vary from the mean response vector $\bar{\beta}_j$ by some v_{ij}; that is, ($\bar{\beta}_k + v_{ik} = \beta_{ik}$). Hence, there is a particular response coefficient for each variable and each sample observation. The model assumes that: E (β_{ij}) = $\bar{\beta}_j$; V (β_{ij}) = $\sigma_j^2 > 0$; Cov (β_{ij}, β_{kj}) = 0 for $i \neq k$ (this implies that β_{ij} are i.i.d with fixed mean $\bar{\beta}_j$; Cov (β_{ij}, u_i) = 0; E (v_{ij}) = 0 for all i and j, E ($v_i v_j'$) = α_k for i = j and E ($v_i v_j'$) = 0 for $i \neq j$). Letting $\Sigma_j v_{ij} \ln x_{ij} + u_i = w_i$ and q = Ln Q, we can rewrite the equation (1) as:[6]

$$q_i = \Sigma_j \bar{\beta}_j \ln x_{ij} + w_i \; ; \quad i = 1,2,....,n \text{ and } j = 1, 2,...., k \quad (2)$$

where X is a N x K matrix of the independent variables and q, $\bar{\beta}$ and w are vectors of order N, K and N respectively. This above is a linear model with constant (or, mean response) coefficients, but heteroskedastic disturbances.[7] OLS estimation of (2) will yield unbiased - but inefficient - estimates of $\bar{\beta}_j$.

Using the iterative procedure suggested in Swamy (1970) one can obtain the feasible GLS estimate of $\bar{\beta}$. Following the presentation in Croppenstedt and Demeke (1997), an unbiased estimate of α can be gleaned from:

$$\hat{\alpha} = (G'G)^{-1} G' r^* \quad (3)$$

where $G = M^* X^*$; $M = I - X (X' X)^{-1}$; $r = q - \bar{\beta}' X$. The * denotes the matrix (vector) which is derived by squaring each element. Then, the estimate of the mean response vector is:

$$\hat{\bar{\beta}} = (X' \Omega^{-1} X)^{-1} X^{-1} X' \Omega^{-1} q \quad (4)$$

where Ω is the covariance matrix of the w's:

$$\Omega = \begin{pmatrix} x_1'\alpha x_1 + \sigma^2 I & 0 & .... & 0 \\ 0 & x_2'\alpha x_2 + \sigma^2 I & .... & 0 \\ . & . & .... & 0 \\ 0 & 0 & .... & x_N'\alpha x_N + \sigma^2 I \end{pmatrix} \quad (5)$$

If we substitute the estimates of α and σ^2, equation (4) will serve as the feasible GLS estimator. Since we are interested in calculating firm-specific response coefficients, the best linear estimates of these are given by:[8]

$$\hat{\beta}_i = \hat{\bar{\beta}} + \phi x_i' [x_i \phi x_i']^{-1} (q_i - x_i \hat{\bar{\beta}}) \quad (6)$$

where $\phi = \text{diag} (\hat{\alpha}_1, \hat{\alpha}_2, ... \hat{\alpha}_k)$.

To reiterate, these are the actual response coefficients for some specific input for each firm in the sample. The highest magnitude response parameter (across all firms) for each input and the intercept term form the production coefficients of the potential frontier production function. That is, we can identify the response coefficients representing the (potential) frontier production function from the above firm- and input-specific response coefficients as follows:

$$\beta_j^* = \max_i \{ \beta_{ij} \} \; ; \quad i = 1, 2,....n \text{ and } j = 1, 2, ..., k \quad (7)$$

It is noted that these parameters need not necessarily coincide with the response coefficients for any single firm. Indeed, they may represent the best combination of response from different firms. From the estimates of the frontier production function, we can calculate the firm specific potential output (q^*_i) as:[9]

$$q_i^* = \ln Q^*_i = \Sigma_j \ \beta^*_j \ln x_{ij} \tag{8}$$

From this, we can easily compute the overall technical efficiency of a given firm as the ratio of the actual output to the potential output, which is given by:

$$TE_i = \exp(q_i) \ / \ \exp(q^*_i). \tag{9}$$

We can also obtain the estimates of input specific efficiency measures for individual firms by calculating the ratio of actual response coefficients to the frontier response coefficients. In percentage terms, the efficiency of using j^{th} input by the ith firm (IT_i) is given by:

$$IT_{ij} = (\beta_{ij} \ / \ \beta^*_j) \times 100; \ \ i = 1, 2, \ldots, n \text{ and } j = 1, 2, \ldots, k \tag{10}$$

In this study, we consider Indian districts to be the firms or decision making units for which we would like to compute input-specific and overall technical efficiencies. We begin by specifying the following Cobb-Douglas production function:

$$\ln(Y_i) = \beta_0 + \beta_{1i} \ln(A_i) + \beta_{2i} \ln(L_i) + \beta_{3i} \ln(F_i) + \beta_{4i} \ln(T_i) + u_i \tag{11}$$

where Y is the value of agricultural output in district i in India, A is the gross cropped area, L is the total labor force devoted to agriculture, F is fertilizer input, T is the number of tractors (a proxy for machinery), and the β's are input specific response coefficients for each i^{th} district.[10] We then employ the procedure outlined above, calculating the overall and input specific technical efficiencies from the estimated β's.

2.2 Modeling Health as a Determinant of Technical Efficiency

To consider the determinants of technical efficiency, we then estimate the following efficiency equation:

$$TE_i = \zeta + a \ IMR_i + \gamma Z + \delta \ \ STATE + e_i \tag{12}$$

where TE represents overall or input specific technical efficiency for the i^{th} district, IMR is the district's rural infant mortality rate, and Z is a vector of variables including rural literacy rate, percentages of villages electrified, pucca roads (a proxy for both infrastructure as well as access to rural markets), agro-climatic conditions, level of irrigation and cropping intensity. STATE is a vector of dummy variables representing the state in which the district is situated and is used to capture institutional, economic, and agro-climatic factors that we cannot control for more explicitly.

As we are primarily interested in the value of *á* in the TE model, we need to be concerned with the possible endogeneity of health, either through omitted variable bias or simultaneity. In order to address this, we looked to find sources of exogenous variation in order to identify IMR in the above model (see Section 5c). We should note that literacy rates and/or other infrastructure variables may also be correlated with the unobservables as well. Given the paucity of data in general, it is rather difficult to find plausible instruments for all variables, which we *a priori* may deem to be endogenous. This is certainly a limitation of the study and is more carefully considered in Sections 5c and 6.

3. Data and Measures

Most of the production inputs and agricultural output data have been taken from Bhalla and Singh (2001), who provide average figures for 281 districts spanning 15 Indian states for the triennial years 1960-3, 1970-3, 1980-3, 1990-3. We have used the data for 1990-3, which is currently the most recent year for which both detailed district-level agricultural production and infant mortality data are available. Our output measure, the value of agricultural output, represents the price-weighted sum of output for 35 crops, which account for over 97% of the total value of agricultural production in India. Similarly, gross cropped area (in ‘000 hectares), our measure of land input, is calculated for the same 35 crops. Our measure for fertilizer use is the tonnage of fertilizer (NPK) consumed and our measure for machinery and capital is the number of tractors per district. For labor, we have used data from the Census of India, 1991. The total work force in agriculture is computed by adding the number of agricultural workers and cultivators. We have weighted males, females and children by 1, 2/3, and 1/3, respectively. Mean values by state for each of the input and output indicators can be found in *table 1*. One can see that there exists a great deal of heterogeneity across states and districts with respect to agricultural production.

Table 1: State-wise Mean Values for Input and Output Variables

State	Gross Cropped Area (000 hectare)	Fertilizer (tonnes)	Labor (male equiv)	Tractors (No.)	Value of Output (000 Rs)
Andhra Pradesh	835816.1	91078.3	1007394	1870.6	6496513
Assam	464199	4712.9	541289	92.4	3747417
Bihar	687740.8	39953.2	1267752	2809.3	3276013
Gujarat	623079.4	42419.9	413818	2670.9	3061675
Haryana	801181.7	84901.1	354971	13471.8	6972447
Karnataka	635578.1	44730.3	491274	1664	3663874
Kerala	407620.9	30641	409908	262.57	5104529
Madhya Pradesh	541618.6	19167.4	374807	920.2	2203879
Maharashtra	905097.7	54385.3	633583	1514.5	3340865
Orissa	872152.7	17928.5	628416	172.6	4337459
Punjab	682657.7	110233.8	297713	21076.1	7788069
Rajasthan	706352.2	14995.6	329802	3252.5	2166225
Tamil Nadu	632106.6	73826.9	1058039	1984.8	6549084
Uttar Pradesh	521149.6	46231.2	566272	4616.1	4117483
West Bengal	529564.5	62048.7	732473	701.5	5570756

Source (Basic Data): Bhalla and Singh (2001) and Census of India (1991).

For the second stage of our analysis, district level rural infant mortality rates are taken from Irudaya Rajan and Mohanachandran (1998). Rural literacy rates, the percentage of villages with pucca roads and electricity for agricultural use, are gleaned by the Census of India (1991) and Government of India (1997), respectively. The percentage of gross cropped area irrigated and cropping intensity (gross cropped area divided by net sown area) are calculated from Bhalla and Singh (2001). State-wise means for each of the second stage variables can be found in *table 2.*

Table 2: State-wise Sample Means for Overall, Land, and Labor TE Values and Other Indicators

State	IMR	Road	Agri. Elec.	Lit. Rate	Crop. Intensity	Irrig	**Overall TE**	**Land TE**	**Labor TE**
Andhra Pradesh	51.73	55.32	73.97	34.46	120.16	40.28	35.55	69.32	69.4
Assam	84.88	26.92	1.86	49.88	126.92	6.4	91.18	58.71	90.93
Bihar	71.28	25.62	25.32	33.9	132.93	36.53	26.39	72.02	63.74
Gujarat	69.48	61.37	85.19	52.9	106.54	26.54	36.74	70.2	67.72
Haryana	61.18	97.97	96.01	48.73	158.74	69.53	56.73	65.28	77.2
Karnataka	65.53	68.52	96.46	49.29	114.65	22.14	45.59	68.36	71.38
Kerala	35.54	98.63	96.34	89.57	137.58	14.4	79.69	62.23	83.78
Maharashtra	65.54	42.89	69.74	55.13	117.1	12.84	28.13	72.14	63.99
Madhya Pradesh	116.65	23.34	57.34	36.26	121.25	19.5	42.7	69.39	69.29
Orissa	113.47	22.89	21.12	42.53	148.23	28.83	64.22	63.64	80.56
Punjab	60.76	95.96	96.79	54.6	176.35	90.29	62.17	64.52	78.64
Rajasthan	90.69	31.11	48.19	28.89	122.92	26.06	43.8	69.91	68.36
Tamil Nadu	56.12	78.3	84.5	56.74	121.19	43.94	44.55	67.34	73.44
Uttar Pradesh	94.9	45.04	56.69	36.98	148.44	59.82	44.68	68.02	71.97
West Bengal	62.17	33.23	19.87	54.89	154.37	46.69	46.96	67.26	73.5

Source (Basic Data): Irudaya Rajan and Mohanachandran (1998), Census of India (1991), Government of India (1997), and Bhalla and Singh (2001).

Note: Efficiency values in last 3 columns are computed by authors.

In order to control for agro-climatic conditions, we have used rainfall data for the year 1991 (generously provided to us by Shenggen Fan at IFPRI), and the National Bureau of Soil Survey and Land Use Planning's (1992) grouping of districts into 20 agro-ecologic zones based on climate, topography, water resources, and soil type.[11]

Next, in Section 5c we consider the percentage of the district population from scheduled castes and scheduled tribes, percentage of villages having tap water and (any) medical facilities, the percentage of villages having a primary school, and the lagged level of irrigation, as instruments for IMR, literacy status, and irrigation. Scheduled caste/tribe and lagged irrigation percentages have been taken from Government of India (1994) and Bhalla and Singh (2001), respectively and the remaining variables from Government of India (1997).[12]

4. Results: Frontier Production Function Estimates

In column 1 of *table 3,* we provide OLS (for comparative purposes) estimates of the Cobb-Douglas production function. Results indicate that fertilizer, land area, and labor, in descending order of magnitude, are statistically significant determinants of agricultural production. Surprisingly, the number of tractors carries a negative elasticity, though this is not significantly different from zero at any acceptable level of confidence.

Table 3: Mean Response Coefficients, and Range of Estimates of the Actual Response Coefficients

	(1)	(2)	(3)	(4)
	OLS Estimate	Mean Response Coefficients	Maximum Value of Actual Response Coefficient	Minimum Value of Actual Response Coefficient
Constant	2.492 (2.979)	2.5863 (1.585)	2.586	2.586
Ln A	0.272 (4.219)	0.219 (2.219)	0.321	0.174
Ln F	0.364 (11.071)	0.423 (7.594)	0.434	0.419
Ln L	0.2506 (4.440)	0.2485 (2.236)	0.347	0.038
Ln T	-0.0004 (0.019)	0.0066 (0.206)	---	---
R^2 [F]	0.667 [127.9]	---		

Robust absolute t-statistics in parenthesis.

The random coefficient estimates (i.e., the mean response coefficients for the reduced form estimates using the iterative GLS procedure) are presented in the second column. We first note that we find the model to be valid (statistically significant at the 5% level or better), as ascertained by the Breusch-Pagan Lagrange multiplier test. The estimated mean response coefficients are generally in line with the OLS estimates. What we are really interested in here, however, is the variation in elasticity coefficients across districts. In the third and fourth columns of *table 3* we find that districts vary greatly in terms of production returns to inputs use and, consequently, the production process as a whole. In particular, there is a great deal of variation in returns to labor, with an order of magnitude difference between the district with the lowest response coefficient and the district with the highest. [13] Land elasticity also varies, though the range is smaller than that for labor.[14] For fertilizer use, while we do find statistically significant heterogeneity between districts, the distribution of response coefficients is quite tight.

5. Results: Technical Efficiency and Its Determinants

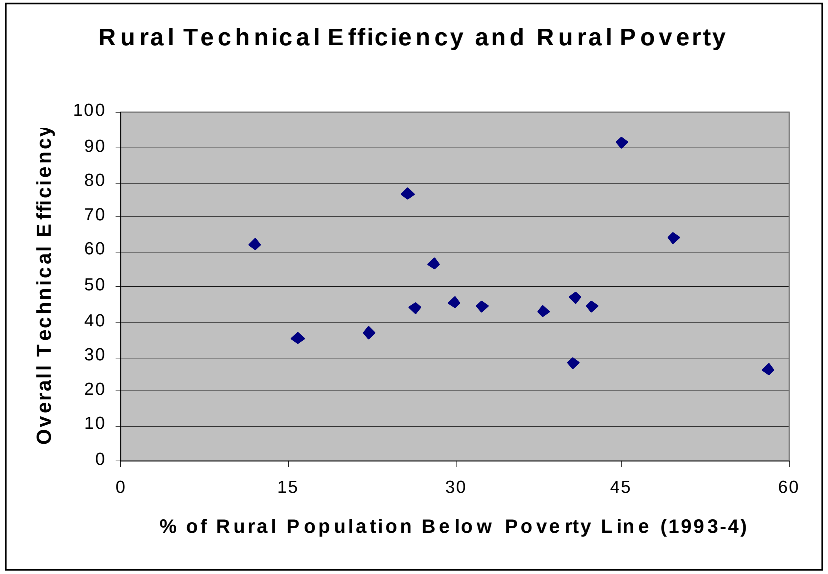
5a. Patterns in Technical Efficiency

Using the procedures outlined in the section 2, we computed the input specific efficiency as well as overall technical efficiency for each district in the sample. In the last three columns of *table 2*, we present state-wise mean values for overall efficiency, land efficiency, and labor efficiency, along with data for various socioeconomic indicators, which we use to explain the variation in efficiency. What should be immediately clear is that the efficiency-overall and specific input use- varies quite widely across the Indian states (and districts). Overall mean efficiency ranges from a low of 26.39, in Bihar, to a high of 91.18 in Assam.

If we look at districts by state and level of technical efficiency, we find that of the 261 districts in our sample, 181 (about 70%) have

overall efficiency scores below 50% (see Appendix I). Of these 181, 99 districts belong to four states: Bihar, Madhya Pradesh, Rajasthan, and Uttar Pradesh. These states are collectively well known for their (relatively) low per capita incomes and human development performance.

Figure 1



However, it is clear that inefficiency is not simply a problem for poorer states. A large number of districts in Maharashtra, Gujarat, and Tamil Nadu, all states with relatively high per capita incomes, perform quite poorly as far as using inputs effectively in agriculture. To present this point differently, we plot overall technical efficiency by the percentage of the rural population living below the poverty line (for the year 1993-4) in *figure 1*. It is clear from the scatter plot that there is no discernable relationship, one way or the other, between poverty and inefficiency.[15]

Another point to note from *table 2* is that overall efficiency seems to be determined more by labor efficiency than land (or fertilizer) efficiency. This can be seen from the GLS results in *table 3* as well, where we pointed out that labor use coefficients vary much more greatly than those on land and fertilizer. Indeed, we find that labor efficiency and overall technical efficiency are positively correlated while land efficiency is negatively correlated to both of these measures. The results appear to suggest that the two inputs are substitutes in terms of efficiency of use.[16]

5b. Determinants of Technical Efficiency

What, then, are the determinants of technical efficiency at the district level and does health play a role in generating more efficient use of inputs? In the first three columns of *table 4*, we present OLS estimates of the technical efficiency model presented in Section 2.[17] For overall efficiency, we find that the infant mortality rate has a negative and statistically significant effect across all specifications. *Our point estimate here suggests that a decline in IMR by 10 deaths per 1000 increases technical efficiency by 1.1 to 1.3 percentage points.* Greater cropping intensity is negatively associated with technical efficiency, with the effect being robust across all specifications. The positive association between literacy and irrigation loses significance with the addition of rainfall controls and agro-climatic zone fixed effects.[18]

Table 4

The Determinants of Overall Technical Efficiency

	(1)	(2)	(3)	(4)	(5)	(6)
IMR	-0.117	-0.112	-0.138	-0.119	-0.116	-0.143
	(2.13)	(2.02)	(2.15)	(2.16)	(2.09)	(2.26)
% Literate	0.223	0.186	0.119	0.458	0.441	0.332
	(2.3)	(1.94)	(1.18)	(2.86)	(2.56)	(2.29)
% Irrigated	0.145	0.138	0.096	0.42	0.423	0.333
	(2.44)	(2.21)	(1.41)	(3.18)	(2.96)	(2.52)
% Village Pucca Road	0.071	0.087	0.063	0.054	0.067	0.054
	(0.87)	(1.09)	(0.79)	(0.69)	(0.87)	(0.67)
% Village Electrified	-0.086	-0.067	-0.084	-0.098	-0.081	-0.095
	(1.98)	(1.44)	(1.33)	(2.28)	(1.76)	(1.51)
Cropping Intensity	-0.135	-0.149	-0.147	-0.137	-0.149	-0.148
	(3.1)	(3.47)	(3.09)	(3.13)	(3.4)	(3.02)
%Irrig*%Literate				-0.007	-0.007	-0.006
				(2.31)	(2.18)	(2.23)
State FE	yes	yes	yes	yes	yes	yes
District Rainfall	no	yes	yes	no	yes	yes
Agroclimatic Zone FE	no	no	yes	no	no	yes
Observations	261	258	253	261	258	253
R-squared	0.62	0.61	0.65	0.63	0.62	0.66

Robust absolute t-statistics in parenthesis

It is possible that the marginal association between any given independent variable and changes with the magnitude of one of the other independent variables. We explore the possibility of three such interaction effects: that between health and irrigation, health and cropping intensity, and literacy and irrigation. The rationale behind the first and second is that the impact of health on efficiency may decrease (increase) as agricultural work becomes less (more) strenuous. The rationale behind the last is that education may become more important as the technology of agricultural production becomes more complicated. Alternatively, it may be that education is more important in marginal conditions, where firms are the whim of seasonal variations in rainfall.

We present the results for this analysis in the final three columns of *table 4*. Of the three postulated effects, only the interaction between literacy rates and irrigation was found to be statistically significant. The results indicate that literacy is somewhat less salient as the percentage of land under irrigation increases. Most important for the purposes of this study, however, is that the coefficient on IMR remains significant regardless of the change in specification.

Finally, in *table 5*, we present OLS results of the equations for labor efficiency.[19] The first thing that should be noted here is how, in general, the coefficients of the independent variables are qualitatively similar to those in the overall efficiency equation. It is clear here that a lower IMR is associated with increased efficiency of labor use and that this association is robust to the various specifications. This is very much in line with the extensive micro-literature on wages returns to investments in health and nutrition. Percent of gross cropped area irrigated is also associated with increased labor efficiency, perhaps indicating complementarities between technical factors and human capital. Finally, literacy appears to have a positive effect on labor efficiency, though the estimates are not robust to changes in specification.

Table 5: The Determinants of Labor Use Efficiency

	(1)	(2)	(3)	(4)	(5)	(6)
IMR	-0.057	-0.052	-0.063	-0.058	-0.055	-0.066
	(2.11)	(1.93)	(2.15)	(2.14)	(2.02)	(2.26)
% Literate	0.051	0.025	0.005	0.196	0.172	0.099
	(1.13)	(0.58)	(0.12)	(2.17)	(1.83)	(1.51)
% Irrigated	0.073	0.071	0.052	0.244	0.235	0.156
	(2.35)	(2.2)	(1.55)	(2.75)	(2.51)	(2.45)
Road	0.017	0.027	0.027	0.006	0.015	0.023
	(0.42)	(0.67)	(0.68)	(0.15)	(0.38)	(0.57)
% Village Electrified	-0.029	-0.018	-0.033	-0.037	-0.026	-0.038
	(1.00)	(0.56)	(1.06)	(1.36)	(0.88)	(1.23)
Cropping Intensity	-0.02	-0.029	-0.038	-0.022	-0.03	-0.038
	(0.77)	(1.20)	(1.55)	(0.87)	(1.23)	(1.58)
%Irrig*%Literate				-0.004	-0.004	-0.003
				(2.28)	(2.03)	(2.04)
State FE	yes	yes	yes	yes	yes	yes
District Rainfall	no	yes	yes	no	yes	yes
Agroclimatic Zone FE	no	no	yes	no	no	yes
Observations	261	258	253	261	258	253
R-squared	0.47	0.46	0.57	0.48	0.47	0.58

Robust absolute t-statistics in parentheses

5c. Endogeneity and Other Robustness Checks

Despite the robustness of the results across various specifications, it is entirely possible that the estimated coefficient on IMR reflects the presence of unobserved factors that are correlated with both IMR and technical efficiency. For example, it may be that there are certain institutional arrangements that affect the efficiency of both agricultural and health production. Or technical efficiency and population health may be simultaneously determined.

In either case, econometric estimation using OLS would lead to biased and inconsistent estimates of the coefficient on IMR. To get at this potential problem, we looked for a source of exogenous variation that is correlated with health but (plausibly) uncorrelated with unobservables in the production efficiency equation. Finding such an instrumental variable (IV) is difficult: there is no guarantee that the typical instruments for health employed at the micro-level (commodity prices, health care institutions, public health infrastructure such as tap water availability) are indeed exogenous at a more aggregate level. Even when we looked past this concern and employed the percentage of villages with tap water and medical facilities, we found that these variables did not meet another criterion of a good IV: both variables were only weakly correlated with health status and insignificant in the reduced form equation for IMR.[20]

Perhaps, a more attractive set of IVs are the percentage of scheduled castes and scheduled tribes in the district population. The idea is that members of scheduled castes and scheduled tribes are generally in poorer health – due to increased poverty, discrimination, or some combination of the two, within these groups - but that the link between the district share of these populations and technical efficiency in agricultural production is less clear. Also, in the reduced form equations for IMR, the scheduled caste percentage is positively correlated with IMR (not shown here).[21] Using these instruments, we are able to reject the endogeneity of IMR in both the overall efficiency and labor efficiency models using the Hausman (1978) test.[22]

These findings should be interpreted with some caution. First, it may be that the scheduled caste/scheduled tribe variables are indeed somehow correlated with both health and production efficiency. For example, the same factors that lead to differential caste based discrimination may be correlated with unobserved institutional characteristics that influence production efficiency. Second, other variables in technical efficiency equation may be endogenous as well. While we were able to instrument for literacy rates and irrigation levels using the percentage of villages with primary schools and lagged percentage of area irrigated (and, again, we were unable to reject the null hypothesis of exogeneity for either), these instruments may not be satisfactory for all the reasons presented above. Unfortunately, given data constraints and the paucity of clearly acceptable instruments, we are unable to do better in this study.

A final consideration involves our choice of methodology. It is possible that an alternate set of methods to calculate technical efficiency will yield different empirical results. To address this possibility, we have used the stochastic frontier method (see section 2.1) to calculate overall technical efficiencies and used these estimates as dependent variables for a second-stage analysis. We present these results in *table 6*. The results are qualitatively similar to those presented in *table 4* and further validate our modeling approach.

Table 6: The Determinants of Overall Technical Efficiency Calculated Using Stochastic Frontier Method

	(1)	(2)	(3)	(4)	(5)	(6)
IMR	-0.263	-0.238	-0.253	-0.267	-0.246	-0.265
	(2.76)	(2.5)	(2.3)	(2.81)	(2.59)	(2.43)
% Literate	0.127	0.018	-0.019	0.692	0.531	0.444
	(0.76)	(0.11)	(0.11)	(2.44)	(1.71)	(1.63)
% Irrigated	0.348	0.348	0.303	1.008	0.921	0.819
	(2.75)	(2.59)	(1.97)	(3.87)	(3.19)	(2.93)
Road	0.23	0.265	0.244	0.189	0.225	0.223
	(1.52)	(1.76)	(1.67)	(1.29)	(1.55)	(1.52)
% Villages Electrified	-0.138	-0.095	-0.081	-0.166	-0.123	-0.105
	(1.78)	(1.12)	(0.72)	(2.16)	(1.45)	(0.93)
Cropping Intensity	-0.155	-0.189	-0.21	-0.162	-0.19	-0.214
	(1.92)	(2.32)	(2.22)	(1.99)	(2.3)	(2.18)
%Irrig*%Literate				-0.016	-0.014	-0.013
				(2.95)	(2.26)	(2.32)
State FE	yes	yes	yes	yes	yes	yes
Rainfall	no	yes	yes	no	yes	yes
Agroclimatic Zone FE	no	no	yes	no	no	yes
Observations	261	258	253	261	258	253
R-squared	0.62	0.6	0.63	0.63	0.62	0.64

Robust t-statistics in parentheses

5d. Magnitude of Health-Mediated Gains in Efficiency

In practical terms, what do our findings mean for Indian districts? That is, just how potent is health in promoting efficiency in input use? In *table 7*, we compute the overall efficiency gains from improvements in IMR, literacy, and irrigation for the 15 *least efficient* districts in the sample, using the coefficients obtained from column (6) of *table 4*. The efficiency gains were computed by first calculating the difference between the given district's IMR, literacy rate, and percent irrigated land, and a target value for each of these variables, which was selected by choosing the lowest and highest values for each of this indicators, respectively, in the overall sample of districts. We then multiplied this difference by the coefficients computed in *table 4*.

Table 7: Efficiency Gains from Improvements in IMR, Literacy and Irrigation for 15 Most Inefficient Districts

District	State	Overall Efficiency	Gains from IMR	Gains from Literacy and Irrigation	Total Gains
Jaisalmer	Rajasthan	4.91	9.30	14.38	23.68
Bikaner	Rajasthan	15.21	6.44	15.85	22.29
Bhagalpur	Bihar	17.40	6.44	18.54	24.97
Muzaffarpu	Bihar	18.76	7.09	17.83	24.92
East Nimar	Madhya Pradesh	19.49	13.01	18.11	31.12
Monghyr	Bihar	20.51	6.96	18.35	25.31
Bangalore	Karnataka	20.80	4.58	19.69	24.27
Ahmednaga	Maharashtra	21.08	3.29	21.70	24.99
Nanded	Maharashtra	21.44	5.86	18.27	24.14
Gaya	Bihar	22.35	6.75	18.21	24.96
Nashik	Maharashtra	22.47	5.86	20.94	26.80
Jalgaon	Maharashtra	22.70	6.86	22.73	29.59
Sholapur	Maharashtra	23.10	4.29	20.99	25.28
Satna	Madhya Pradesh	23.19	17.02	18.63	35.64
Pune	Maharashtra	23.31	2.86	22.24	25.10

In general, we see that efficiency gains from health are lower than (combined) gains from improved literacy and irrigation. However, these gains are certainly not trivial. For example, the most inefficient district in the sample (Jaisalmer) would more than *triple* its efficiency with a decrease in IMR to the pre-specified target value.

6. Discussion and Conclusions

Over 60% of India's population relies on agriculture to make a living. In recent years, the low productivity and lack of efficiency in this sector has concerned many policymakers, especially in the context of increasing global competition. Given the large labor force concentrated in agriculture, and the many forward and backward linkages between agriculture and other sectors, ensuring the prosperity of this sector is of paramount importance.

At the same time, a large number of empirical studies at the micro and macro levels have shown that there are potentially large economic benefits from improvements in individual and population health, respectively. In particular, many of the micro studies have focused on the effect of various health status indicators on agricultural wages or farm level production.

In this study, we have considered the impact of health on agricultural production at a more aggregate level. In doing so, we have diverged from the typical approach of including health status variables as inputs in the production function. Rather, we have modeled health (IMR) as a factor which influences the efficiency of the production process, a strategy we feel better captures the role of health human capital in the production process.

We have found that Indian districts are quite inefficient on the whole in producing agricultural goods, and that the range of (in) efficiency across states and districts is quite wide. We have also found that the level of rural poverty in a given state has no bearing on the average efficiency of districts within that state. There is an important policy point here: inefficiency in agricultural is not just a problem of the poor.

Going back to the main objectives of this study, we have found that better health has a practically and statistically significant impact on technical efficiency in agricultural production, and that a large portion of this effect is likely due to the augmentation of labor efficiency. These results are not only consistent with the growing body of cross-country and micro-level studies on the economic benefits of good health, but are also in line with emerging literature in India which illustrates the role of health in reducing poverty and promoting economic growth (Mitra et al., (2002), and Gupta and Mitra (2003)). In the context of this literature, our study adds to the argument that population health investment is an important component of policy packages seeking to promote better economic performance (both in growth and efficiency terms) and reduce poverty.

There are several limitations in our study. First, we consider a production function, which essentially aggregates across a large, diverse set of crops. To the extent that the efficiency of production of different crops varies (both across crops and within districts) our estimates of technical efficiency may be imprecise. While we try to control for crop mix in the equations for technical efficiency, we are constrained from disaggregating the production function due to limitations in the available data.

Second, we are unable to properly address any potential endogeneity issues due to the paucity of acceptable instruments. While we are able to reject the endogeneity of IMR using the percentage of the population belonging to scheduled castes and tribes, it would perhaps be desirable to find other instruments, which are perhaps more clearly exogenous, in order to be more certain of this conclusion.

Despite these limitations, our study proposes an interesting mechanism through which health can exert a positive influence on economic performance. Future studies should seek to further define the role of health in influencing efficiency in production.

Appendix I District Specific Efficiency Values

District	Overall TE	Land Eff.	Fertilizer Eff.	Labor Eff.	District	Overall TE	Land Eff.	Fertilizer Eff.	Labor Eff.
Andra Pradesh					Bijapur	40.39	67.25	97.39	73.43
Adilabad	28.19	72.42	97.81	63.28	Chikmag	65.04	65.05	97.19	77.94
Anantapur	41.82	67.15	97.38	73.72	Chitradurg	42.91	67.82	97.44	72.39
Chitoor	48.24	66.57	97.33	75.01	Dak. Kan.	71.45	63.94	97.10	80.41
Cuddapah	34.16	70.49	97.68	67.00	Dharwad	27.65	71.56	97.76	65.21
East Godar	37.62	68.48	97.50	71.09	Gulbarga	26.51	72.31	97.80	63.87
Guntur	34.23	66.61	97.32	74.73	Hassan	45.58	68.11	97.46	71.85
Karimnagar	41.97	67.78	97.43	72.51	Kodagu	100.00	58.55	96.59	90.90
Khammam	39.01	68.95	97.54	70.14	Kolar	25.96	73.32	97.92	61.18
Krishna	37.77	68.51	97.50	71.05	Mandhya	41.54	69.10	97.55	69.80
Mahbubnagar	23.99	72.76	97.87	62.66	Mysore	36.23	69.30	97.57	69.44
Medak	28.97	71.88	97.79	64.12	Raichur	26.58	72.34	97.84	63.72
Nalgonda	31.88	70.40	97.67	67.25	Shimboga	54.56	66.18	97.29	75.79
Nizambad	30.94	71.63	97.80	64.56	Tumkur	46.98	66.77	97.35	74.52
Srikakulam	38.97	66.82	97.35	74.41	Uttarkannada	71.18	65.76	97.27	76.89
Assam					***Kerala***				
Silchar	100.00	58.59	96.79	91.63	Alappuzha	74.21	64.48	97.15	79.36
Darrang	100.00	56.97	96.72	94.50	Ernakulam	63.19	63.46	97.04	81.00
Goalpara	77.35	62.08	97.05	84.14	Kannur	100.00	57.51	96.58	92.95
Kamrup	83.19	60.87	96.95	86.23	Kollam	82.29	62.18	96.95	84.00
N. Lakhimpur	100.00	54.20	96.50	100.00	Kozhikode	68.45	62.26	96.95	83.46
Nagaon	81.75	61.95	97.00	84.40	Thirssur	80.45	63.46	97.06	81.45
Jorhat	100.00	56.30	96.61	95.61	Thiruvan.	90.26	62.26	96.97	84.23
Bihar					***Maharastra***				
Bhogalpur	17.40	76.60	98.22	53.90	Ahmadnagar	21.08	74.29	98.01	59.98
Bhojpur	34.59	68.83	97.53	70.41	Akola	31.37	70.68	97.68	66.90
Darbhanga	23.69	72.10	97.80	63.65	Amravati	30.88	71.04	97.71	66.17
Dumka	28.69	71.13	97.70	65.54	Aurang.	25.27	72.05	97.80	64.38
Gaya	22.35	72.86	97.89	62.14	Bhandara	29.37	71.93	97.78	64.08
Hazar.	25.67	72.32	97.79	62.94	Bid	30.75	71.17	97.71	65.98
Munger	20.51	72.92	97.89	62.10	Buldana	26.86	72.58	97.84	63.14
Muzaffar.	18.76	74.45	98.01	58.83	Chandrap.	30.65	70.96	97.69	66.21
P. Champ	32.06	69.64	97.60	68.72	Dhule	27.69	72.13	97.81	63.93
Palamu	23.64	74.26	97.94	58.68	Jalgaon	22.70	73.72	97.98	60.92
Patna*	24.35	72.85	97.89	62.06	Kolhapur	40.21	68.70	97.52	70.64
Purnia	23.73	72.09	97.80	63.81	Nagpur	26.93	73.20	97.90	61.85
Ranchi	27.21	72.05	97.77	63.89	Nanded	21.44	75.59	98.13	57.46
Saran	35.30	68.80	97.52	70.46	Nashik	22.47	73.75	97.96	60.76
Singhbum	38.05	69.46	97.56	69.03	Osman	28.88	70.92	97.70	66.49
Gujarat					Parbhani	28.94	71.42	97.73	65.56
Ahmadabad	29.75	72.62	97.86	63.04	Pune	23.31	73.43	97.93	61.62
Amreli	49.75	67.37	97.40	73.26	Ratnagiri	40.78	68.73	97.51	70.58
Ban. Kanth	29.72	71.73	97.77	64.96	Sangli	24.48	74.01	98.00	60.23
Bharuch	40.79	69.41	97.57	69.25	Satara	33.26	70.49	97.67	67.15
Bhavnag.	32.25	71.44	97.75	65.40	Sholapur	23.10	73.71	97.95	61.12
Jamn.	35.02	71.19	97.72	65.97	Thane	28.54	72.54	97.83	62.65
Junagadh	44.73	67.60	97.42	72.84	***Madhya Pradesh***				
Kachchh	50.13	67.34	97.40	73.28	Balaghat	37.22	70.21	97.63	67.57
Kheda	36.92	69.22	97.56	69.62	Bastar	68.71	62.41	97.09	82.99
Mahesana	38.42	68.74	97.52	70.61	Betul	29.13	72.90	97.85	62.40
P. Mahals	26.14	72.67	97.84	62.58	Bhind	62.84	72.87	97.84	62.36
Rajkot	30.08	72.14	97.82	64.13	Bilaspur	37.68	68.00	97.45	72.05
Sab. Konth	31.00	71.98	97.81	64.16	Chhatarp.	37.49	70.57	97.65	66.96
Surat	42.12	68.50	97.50	71.05	Chhindw	59.40	64.92	97.21	78.11
Surendr.	41.48	69.33	97.56	69.52	Damoh	50.32	68.25	97.47	71.56
Vadodara	33.64	70.59	97.68	66.89	Datia	58.50	68.32	97.48	71.42
Valsad	32.66	71.55	97.77	64.77	Dewas	43.67	68.91	97.53	70.27
Haryana					Dhar	27.14	73.23	97.91	61.83
Ambala	60.05	65.76	97.24	76.43	Durg	32.55	69.29	97.56	69.55
Gurgaon	46.68	67.17	97.38	73.66	E. Nimas	19.49	77.10	98.25	54.09
Hisar	61.73	61.80	96.88	83.49	Guna	54.38	65.95	97.30	75.97
Jind	59.59	65.95	97.27	76.02	Gwalior	52.86	67.81	97.44	72.44
Karnal	54.22	64.82	97.15	78.08	Hoshang.	42.03	68.93	97.53	70.25
Mahendragarh	58.14	66.19	97.30	75.50	Indore	46.17	68.79	97.52	70.50
Karnataka					Jabalpur	28.04	72.92	97.85	62.30
Bangalore	20.80	75.52	98.13	56.72	Jhabua	39.93	69.62	97.58	68.81
Belgaum	48.34	65.71	97.25	76.51	Mandla	41.32	68.81	97.51	70.44
Bellary	28.25	72.34	97.85	63.58	Mandsaur	32.32	71.08	97.72	66.09
Bidar	47.31	67.96	97.45	72.13	Morena	50.99	66.83	97.35	74.39

District	Overall TE	Land Eff.	Fertilizer Eff.	Labor Eff.	District	Overall TE	Land Eff.	Fertilizer Eff.	Labor Eff.
Madhya Pradesh					Udaipur	51.47	66.53	97.35	75.07
N'Simpur	60.05	66.00	97.29	76.00	Tamil Nadu				
Panna	44.08	69.91	97.60	68.22	Cheng. MGR	49.87	66.66	97.33	74.87
Raigarh	37.40	69.49	97.57	69.09	Coimbatore	41.34	67.35	97.39	73.40
Raipur	43.52	66.31	97.31	75.35	Kanniyakumari	72.91	65.28	97.21	78.05
Raisen	60.03	65.47	97.25	76.90	Madurai	35.11	68.98	97.54	70.05
Rajgarh	35.24	71.19	97.71	65.84	N. Arcot Ambed.	41.25	67.66	97.42	72.79
Ratlam	31.24	72.77	97.86	62.71	Ramanath	39.06	68.48	97.49	71.09
Rewa	23.69	75.06	98.05	58.02	S. Arcot	53.30	64.79	97.16	78.63
Sagar	37.69	70.30	97.64	67.60	Salem	49.94	64.80	97.17	78.48
Satna	23.19	75.48	98.07	57.17	Thanjavur	41.78	66.77	97.34	74.50
Sehore	50.15	66.74	97.35	74.46	Tiru. Kalla	34.06	70.08	97.64	67.77
Seoni	47.80	67.96	97.45	72.13	Tiruchch	31.47	69.89	97.63	68.21
Shahdol	37.42	69.86	97.58	68.36	Uttar Pradesh				
Shajapur	55.68	65.76	97.27	76.37	Agra	46.60	67.72	97.43	72.61
Shivpuri	54.16	66.50	97.34	75.01	Aligarh	42.27	68.07	97.46	71.92
Sidhi	32.60	71.70	97.71	64.68	Allahabad	32.14	70.14	97.64	67.73
Surguru	41.62	68.17	97.47	71.72	Azamgarh	41.12	67.84	97.44	72.37
Tikamgam	38.10	70.67	97.68	66.73	Bahraich	38.53	68.57	97.50	70.92
Ujjain	45.15	67.95	97.45	72.14	Ballia	33.11	71.18	97.74	65.55
Orissa					Banda	49.14	66.70	97.35	74.64
Balangir	77.34	61.68	96.99	84.40	Bar. Ban.	34.04	70.38	97.66	67.20
Baleshwar	39.23	68.43	97.49	71.20	Bareilly	41.37	68.55	97.50	70.95
Cuttack	55.15	62.84	97.04	81.95	Basti	32.91	69.52	97.59	69.02
Dhenkanal	84.20	60.59	96.92	86.44	Bijnor	70.77	63.32	97.02	81.48
Ganjam	47.53	65.88	97.28	76.19	Budaun	35.90	69.67	97.60	68.73
Kalahandi	60.81	61.73	96.99	83.96	Bulandsh.	47.72	66.96	97.36	74.12
Kendujhar	68.86	63.89	97.17	80.21	Dehra Dun	59.46	68.56	97.50	70.91
Mayurb	59.21	64.74	97.22	78.55	Deoria	36.56	68.89	97.53	70.26
Phulanb.	96.63	60.57	96.97	86.84	Etah	48.69	66.95	97.36	74.16
Sambalp.	48.82	65.76	97.26	76.38	Etawah	50.49	66.66	97.34	74.78
Sundargarh	68.67	63.95	97.17	80.04	Faizab.	35.89	69.81	97.62	68.37
Punjab					Farrukha.	55.80	65.75	97.25	76.66
Amritsar	54.21	65.32	97.20	77.19	Fatepur	41.41	68.95	97.54	70.14
Bhatinda	65.42	62.43	96.93	82.50	Ghazipur	38.56	69.63	97.60	68.74
Firozpur	66.32	61.38	96.82	84.33	Gonda	40.13	67.88	97.44	72.31
Gurdasp	51.79	66.90	97.35	74.21	Gurakhp.	37.16	68.60	97.50	70.86
Hoshiarpur	54.85	66.79	97.34	74.42	Hamirpur	60.83	64.90	97.22	78.14
Jalandh.	64.22	64.35	97.11	79.12	Hardoi	46.49	66.98	97.37	74.12
Kapurthal.	70.61	65.16	97.18	77.60	Jalaun	58.95	66.04	97.29	75.97
Ludhiana	60.95	64.60	97.12	78.58	Jaunpur	36.85	69.72	97.61	68.56
Patiala	66.65	63.19	96.99	81.19	Jhansi	48.31	66.98	97.37	74.04
Ropar	53.66	68.45	97.49	71.16	Kheri	51.72	65.82	97.25	76.41
Sangrur.	75.23	61.27	96.82	84.75	Kanpur*	47.86	66.80	97.35	74.47
Rajasthan					Lucknow	30.87	72.96	97.90	61.69
Ajmer	32.03	72.26	97.77	63.69	Mainpuri	49.22	67.07	97.37	73.94
Alwar	50.88	66.25	97.32	75.44	Mathura	52.02	66.91	97.36	74.23
Banswara	35.11	71.38	97.73	65.23	Meerut	51.45	64.96	97.18	77.94
Barmer	33.99	70.64	97.60	67.21	Mirzapur	32.12	71.04	97.70	65.94
Bharatpur	49.06	66.73	97.35	74.54	Moradab	45.98	66.72	97.33	74.61
Bhilwara	34.89	70.69	97.67	66.71	Muzzafarn	66.03	63.76	97.06	80.62
Bikaner	15.21	82.14	98.68	45.95	Nainital	49.24	67.87	97.44	72.32
Bundi	46.00	69.23	97.56	69.63	Pilibit	54.86	66.60	97.32	74.87
Chittaurg	39.87	69.06	97.54	69.95	Pratapg	35.35	70.58	97.68	66.75
Churu	85.80	59.45	97.07	87.59	Rae Boreli	34.57	70.51	97.67	66.95
Dungarpur	49.28	69.03	97.53	69.92	Rampur	48.72	67.97	97.45	72.14
Ganganagar	92.45	57.34	96.62	91.57	Saharanpur	57.79	64.88	97.16	78.27
Jaipur	47.17	66.37	97.33	75.19	Shahjahanp.	48.22	67.00	97.36	74.06
Jaisalmer	4.91	100.00	100.00	11.07	Sitapur	42.40	67.92	97.45	72.22
Jalor	34.16	71.44	97.70	65.54	Sultanpur	37.32	69.64	97.59	68.72
Jhalawar	40.03	69.99	97.61	68.12	Unnao	34.18	70.67	97.68	66.62
Jhunjhunun	41.21	69.66	97.57	68.85	Varanasi	32.85	70.66	97.70	66.67
Jodhpur	34.85	70.40	97.62	67.58	West Bengal				
Kota	43.18	68.59	97.50	70.89	24 Parag.	36.28	68.02	97.45	72.02
Nagaur	40.20	68.15	97.47	71.75	Bankura	57.30	64.99	97.18	78.21
Pali	33.26	71.69	97.74	65.03	Birbhum	40.17	68.68	97.51	70.69
S. Madhop	58.38	64.99	97.22	77.92	Burdwan	60.58	63.77	97.07	80.55
Sikar	38.11	70.28	97.61	67.67	Haora	33.31	72.23	97.86	63.06
Sirohi	49.33	69.54	97.57	68.99	Hugli	54.13	65.87	97.25	76.46
Tonk	57.98	65.84	97.30	76.19					

Foot Notes

1. This measure allows us to compare efficiency across similar economic units such as firms/ aggregation of firms, for example at industry level or geographically. If efficiency varies across units, then further analysis can reveal the factors that cause such variations. Such analysis has important policy implications for the improvement of efficiency (Kalirajan and Shand, 1994).

2. The cross-country study by Bloom et al., (2004), models health, experience, and education, as labor augmenting inputs in an aggregate Mincer production model.

3. Croppenstedt and Muller (2000) consider the production effects of weight-for-height and morbidity in a firm-level study on Ethopian agriculture. While they employ the stochastic frontier method to calculate technical efficiency, they still model health as an input in the production function. In another part of the study, the effect of height and weight-for-height on wages, assuming the latter fully reflects the productivity of labor, is estimated. Croppenstedt and Demeke (1997) use the random-coefficients approach to estimate input-specific and overall technical efficiencies for Ethopian farms. While they do allude to the possibility that health and nutrition can drive differential labor efficiency, they do not model these explicitly in their technical efficiency equations. Mitra et al., (2002), study the effect of various infrastructure on total factor productivity (TFP) in Indian industry at the state-level. After calculating TFP for different industrial categories, they study its determinants by considering the effect of an aggregate infrastructure index (which includes IMR as one of its components). Decomposing the effect of each infrastructure variable on industry-specific TFP, they find IMR to be an important determinant of TFP in several industries. However, because of multicollinearity they are unable to estimate the returns to this variable explicitly.

4. All these extensions require the functional form of the frontier and distribution of the one-sided residual term to be specified. This can result in errors of mis-specification if the above specifications are incorrect.

5. Their approach is a slight modification of the Hildreth and Houck (1968) model, which is a special case of Swamy's (1971) panel data model.

6. $E(w_i) = 0$; $V(w_i) = \sigma^2 + \Sigma_j \sigma_{ii} x_{ij}^2$ and $Cov(w_i w'_i) = 0$ for $i \neq i'$.

7. Hildreth and Houck's (1968) random coefficient model belongs to this class of heteroscedastic error models.

8. See Griffiths (1972).

9. In special cases of production process in which constant returns to scale are imposed on the individual response coefficients, β_{ij} , the estimation of $\beta_j^{*'s}$ would be complicated and intractable. Even when the condition of constant returns to scale is imposed on the mean response coefficients, $\overline{\beta}_j$'s then due to the relationship

$$\beta_j^* = \max_i (\overline{\beta}_j + \hat{v}_{ij})$$

the possibility that $\sum \overline{\beta}_j^* > 1$ cannot be ruled out. In either case, the problem that remains is that the "best practice" production outcome might not be feasible if all production processes had to have constant returns to scale by some strict technical rule.

10. We also tried to use the more flexible trans-log specification, but found that the Cobb-Douglas approach fit the data best.

11. Also, we attempt to control for crop mix by designating districts as rice, wheat, or "other districts", based on plurality of land use towards a particular crop. We have constructed these indicators from CMIE (2000). We have also tried to control for institutional climate using per capita credit to agricultural enterprises as a proxy variable. This measure is taken from CMIE (2000) as well.

12. Because the number of districts and district boundaries has changed from census year to census year, roughly a quarter of Bhalla and Singh's district units are actually amalgams of currently existing districts. For that reason, we use appropriate population, area, and village weights to align our second stage data with the district combinations used by Bhalla and Singh.

13. A wide variation in labor elasticity was also computed in a firm-level of study of Ethiopian farms conducted by Croppenstedt and Demeke (1997). The lowest and highest computed values in their study were 0.04 and 0.37, which is quite similar to our results. They postulate the wide variation in labor elasticity is due to differential labor quality (such as health status). They do not, however, explicitly control for health and nutritional status as a determinant of labor (or overall) technical efficiency.

14. The distribution of mean response coefficients for land falls below those typically found in the literature (0.4-0.6, see Kalirajan and Obwona, 1994). This could perhaps suggest that farmers are not necessarily constrained by the availability of land.

15. The correlation between the two variables is -0.054, and is not statistically significant.

16. Neither production theory nor the random coefficients estimation method imposes any restrictions on the input coefficients, such that a low elasticity for one variable means a higher elasticity for another. Hence, it is quite interesting that we are seeing a negative correlation between input efficiencies in our results.

17. Many of the state-level control variables (not shown here) are statistically significant, indicating that the dummy variables are capturing a number of determinants, which we cannot adequately control for more explicitly. Some of the state-level effects are

especially large: Assam and Orissa - two of the poorest performing states in our sample with respect to poverty and human development - sport fairly gaudy coefficients, relative to the other states. It is likely that available data does not quite capture the differential state-wise environmental and institutional factors and, as a result, we are limited in our attempts to disaggregate the large unexplained effects that state membership confers on technical efficiency.

18. As mentioned in Section 2, we also attempted to control for state level institutional characteristics, as proxied for by per capita credit to agriculture enterprises, and regional crop mix (rice, wheat, or “other” plurality in land use). Analysis using these variables was quite unrevealing: none of these were statistically significant, and their inclusion did not change the coefficients on any of the other variables appreciably. Hence, we do not present these estimates in the ensuing tables.

19. We do not show models for land and fertilizer efficiency, as the results were quite unrevealing: few of the explanatory variables were statistically significant.

20. See Bound et al., (1995) for a discussion on the difficulties that arise from using weak instruments.

21. We see that a 1 percentage point increase in the scheduled caste percentage is associated with an increase in IMR of nearly 0.8 deaths per 1000.

22. The 2SLS instruments were nearly identical to the OLS estimates presented in Table 4.

References

Acemoglu, D. and Smith (2006), Disease and Development: The Effect of Life Expectancy on Economic Growth, *NBER Working Paper 12269*.

Aigner, D.J., C.A.K. Lovell and P. Schmidt (1977), Formulation and Estimation of Stochastic Frontier Production Function Models, *Journal of Econometrics* 6 (1): 21-37.

Alderman, H., J.Hoddinot, and B. Kingsley (1996), The Returns to Endogenous Human Capital in Pakistan’s Rural Wage Labor Market, *Oxford Bulletin of Economics and Statistics* 58 (1): 29 – 55.

Alderman, H., J.R. Behrman, V. lavy, and R. Menon (2001), Child Health and School Enrollment: A Longitudinal Analysis. *Journal of Human Resources* 36 (1): 185 – 205.

Alsan, M., D.E. Bloom, and D. Canning (2004), The Effect of Population Health on Foreign Direct Investment. *NBER Working Paper 10596*.

Barro, R. and X. Sala-i-Martin (2004), *Economic Growth*, 2nd ed. Cambridge, Mass: The MIT Press.

Battese, E. C. (1992), Frontier Production Functions and Technical Efficiency: A Survey of Empirical Applications in Agricultural Economics, *Agricultural Economics* 7 (3): 185-208.

Bauer, P. W. (1990), Recent Developments in the Econometric Estimation of Frontier, *Journal of Econometrics* 46: 39-56.

Behrman, J.R., and M.R. Rosenzweig (2004), Returns to Birthweight, *Review of Economics and Statistics* 86 (2): 586 – 601.

Bhalla, G.S., and G. Singh (2001), *Indian Agriculture: Four Decades of Development,* New Delhi, India and London, England: Sage Publications.

Bhargava, A., D.T. Jamison, L.J. Lau, and C.J.L. Murray (2001), Modeling the Effects of Health on Economic Growth, *Journal of Health Economics* 20 (3): 423-440.

Bloom, D.E. and D. Canning (2000), The Health and Wealth of Nations, *Science* 287: 1207-1209.

Bloom, D.E., D. Canning, and B. Graham (2003), Longevity and Life-Cycle Savings, *Scandinavian Journal of Economics* 105 (3): 319 – 338.

Bloom, D.E., D. Canning, and J. Sevilla (2004), The Effect of Health on Economic Growth: A Production Function Approach, *World Development* 32 (1): 1 – 13.

Bound, J.D., A. Jaeger, and R.M. Baker (1995), Problems with Instrumental Variables Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak, *Journal of the American Statistical Association* 90: 443 – 50.

Center for Monitoring the Indian Economy (CMIE) (2000), *Profiles of Districts.* Mumbai: CMIE.

Croppenstedt, A. and M. Demeke (1997), An Empirical Study of Cereal Crop Production and Technical Efficiency of Private Farmers in Ethiopia: A Mixed Fixed-Random Coefficients Approach, *Applied Economics* 29 (9): 1217 – 1226.

Croppenstedt, A., and C. Muller (2000), The Impact of Farmers' Health and Nutritional Status on Their Productivity and Efficiency: Evidence from Ethiopia, *Economic Development and Cultural Change* 48 (3): 475 – 502.

Deolalikar, A. (1988), Nutrition and Labor Productivity in Agriculture, *Review of Economics and Statistics* 70 (3): 406-413.

Farrell, M.J. (1957), The Measurement of Productive Efficiency, *Journal of Royal Statistical Society* 120 (3): 253-81.

Gallup, J.L., and J.D. Sachs (2001), The Economic Burden of Malaria, *American Journal of Tropical Medicine and Hygiene* 64 (suppl.): 85-96.

Glewwe, P., H. Jacoby, and E.M. King (2001), Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis, *Journal of Public Economics* 81 (3): 345 – 368.

Government of India (1994), Availability of Infrastructural Facilities in Rural Areas of India: An Analysis of Village Directory Data, *Census of India 1991,* New Delhi: Office of the Registrar General.

Government of India (1997), District Level Estimates for Fertility and Child Mortality and their Interrelations with Other Variables. *Census of India 1991, Occasional Paper 1 of 1997*, New Delhi: Office of the Registrar General.

Greene, W. H. (1993), The Econometric Approach to Efficiency Analysis., in H. O. Fried, C.A.K. Lovell and S. S. Schmidt (Eds.), *The Measurement of Productive Efficiency: Techniques and Applications*, London: Oxford Univ. Press, 68-117.

Griffiths, W.E. (1972), Estimation of Actual Response Coefficients in the Hildreth-Houck Random Coefficients Model, *Journal of the American Statistical Association* 67 (339): 633-35.

Gupta, I., and A. Mitra (2004), Economic Growth, Health and Poverty: An Exploratory Study for India, *Development Policy Review* 22 (2): 193 – 206.

Hausman, J.A. (1978), Specification Tests in Econometrics, *Econometrica* 46: 1251 – 71.

Hildreth, C. Houck, J.P. (1968), Some Estimators for a Linear Model with Random Coefficients, *Journal of the American Statistical Association* 63: 548-95.

Irudaya Rajan, S. and Mohanachandran, P. (1998), Infant and Child Mortality Estimates: Part 1, *Economic and Political Weekly,* May 9: 1120-1140.

Kalirajan, K.P. and R.T. Shand (1994a), *Economics in Disequilibrium: An Approach from the Frontier*, New Delhi: Macmillan India Ltd.

Kalirajan, K.P. and R.T. Shand (1994b), Frontier Production Function: The Stochastic Coefficients Approach, *Oxford Bulletin of Economics and Statistics* 56: 87-96.

Kalirajan, K.P., and M.B. Obwana (1994), A Measurement of Firm- and Input-Specific Technical and Allocative Efficiencies, *Applied Economics* 26 (4): 393 – 398.

Meeusen, W. and J.V.D. Broeck (1977), Efficiency Estimation from Cobb Douglas Production Function with Composed Error, *International Economic Review*, 18: 435-44.

Miguel, R., and M. Kremer (2004), Worms: Identifying Impacts of Education and Health in the Presense of Treatment Externalities, *Econometrica* 72 (1): 159 – 217.

Mitra, A., A. Varoudakis, M. Veganzones-Varoudakis (2002), Productivity and Technical Efficiency in Indian States' Manufacturing: The Role of Infrastructure, *Economic Development and Cultural Change* 50 (2): 395-426.

National Bureau of Soil Survey and Land Use Planning (NBSS and LUP) (1992), *Agro-Ecological Regions of India.,* 2nd ed. (NBSS Publication No. 24.) Nagpur: NBSS and LUP.

Nerlove, M. (1965), Estimation and Identification of Cobb-Douglas Production Functions, *Rand MCNally*, Chicago IL

Ruger, J.P, D.T. Jamison, and D.E. Bloom (2001), Health and the Economy, in M.H. Merson, Black, R.E., and Mills, A.J., eds. *International Public Health*. Gaithersburg, MD: Aspen, 617-666.

Schultz, T.P and A. Tanzel (1997), Wage and Labor Supply Effects of Illness in Cote D'Ivoire and Ghana: Instrumental Variable Estimates for Days Disabled, *Journal of Development Economics* 53: 251 – 286.

Schultz, T.P. (2002), Wage Gains Associated with Height from Health Human Capital, *American Economic Review* 92 (2): 349 – 353.

Schultz, T.P. (2003), Wage Rentals for Reproducible Human Capital: Evidence from Ghana and Ivory Coast, *Economics and Human Biology* 1 (3): 331 – 366.

Strauss, J., and D. Thomas (1998), Health, Nutrition, and Economic Development, *Journal of Economic Literature* 36 (2): 766 – 817.

Swamy, P.A.V.B. (1970), Efficient Inference in a Random Coefficient Regression Model, *Econometrica* 38 (2): 311-23.

Swamy, P.A.V.B. (1971), *Statistical Inference in Random Coefficient Regression Models*, New York: Springer-Verlag.