"Population Growth, Income Growth and Deforestation: Management of Village Common Land in India"

by

Andrew Foster, Mark Rosenzweig and Jere R. Behrman
Population Growth, Income Growth and Deforestation:

Management of Village Common Land in India

August 1997

Andrew D. Foster
Brown University
Mark R. Rosenzweig
University of Pennsylvania
Jere R. Behrman
University of Pennsylvania

The research reported in this paper was supported in part by grants NIH HD33563 and NIH HD30907. We are grateful to Nauman Illias for able research assistance.
I. Introduction

With increasing concern about the phenomena of global warming and declining bio diversity, there has been increased attention paid to the disappearance of the world’s forests. The use of forest land has been a particular focus of attention among those concerned about the environmental consequences of population growth. As Lee (1991) and others have shown, the existence of common-property resources such as forest land in the absence of appropriate incentives may be the primary source of population externalities that might justify natalist interventions. Economists and other social scientists have also debated, however, whether common resources have been sufficiently well managed by local institutions in traditional societies to effectively eliminate inefficiencies such as those due to externalities that are potentially associated with these resources (DasGupta, 1993; Jodha, 1985). Moreover, while it is widely recognized that improvements in agricultural productivity permit population growth without adverse Malthusian income consequences, the effects on forest resources of agricultural technical change remain an open question.

One important barrier to improved understanding of the technological, behavioral and organizational forces affecting global deforestation has been the absence of adequate data. Many empirical studies of land and forest use are based on case studies over short time periods that do not permit identification of the longer-term consequences of population change or technical progress and cannot easily be generalized. In a recent article, however, Cropper and Griffiths (1994) have assembled a data base consisting of a time-series of non-OECD countries over the period 1961-91 that serves to describe associations among forest land, per-capita GDP, and population density within major continents of the world. But the very useful descriptive findings from this study do not measure the underlying causes of income growth across the countries and thus do not identify the potentially differential effects on forest use of such income-augmenting forces as technical change, rural industrialization or serendipitous natural resource discoveries. Moreover, as population change itself also affects income growth (Evenson, 1993),
interpretation of population effects net of income effects requires care. And, the evidence that population
growth adversely affects forests does not imply that a local population externality is present due to an
inefficient local land-management regime.

In this paper, we utilize a newly-assembled data set that combines at the village level longitudinal
household survey data and satellite images of land use that cover a wide area of rural India over a twelve-
year period to partially address limitations to knowledge about population and income effects on land
management in general and the management of forest exploitation in particular. We use the data to obtain
estimates of population and income effects on deforestation, to identify the mechanisms by which these
factors affect land use, and to address the question of whether forest areas are first-best locally managed.
We show, based on the solution to a three-sector general-equilibrium representative agent model, that the
proposition that forest area is first-best efficiently managed, as would be the case given complete markets
for land and labor, implies an exclusion restriction, namely that population growth does not have any
effects on forest area or forest resource extraction net of its effects on equilibrium wages, land rents and
incomes.¹ A key feature of the theoretical model is thus the distinction between the effects of population
growth on forest area and density that operate through the influence of wages, land rents, and income and
those that operate net of these economic or "pecuniary" effects. Using a specific parameterization of the
model we also establish that the exclusion test has power against the alternative hypothesis that forest labor
is unmonitorable but forest land allocations are second-best efficient given this limitation.

India is a particularly interesting and useful setting in which to study the determinants of
deforestation. India is a country that has experienced both rapid population growth and relatively high rates
of deforestation in the last 30 years. In addition, since the late 1960's, India has been a major beneficiary of
the "green revolution," which has substantially augmented crop productivity growth. There are two

¹The general-equilibrium exclusion restriction test is analogous to tests of separability using farm
household models, as in Benjamin (1992).
important characteristics of the Indian green revolution relevant to the study of income growth effects on
the environment and its relationship to population change. First, the new technologies, in the form of new-
high yielding seeds (principally, for wheat, rice, corn and sorghum), were developed outside of India and
imported. Second, because the new seeds were particularly sensitive to soil properties and water
availability, due to the wide spatial variation in agroclimatic conditions across India, realized productivity
gains in agriculture differed widely across India. Thus, for at least the initial stages of the green revolution,
technical change was not itself induced by area differentials in population growth or density and was
sufficiently varied to potentially identify its consequences. Evenson (1993) has exploited this variation to
examine the effects of population growth rates and agricultural technical change on rural wage rates and
land prices across districts of India.2

Large-scale studies of the effects of agricultural productivity growth or population change on
deforestation in India have been limited by the fact that governmental data sources provide information only
on designated forest lands, including newly-designated land for forests on which tree planting may have
only been recently initiated rather than on actual tree cover or forest biomass.3 A recent study (Richards
and Flint, 1994) has provided estimates, in part based on satellite imagery, of forest biomass at the state
level for India for 1970 and 1980. These estimates show the existence of wide spatial variation across India
in forest devastation and permit a look at its association with population trends and agricultural
productivity growth.4 Figure 1 provides, by Indian state, the proportionate changes in forest biomass from
the Richards and Flint data, population size, from the Indian censuses of 1971 and 1980, and a Laspeyres-

------
2Foster and Rosenzweig (1996) exploited these features of the green revolution to quantify the
relationship between technical change and returns to schooling in India.

3Cropper and Griffiths (1994) were unable to estimate any significant relationships for Asia or
South Asia using official government statistics on forests.

4Their estimates of “forests” also are quite different from those available from government
documents. See section III below.
India State-Specific Rates of Forestation, Output Growth and Population Growth: 1971-82
weighted crop productivity index for wheat, rice, corn and sorghum, from Vanneman and Barnes (1995) over the period 1970-1980, where the states are arrayed by the amount of deforestation. As can be seen, all but one Indian state (Jammu and Kashmir) experienced a loss in forests in this first decade of the green revolution, with losses ranging from 2% in Bihar state to 55% in Haryana. The figure also suggests there is a slight positive correlation between agricultural productivity growth and deforestation (.15), but no evident correlation between population change and the loss in forest area across states (.03).

The Richards and Flint forest biomass estimates provide too few observations to permit econometric analyses of the effects of population and income change on deforestation in India. Moreover, given the local nature of the management of land resources that is emphasized in the literature, state- or district-level analyses of relationships between population densities, income growth and forest use may obscure key relationships to the extent that they are mediated at the local level. For India, the most important governmental unit managing common property resources is the village. We have therefore constructed a village-level data set covering the period of the 1970's that makes use of direct satellite image data to construct estimates of forest biomass change and national rural household survey data describing agricultural productivity, wage rates, land rents and incomes.

Section II of the paper sets out the general equilibrium, three-sector model of land and labor allocations that contrasts the perfect-markets and unmonitorable labor regimes and establishes the exclusion test. Section III describes the data sources and the construction of the variables. Section IV discusses econometric specification and estimation issues, and section V presents the reduced-form estimates and the results of the structural exclusion tests. The reduced-form econometric results confirm earlier findings that higher population densities lead to reductions in forest area and density, but indicate that the effect of income growth on forests is importantly conditioned by the specific mechanisms causing incomes to increase. In particular, we find that both agricultural technical change and rural industrial growth increase equilibrium land rents, wages, and incomes but, due to differentials in the relative
magnitudes of these effects, the former results in decreased forest area while the latter has the opposite effect. These results suggest that agricultural technical change, while forestalling the negative income consequences of a Malthusian equilibrium, exacerbate the negative consequences of population growth on forest exploitation. Our estimates also reject the restriction implied by the first-best model of land allocation for a variety of estimation procedures and specifications, and appear to conform reasonably well to the second-best model in which common forest land is optimally managed subject to limitations on controlling labor allocations, so that forests are overexploited. However, this second-best model is not easily distinguished empirically from a broad class of alternative models including ones in which forest land allocations are not even second-best efficient.

II. Theory

To elucidate the paths by which agricultural technical change and population growth affect deforestation and to establish a test for whether the amount of forest area and deforestation reflect suboptimal forest management, we formulate a simple model consisting of a set of regional economies (villages) in which households attempt to optimally allocate labor across forest, agricultural and manufacturing activities. The three-sector model is specified under two regimes: a benchmark model in which all markets for production inputs are perfect and a model in which labor time spent in the extraction of forest resources cannot be monitored. For clarity of exposition we suppress regional subscripts and assume that in each village (1) the number and size of households and thus population growth is exogenous and (2) households are identical. We also assume that periods are of sufficient length such that the extraction of forest resources in one period does not influence the output of forest products in

---

5López (1997) also rejects first-best efficient forest management, based on data from 20 villages in Côte d'Ivoire.

6We allow for the possibility that population growth responds to environmental shocks when we estimate the model.
subsequent periods,

Total available land per household $a_t$ is divided at time $t$ into forest land, $a_{fr}$, and agricultural land $a_{ag}$. Per-household total labor supply, i.e., family size, is denoted $l_t$, and labor supply by the household in the forest, agriculture and manufacturing sectors are denoted $l_{fr}$, $l_{ag}$, and $l_{mf}$, respectively. The technologies associated with the extraction of forest products - fuelwood and fodder\(^7\) - and the agricultural and manufacturing sectors are, initially, $f(v_t,a_{fr},l_{fr})$, $g(\theta_t,a_{ag},l_{ag})$, and $m(\eta_t,l_{mf})$, respectively, where $v_t$ reflects the endowment of forest products in the village determined by the local environment (rainfall, temperature, soil quality, etc.), $\theta_t$ indicates technologically-determined agricultural productivity (the suitability and extent of adoption of high-yielding variety crops), and $\eta_t$ captures labor productivity in manufacturing (the availability of infrastructure, capital, and knowledge relevant to the generation of non-agricultural non-forest employment). We assume that manufactured and agricultural goods are traded across villages and labor is sectorally mobile. Neither labor nor forest products are traded across regions (villages), although they may be traded by households within villages. The intervillage nontradeability of forest goods is motivated by the observation that goods such as firewood or the forage consumed by livestock in the forest, the principal uses of forest products in the Indian context that we study, are bulky and thus are not easily transported.

a. **Complete Input and Output Markets and Private Land Ownership: Efficient Land Management.**

We first consider the problem of a representative household in an environment of complete and perfect input and output markets to illustrate how shifts in population size and technical change influence forest use when land use is allocated efficiently. This “complete-markets” model thus corresponds to one as well in which forest-land commons are first-best efficiently managed by a village council or other institution. We assume that a representative household maximizes expected discounted utility and in each

---

\(^7\)Bowonder (1982) estimates that 90% of the wood extracted in India is used for fuel.
period chooses allocations of its land to forest and agricultural production, allocations of labor to forest-product extraction, agriculture, manufacturing, and the labor market, and its consumption of forest and non-forest goods. Formally, the household maximizes in each period $\tau$,

$$V_\tau = \mathbb{E}_\tau \sum_{t=\tau}^{\infty} \beta^t u(c_{\beta t}, c_{n t})$$  \hspace{1cm} (1)$$

subject to a single-period budget constraint in each period $t$

$$p_{ij} c_{j t} + c_{nt} = p_j f(v, a_{jt}, l_{jt}) + g(\theta, a_{jt}, l_{jt}) + m(\eta, l_{nt}) + w_i x(l_{jt} - l_{jt} - l_{gt} - l_{nt}) + r_i x(a_{jt} - a_{jt} - a_{gt}) + \omega_i$$  \hspace{1cm} (2)$$

where $\beta$ is the discount factor, $u(\ldots)$ is the single-period utility function; $c_{nt}$ and $c_{nt}$ are the consumption of forest goods and non-forest goods; $p_{jt}$, $w_i$, and $r_i$ are the local equilibrium prices of forest goods, labor and land; $a_i$ and $l_i$ are the land and labor endowments; and $\omega_i$ is income from sources other than land and labor ("endowed income"). Uncertainty is assumed to arise from stochastic shocks to the technology and endowment variables. Note that we have not allowed households to smooth consumption. We show below that the exclusion test that we derive is robust to this assumption.

This dynamic stochastic model can be separated into a series of distinct single-period optimization problems. Moreover, each single-period problem can be separated into a consumption allocation decision given income and an income-maximization problem. The income-maximization problem involves (i) the determination of optimal allocations of land to forest and labor to the three sectors, given $w_i$, $p_{jt}$, and $r_i$, the technologies of agricultural production $\theta$, and manufacturing $\eta_i$, and the endowment $\omega_i$. The equilibrium condition for labor is just that the sum of per-household labor allocations by sector, given the wage, rental rate and the price of forest goods, must equal the per-household labor endowment $l_i$. In order to construct an equilibrium condition for forest goods we may solve the single-period consumption allocation problem of the form.
\begin{equation}
\nu(y_1, p_R) = \maximize_{c^R_n, c^m_n} \ u(c^R_n, c^m_n) \text{ such that } p_R c^R_n + c^m_n = y_1
\end{equation}

which yields demand equations for forest, \(c_i(y_1, p_R)\), and non-forest, \(c_n(y_1, p_R)\), goods. Thus the village equilibrium in forest products, given the assumed non-tradeability across villages in forest products, requires that \(c_f(y_1, p_R) = f(a_f, l_f)\).

Solving the first-order conditions determining the allocation of labor and land along with the equilibrium conditions for labor markets and forest products yields reduced-form expressions for how forest area as well as wages, rent on land and income are influenced by technology, infrastructure, endowment income, local environment and population density as determined by household size and land per household (or, its inverse, household density, the number of households per unit area):

\begin{equation}
z_i = z(\theta, \gamma, \omega, v, a_i)
\end{equation}

for \(z = \{a, w, r, y\}\). It is easy to demonstrate that the model implies that in (4), improvements in (factor-neutral) agricultural technology \(\theta\), raise land rents and wages, manufacturing productivity raises wages and may raise rents by increasing demand for forest products, and increases in household size and density (the number of households per unit area) raise land rents but lower wages. It may also be shown that these latter effects can differ even though household size and household density both influence land and labor rents through their effects on population density; this is a consequence of the assumption that the household, rather than the individual, is the relevant decision-making unit. Only when per-capita income and per-capita forest goods consumption are not affected by household size will the elasticities of
household size and number be identical. In the case of the determination of household income $y$, different effects of household size and density are especially likely to be observed because increases in household size directly increase household income $y$, although less than proportionally by augmenting household labor supply.

These implications from the reduced-forms of the complete-markets model, however, may not be different from those derived from a variety of models that incorporate market failures (see below). A strong prediction from the complete-market framework, however, can be constructed by deriving equilibrium conditions that condition on three endogenous variables - income, wages and land rentals. The equations may be obtained by inverting equation (4) for $z=\{w,r,y\}$ with respect to the technology level $\theta$, manufacturing infrastructure $\eta$, and endowment income $\omega$, and substituting into (4) for $z=\{a_t\}$. This yields equations of the form:

$$a_t = a_t(w, r, y, \theta, \eta, \omega)$$

The fact that equation (5) can be informative about deviations from the assumption of complete markets may be seen by noting that, under the complete-markets, private-forests model, $\partial a_t/\partial l = \partial a_t/\partial = 0$ - changes in household size and density have no direct effect on forested area net of the land price, wage

---

8Sufficient conditions for this are that $c(y, p)$ is linear in income, that the income endowment is zero, and that there are zero non-labor manufacturing earnings. Alternatively, the model could be modified so that the income endowment and any non-labor manufacturing earnings accrued to individuals rather than the household so that, for example, total household endowment income would be $\omega l$.

9This result raises the question of whether it would have been more attractive to carry out the analysis on a per-capita rather than a per-household basis. While this would in principle allow us to ignore the distinction between household size and density the simplification would have come at a substantial cost. A household-based model conforms more reasonably to the setting and to the data in that the household is normally considered to be the locus of decision making in rural South Asia and that, like most survey data from developing countries, the survey used in this study primarily presents household-level information. In order to carry out the analysis at an individual level it would also have been necessary to impose substantial additional structure on the model such as imposing the assumption that forest product demand is linear in income and that household members are identical.
rates and income when there are markets for land and labor that operate perfectly. To see this, we can rewrite the income-maximizing problem solved by a representative household facing competitive markets for land and labor at rates $r_t$ and $w_t$ as

$$
\begin{align*}
\max_{a_h, l, l_m} & \quad y_t = a_h l_t l_m + p_n f(v_t, c_{pt} l_t) + g(t, c_{gt} l_m) + m(l_m) + \omega_t \\
& + w_t (l_t - l_{pt} l_m) - r_t (a_t - a_{pt} - a_{gt})
\end{align*}
$$

(6)

The first-order condition with respect to labor devoted to forest goods implies that the marginal revenue product of forest labor is set equal to the wage, $p_n f(v_t, a_{pt} l_m) = w_t$. Similarly, forest land will be allocated so that $p_n f(v_t, a_{gt} l_h) = r_t$. These two equations, coupled with the forest-products equilibrium equation $c_f(y_t, p) = f(a_{pt} l_h)$, yield a system of three equations in three unknowns, $a_h$, $l_h$, and $p_h$. Solving, in particular, for forest area yields

$$
a_p = a_p(w_t, r_t, y_t, v_t)
$$

(7)

which is a special case of equation (5), in which the population measures $l_t$ (household size) and $a_t$ (land per household) are excluded. Thus, in the case of (locally) complete land and labor markets, changes in the size of the population should have no effect on forest area net of wages, land rentals, and household income. Estimation of (5) and the test that the exclusion restrictions embodied in (7) are appropriate provides a test for efficient management of forest resources.

---

10Note that the exclusion of household size depends on the assumption that, given income and forest product prices, household size does not affect household demand for forest products. If one takes the position that the household utility function reflects the underlying preferences of individual household members, then this exclusion is easily justified if both forest and other goods are private goods within the household. If forest goods are public within the household (e.g., due to the public-good nature of heating and cooking), then household size may enter the forest-products demand function since the return to public good consumption is increasing in household size (see Foster and Rosenzweig 1997). However, under such circumstances an increase in household size would increase forest area net of wages, rentals and income, a prediction that is rejected by the data below.
Although we have initially assumed away the possibility of intertemporal resource transfers, the test can be generalized to account for saving and/or borrowing opportunities by replacing income in equation (7) with household expenditures. This may be established by first noting that the first-order conditions determining forest area and labor under complete markets, which are used to derive (7), are not influenced by the presence or absence of intertemporal markets. What is affected is the condition equilibrating the market for forest products, because forest product demand will be influenced in general by assets and the return on savings, not just current period income. However, note that, regardless of credit-market structure, the first order conditions determining period \( t \) consumption are \( \partial u / \partial c_{it} = \lambda_t p_{it} \) and \\
\( \partial u / \partial g_t = \lambda, \) where \( \lambda_t \) is the period \( t \) marginal utility of income. Combining these with the definition of total expenditures \( e_t = p_{it} f_{it} + c_{gt}, \) yields an expenditure-conditional demand function \( c_{it}^x (e_t, p_{it}), \) which must be equated to forest output in equilibrium. Solving the resulting equation along with the forest allocation equations results in an alternative version of (7), \( \alpha_t = \alpha_t^x (w, r, e, v_t), \) which excludes, as before, household size and density.

The special structure of the complete-markets case that permits derivation of (7) also allows the development of formal expressions characterizing the effects of wage \( w \) and rental rate \( r \) on forest area. In particular, implicitly differentiating the three equation system yields:

\[
\frac{\partial a_f}{\partial r} = \frac{1}{D} [f_{f} f_{f}^2 + p_{f} f_{f} \frac{\partial c_{f}}{\partial p_{f}}] = -\frac{f_{f} f_{f} \epsilon_{f} + f_{f} \epsilon_{f} \epsilon_{c} e_{c}}{f_{f}^{2}}
\]

(8)

and

\[
\frac{\partial a_f}{\partial w} = -\frac{1}{D} [f_{f} f_{a} + p_{f} f_{a} \frac{\partial c_{f}}{\partial p_{f}}] = -\frac{f_{f} f_{a} \epsilon_{f} + f_{f} \epsilon_{f} \epsilon_{c} e_{c}}{f_{f}^{2}}
\]

(9)

where \( f_i \) is the marginal product of forest labor and so forth, \( \epsilon_{it} \) is the elasticity of forest products with
respect to forest labor, $\varepsilon_{ml}$, the elasticity of the marginal product of forest land with respect to forest labor, $\varepsilon_{cp}$ is the price elasticity of demand for forest goods,\(^{11}\) and

\[
D = p\frac{\partial f_{fa}}{\partial f_a} f_a^2 - f_a p\frac{\partial c_f}{\partial p} + f_a p\frac{\partial c_f}{\partial p} - p f_a \frac{\partial c_f}{\partial p} \tag{10}
\]

with $D<0$ for a stable equilibrium. It is clear from (8) that the effect of the land rental rate on forest area is unambiguously negative under the plausible conditions that $f_{la}>0$ and $\varepsilon_{cp}<0$. On the other hand, (9) indicates that the sign of the wage-rate effect on forest area is ambiguous, depending on the relative magnitudes of the price elasticity of demand for forest goods and the structure of the forest production function.\(^{12}\)

b. Unmonitorable Forest Labor

One of the most important problems of forest management identified in the literature is the impossibility of monitoring labor use on open-access land such as forests, grazing land or fisheries. We can modify the complete-market model to incorporate this constraint, as in Dasgupta (1995), for the purpose of establishing that the exclusion restriction associated with the complete input and output markets model has power against at least one coherent model incorporating inefficiency in forest management. We assume that individuals who extract resources on unmonitored forest land in the village receive output that is proportional to their share of total labor on the forest land. Then assuming that each household exhibits Nash behavior, if $a_n$ and $l_n$ are the average amounts of forest area and time spent in forest production per household in the village so that the average product is $f(a_n^*, l_n^*)/l_n^*$, for a sufficiently large number of households (such that the labor contributions of a particular household to the community average may be

\(^{11}\)These expressions hold whether or not intertemporal markets are present if the price effects and price elasticities pertain to demand equations conditioning on expenditure.

\(^{12}\)If the forest production function is Cobb-Douglas it is easy to show that the sign of the wage effect depends on the sign of $1 + \varepsilon_{cp}$. Thus forest area is decreasing in the wage if and only if the price elasticity of demand for forest goods exceeds one in absolute value.
ignored), we may write the household's income-maximizing problem, for a fixed allocation of forest land per household $a_{it}$, as

$$ y_i(a_{it}) = \max_{l_{it}, r_{it}, l_{mt}} \frac{p_{it} f(v_{it} a_{it}, l_{it})}{l_{it}} - g(\theta_i a_{it}, l_{it}) + m(\eta_i r_{it}, l_{mt}) + \omega_i \rho_{it} \rho_{it} - l_{it} - l_{mt} - l_{gt} + r_{it} (a_{it} - a_{gt}) $$  

(11)

Differentiating (11) yields a first-order condition for forest labor of the form $p_{it} f(a_{it}^*, l_{it}^*)/l_{it}^* = w_i$, which indicates that for an interior solution the average product of labor must equal the market wage. If all households behave identically, so that we have a symmetric Nash equilibrium, then each household is at the interior solution $l_{it} = l_{it}^*$. Because for any given allocation of land the average product of forest labor exceeds the marginal product of forest labor, we obtain the standard result that the marginal product of forest labor is less than that of agricultural labor (which is set equal to the wage) and thus that labor is allocated inefficiently, with too much labor being dedicated to the extraction of forest products.

In a setting in which forest extraction cannot be directly monitored, it is obvious that no household would allocate any land to forests. A second-best solution is for a social planner to maximize the utility of the representative household by choosing the allocation of land area directly. We now characterize land and labor use in the case in which forest exploitation cannot be directly regulated by any household or by a social planner, derived from the second-best efficient social planner's optimum. The social planner maximizes (3) subject to the private allocation of labor by village residents conditional on the allocation of forest area, given by the solution to (11), and to equilibrium conditions for the wage, the rental price on agricultural land, and the price of forest goods. The first-order condition for the social planner's optimal forest area is:

---

13 In order to derive solutions we continue the assumption of absent intertemporal markets.
\[
\frac{\partial v_t}{\partial a_{ft}} = \frac{\partial v_t}{\partial y_{ft}} \left( f_t(v_t, a_{ft}, l_{ft}) + g_t(v_t, a_{ft}, l_{ft}) \frac{\partial l_{ft}}{\partial a_{ft}} \right) - \frac{\partial p_{ft}}{\partial a_{ft}} = 0
\]

Because the average product of labor exceeds the marginal product of labor and because labor allocations to forests are increasing in average forest land, equation (12) implies that at the social-planner's optimal allocation of forest land, the marginal revenue product of forest land exceeds the rental rate on agricultural land in contrast to the perfect markets case in which the marginal revenue product of land is equated across sectors. Thus, given the prevailing wage and land rental rate and the marginal products of land and labor in forests, there is too little forest land and too much extraction of forest resources per unit land compared to the competitive case, even when land use is managed (second-best) optimally. The intuition is straightforward: because forest land is used inefficiently (i.e., it attracts too much labor) it is optimal to reduce the amount of forest land relative to what would be dictated by equating the marginal revenue products of land in the two sectors. Because of the inability for the social planner to determine labor allocations directly the resulting equilibrium is inefficient relative to the competitive case.

c. Parametric comparisons of the two regimes.

An important implication of the fact that the marginal product is not equated to the land rental rate in the social-planner's optimum, given unmonitorable forest labor, is that, in contrast to the complete markets regime, the system of equations determining equilibrium in the forest-products sector cannot be decentralized, so that forest area cannot be solved as part of a three-equation system conditional on wages, rentals, and household income. While this implies that the exclusion restriction implied by (7) is unlikely to

---

14Note that for the same set of parameters and endowments this need not imply less forest area under unmonitorable forest labor than under competitive markets case because the equilibrium wage and rental rates will, in general, differ. Below we provide an example in which equilibrium forest area is higher under unmonitorable forest labor.
hold given this alternative model, it also makes the analytic derivation of comparative static results prohibitively complex. Thus, to establish that the test of the exclusion restriction associated with the complete-markets regime has power against the alternative of unmonitorable forest labor, we specify a parametric model for which interpretable analytic derivatives may be constructed for the key results of the model, those relating to the effects of household size and density on forest area given household income, wages, and rents.

For analytic convenience we assume that the utility and production functions are Cobb-Douglas and that technological parameters appear as multipliers on their respective production functions, so that \( f(a, l) = a_f l^{1-\beta} \), \( g(a, l) = a_g l^{1-\gamma} \), \( m(l) = \eta l^\delta \), and \( u(c, c') = c^\gamma c'^{1-\alpha} \). Technical change is thus factor-neutral. Given this parameterization, the solution to the forest-area equilibrium condition (5) for the complete markets case is:

\[
a_f = \alpha \beta y/r
\]

(13)

Expression (13) demonstrates that, given complete markets, greater preferences for forest goods, a higher output elasticity for forest land, and higher incomes lead to greater amounts of forested area while higher land rents reduce the amount of forested land. However, as demonstrated for the general case (7) neither household size \( l \) nor household density \( l/a \) influence forest area net of income and prices (in this case, the land price). By contrast, the comparable equation under the regime of unmonitorable forest labor, the derivation of which is provided in Appendix I, for \( \beta = .5, \gamma = .5, \) and \( \delta = .5, \) is

\[
a_f = (\frac{2 - \gamma}{\alpha} - \frac{
}{2wl} - 1) \gamma
\]

(14)

in which the equilibrium amount of forest area depends on preferences, income, the land price, wages and household size. In this Cobb-Douglas case, it is also true that under the two regimes variation in the
number of households per land area does not affect the amount of forested land. This latter result, while it does not necessarily generalize, emphasizes the importance of distinguishing between the effects of household size and density when carrying out tests of the efficiency of forest-labor markets. The appearance of the wage in (14) but not in (13) suggests that a non-zero wage effect on forest area may be observed even under specifications for which a zero effect would be observed given complete markets. However, unlike for the household size variable, this result does not provide a strong test because the absence of a wage effect observed under complete markets results from the specialized assumption of Cobb-Douglas preferences and technologies.

Differentiation of (14) yields the effect of household size on forest area (net of the effects of changing wage rates, land prices and income) when forest labor allocations are not monitorable for the Cobb-Douglas case:

$$\frac{\partial a_t}{\partial l} = \frac{y^2}{2} \left( \frac{2 - \frac{y}{\alpha} - \frac{y}{2wl} - 1}{\alpha} \right)^2 w^2 l^3$$

(15)

Expression (15) is unambiguously negative as long as equilibrium income is positive and there is an interior solution for forest area. Given the null hypothesis of complete markets, the proposed test thus appears to have substantial power against the alternative hypothesis of unmonitorable labor.

While closed-form expressions for the reduced-form relationships in (4) can be derived for each regime, they are, in general, complex and unilluminating, particularly for the unmonitorable labor regime. To get some sense of how the equilibrium values for forest area, incomes, land prices and wages differ across the perfect-market and unmonitorable labor regimes and also to assess whether the reduced-form relationships under the two regimes are distinctive, we can substitute values for each of the parameters.

\[15\] The programs carrying out the calculations and presenting the results are available from the authors on request.
Based on these assumed values, we can then solve for equilibrium values and obtain numerical partial derivatives corresponding to the reduced-form effects of changing agricultural technologies, population growth, industrial labor demand and endowment incomes on equilibrium prices, incomes and per-household forest area.

Table 1 reports for each market regime in the first row the equilibrium values and in the rest of the table the reduced-form (4) effects. These were obtained by setting each of the technology parameters, θ, η, and ν equal to one, setting the income endowment ω at .5, setting α, β, γ, and δ to .5, and assigning to each household three members (l=3) and two units of land endowment (a=2). Interestingly, the parametric results indicate that in this parameterization the equilibrium amount of forest area is actually higher under the second-best regime in which forests are optimally managed but forest labor is not monitorable compared with the complete markets regime. Thus going to a first-best solution does not necessarily increase forest area. However, the price of the forest good, income and the wage are higher and land rent lower when markets are complete compared with the case in which forest labor cannot be monitored. The lower price of the forest good when forest labor cannot be monitored results from the fact that too much labor is allocated to forest-product extraction, even though the land allocation is optimally set, resulting in higher supply of the forest good.

An important general feature of Table 1 is that despite the cross-regime differences in equilibrium values, the reduced-form effects of agricultural technical change and population growth on wages and land prices for the unmonitorable labor and complete-markets regimes are of the same sign and roughly the same magnitude. Evidently, for at least one parametrization, neither the amount of forest area nor information on the price effects of population and technical change provides a clear signal of market failures associated with land management. Indeed, under either regime the signs of the reduced-form effects are those that conform to the perfect-markets model: increases in (factor-neutral) agricultural technology (θ) raise land rents, wages, and income; an increase in the demand for non-agricultural labor (changes in η) raises wages
<table>
<thead>
<tr>
<th>Equilibrium values</th>
<th>Complete Markets</th>
<th>Unmonitorable Forest Labor</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta )</td>
<td>[.667] [.333]</td>
<td>[.125] [.745]</td>
</tr>
<tr>
<td>( \eta )</td>
<td>[-.167] [.433]</td>
<td>[-.292] [.119]</td>
</tr>
<tr>
<td>( l )</td>
<td>[.083] [.5]</td>
<td>[.441] [.441]</td>
</tr>
<tr>
<td>( v/a )</td>
<td>[.333] [.333]</td>
<td>[.373] [.373]</td>
</tr>
<tr>
<td>( \omega )</td>
<td>[0] [0]</td>
<td>[0] [0]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Partial derivatives</th>
<th>Complete Markets</th>
<th>Unmonitorable Forest Labor</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r )</td>
<td>2.5</td>
<td>2.42</td>
</tr>
<tr>
<td>( w )</td>
<td>.982</td>
<td>[.782] .319</td>
</tr>
<tr>
<td>( y )</td>
<td>2.15</td>
<td>1.75</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>1.38</td>
<td>1.19</td>
</tr>
</tbody>
</table>

*Elasticities reported in square brackets.*
and incomes, but lowers land rents, and population growth, whether via an increase in household size or in the number of households, increases land prices but decrease wages. Note that an increase in household size also increases household income, although per-capita income is clearly decreasing in household size since the elasticity is less than one. Increasing the number of households per unit area (1/a) decreases household incomes.

The simulation results also indicate that increases in population, from either increases in the size or number of households, lead to a decrease in forest area, with the quantitative effects almost identical across the two regimes. The results in Table 1 also illustrate the point that the relationship between income change and deforestation, while not dependent on regime, depends importantly on the source of the growth in incomes. Income growth in the model arises from either improvements in agricultural or industrial technologies (θ or η) or from increases in the income endowment ω. However, agricultural technological progress lowers the amount of forested area, somewhat more so when forest labor is unmonitorable, while increases in industrial productivity or in the household income endowment increase the amount of forest area in equilibrium. Given variation across areas in the sources of income change, relationships between income growth and forest devastation can vary widely across different economies even if the underlying structure is similar.

III. The Village Panel Data Set

As noted, India represents an interesting and potentially useful setting in which to examine the determinants of deforestation. The spatially-differentiated growth in agricultural productivity resulting from the exogenous importation of new seed technologies applied to differentially-suitable agroclimates combined with relatively high rates of population growth and deforestation experienced by India in the 1970's would appear to have the potential to provide insights into the roles of agricultural technical change and population growth in affecting the levels and changes in forest area. Given the importance of local land use management for the maintenance of forests and heterogeneity in agroclimatic conditions that affect both
land productivity and forest health, an assessment of the impact of technical change and population growth on deforestation would ideally make use of changes in these variables over time across administrative units which correspond to local authority. In India the relevant local authority is the village.

We construct a panel data set at the village level using newly-available repeated cross-section and panel Indian household survey data covering the period 1970-1982 and satellite data describing land use that conforms to the requirements of the model. In particular, we merge survey-based information on household incomes, household consumption, household size, numbers of households, crop productivity, land prices, wage rates, and industry presence with governmental statistics on weather and satellite-based information on locale-specific changes in the density of forests. The constructed data set comes from four sources: (i) the 1970-71 National Council of Applied Economic Research (NCAER) Additional Rural Incomes Survey (ARIS), (ii) the Resources for the Future Supplement to the ARIS (RFF), (iii) the 1981-82 NCAER Rural Economic Development Survey (REDS), and (iv) Landsat satellite spectral images for India from 1972-1980.

a. Household survey data

The ARIS data set consists of information collected in a survey of 5,115 rural households in India in the crop year 1968-69. These households were reinterviewed in 1969-70 and again in 1970-71. This national survey was based on a stratified design in which households were randomly sampled from three strata: (1) 1,107 households in a random sample of villages in each of the 16 districts participating in the Indian Intensive Agricultural District Program (IADP). This program was set in place in 1961 in one district in every Indian state (except two districts in Kerala), to encourage farmers to adopt modern production practices, provide advice on efficient use of new technologies, and help to assure adequate supplies of production inputs such as fertilizer and seeds. (2) 2,040 households from a random sample of villages in the 91 districts (in 11 states) participating in another agricultural program, the Intensive Agricultural Area Program (IAAP), a program initiated in 1965-66 with similar goals to those of the IADP
but with considerably less resources. (3) 1,968 households from a random sample of villages in the rest of India not covered by either program. There are 250 villages represented in the ARIS survey, with approximately 20-30 randomly-selected households in each village. The villages are identified by name and by location in administrative units, permitting the geocoding of the data (see below).

The households interviewed in the third round of the survey that we use to construct village-level aggregates represent 91.1 percent (4,659 households) of the original sample, with response rates almost equal for all three strata. The 1970-71 ARIS data set contains sample weights reflecting the stratified sample design and the non-response rates so that population statistics, necessary for aggregation and merging, can be obtained from the survey data. These data contain information on household structure (age-sex composition), income by source, wealth, and agricultural inputs, outputs and costs, by item. Also provided is information on village-level land prices, for irrigated and unirrigated land, and on village infrastructure, including the presence of schools and medical centers and distances to urban areas, regional markets, and transportation links. The data also provides information on whether there is a factory present in the village. The RFF data supplement consists of additional information about the individuals, fertility and farming practices in 4,118 households in 1970-71 that was not originally coded by NCAER. In particular, there is information on the wage rates, by activity, for all individuals who work in the sample households enabling the construction of adult, agricultural wage rates by village using household sample weights.

NCAER conducted a resurvey (REDS) of the 1970-71 survey households in crop year 1981-82. The REDS data provide information on a subset of the original 1970-71 households as well as data on a new, random sample of households based on the same survey design as in the ARIS and on a complete census of households in the original 250 ARIS villages. Because of political constraints, all households in the state of Assam were dropped from the sampling frame. The panel and the new households together number 4,947 and, based on the sample weights, are representative of the entire national rural Indian
population (except for Assam) in 1981-82. The non-response rate of the new sample households was 6.0%. The REDS data provides sampling weights for all households reflecting both the non-response and the survey design, as described in more detail in Vashishtha (1989), thus permitting construction of a representative data set at the village level for 1981-82 that can be matched with that from 1970-71.

The REDS data contain information similar to that coded in the combined 1970-71 ARIS and RFF files in addition to providing a detailed village inventory inclusive of village wage rates and the presence of rural industry. The ARIS, REDS (representative component) and RFF survey and village data combined at the village level thus provide a consistent set of rural agricultural wage rates, land prices, crop productivity, and rural industry measures and information on the size and numbers of households for 240 villages spread all over India for the crop years 1970-71 and 1981-82.

The ARIS-REDS surveys provide sufficiently detailed data on agricultural profits and quasi-fixed inputs for the individual panel households to make it possible to obtain estimates of area-specific rates of agricultural technical progress - farm profit growth net of the influence of investments in human and physical capital - over the 1971-1982 period, as in Foster and Rosenzweig (1996). In that study, rates of technical progress were obtained at the district level. For this study, we employ the same non-linear profit-function methodology and data, but obtain technical change measures at the village level. As discussed in the next section, we use these estimated measures as instruments to predict changes in village-level price-weighted measures of output per acre of four crops whose productivity was enhanced substantially by the introduction of new high-yielding seed varieties - corn, rice, sorghum and wheat (HYV crops). These data are also supplemented with district-level time-series information on annual rainfall for the relevant districts in which the ARIS-REDS villages are located.

b. Landsat satellite data

Official Indian data sources for the period spanning the two NCAER surveys provide information only on land officially classified as forest. This includes land newly set aside for growing trees but which
may not yet exhibit any vegetative growth as well as designated forest areas that have been encroached upon by local growers. Indeed, in contrast to the Richards and Flint (1994) state-level estimates of forest cover, which suggest that the amount of forest biomass - trees - declined in India from 1970-1980, total forest area reserved by Indian governmental authorities for trees actually exhibits an increase over the same period, and the cross-state correlation in the growth rates of official forest area and estimated forest (tree) biomass is negative. The most dramatic example is the state of Haryana, one of the states with the highest rates of increase in agricultural productivity over the period, as noted. In this state official forest area increased by over 60%, but actual tree biomass per hectare according to the Richards and Flint data decreased by 58%.

The absence of ground-level censuses of trees for the relevant period covered by the survey means that in order to obtain a measure of the changes in actual forest or tree cover for the specific “micro” regions surrounding each of the survey villages it is necessary to employ satellite images. Satellite images based on specific light-frequencies enable the construction of indices that measure reasonably accurately area vegetation for relatively small geographic areas. The index we use is the normalized differentiated vegetation index (NDVI) (Rouse et al., 1974), which is the ratio of the difference in reflectance in the near infra-red and red bands in the light spectrum to the sum of these reflectances. This index correlates well with the presence of plant matter because vegetation tends to reflect infra-red light and absorb red light. It is among the most commonly used measures of vegetative cover because it is simple to compute and filters out topographic effects, variations in the illumination angle of the sun, and other atmospheric elements such as haze. The NDVI is bounded between -1 and 1, with vegetation associated with trees achieving values of .2 or greater.

To match the satellite and survey data we first had to geocode the survey data. To do this we

---

16Roughly speaking, the high infra-red reflectance accounts for the relative coolness of vegetation and the low red reflectance accounts for the green color.
obtained geographic position codes for each of the 250 villages in the ARIS survey based on maps from the
district-level volumes of the 1971 and 1981 Indian censuses. To ensure that survey village names
corresponded to those in the census, we used the survey information on village, tehsil, and district names.
Seven villages had to be dropped from the original 250 because a village of the name specified in the ARIS
sample was not found (4 of 7) or because more than one village of the same name was found in the
corresponding tehsil and district (3 of 7). The tehsil maps, which plot the locations of each village, were
then geo-registered using the district-level maps, which contain latitude and longitude information.

The geo-registered tehsil maps, which provide the exact location in degrees and minutes of the
sample villages, were used to select the relevant path and row data for Landsat satellites 1-3, which provide
Multispectral Scanner (MSS) data for the period 1972-1983 and yield a spatial resolution of approximately
80 meters in four bands of the spectrum. Each "path" and "row" pair uniquely determine a geographical
area of approximately 185 kms square for the MSS images. The 243 villages for which location
information was available were contained within 73 path-row locations, indicating that on average each
selected area or "scene" contained 3.3 study villages. Satellite scenes tended to include more than one
sample village largely because of the sample clustering of villages in the 100 districts chosen for the
original ARIS survey.

The Global Land Information System (GLIS) of the EROS Data Center in Sioux Falls, South
Dakota provides information on all of the available individual scenes (i.e., images for particular dates) for
each selected path and row. A number of criteria were used to select specific scenes. First, we needed two
scenes for each path-row area that corresponded as closely as possible to the crop-years covered by the
ARIS (1970-71) and REDS (1981-1982) surveys. The most important constraints in matching by crop-
year are that the first Landsat satellite, Landsat 1, was not launched until late in 1972 and the availability
of the relevant scenes for India is limited after 1980. Second, in order to control for seasonal variation in
vegetative cover, we sought pairs of scenes for a given area that were at similar points within the crop-
cycle and corresponded to periods within the year during which there is minimal presence of standing crops, which can be difficult to distinguish from forest area using satellite imagery. Preliminary analysis suggested that the months of January and February were best. Third, because areas covered by clouds cannot be used to assess vegetative cover, we tried to minimize the extent to which cloud cover obscured the images (GLIS reports cloud cover for all scenes). The choice of the winter months was also useful in this regard because the cloud-laden monsoon period was excluded.

We were reasonably successful in meeting all of these criteria. Ninety-six percent of the scenes corresponding to the ARIS survey came from late 1972 and early 1973, with the scenes corresponding to the REDS survey distributed between years 1977 and 1980. The average number of years between scenes across path-row combinations is 5.1. While this is significantly less than the 11-year period between surveys, the measures of forest biomass created from the images indicate a significant drop-off of forests (see below). In addition, 81% of the selected scenes came from January and February, with all scenes coming from the November-April period. Finally, the level of cloud cover for the selected scenes never exceeds 2 on a 0-7 scale, with 0 denoting complete absence of clouds and 7 complete cloud cover.

For each of the selected 146 scenes corresponding to a specific path-row-day combination we obtained positive transparency images for both the near infra-red and red bands of the spectra. Each of these 292 9" square transparencies was scanned at a resolution of 300 dpi. The scanned images were then geo-registered using as registration points the four corners of the image, at which point the resolution was reduced so that each image consisted of a 500 by 500 pixel square. The two images were combined to construct the NDVI for each of the 250,000 pixels in each scene. A mask was then created for each image that screened out the margins of each image and for each village that screened out all areas more than 10km from the village location as determined by its latitude and longitude. By combining these two masks with the corresponding NDVI scenes we obtained a distribution of the values of the NDVI pixels within a 10km radius of each village for each of the surveys. The two primary summary measures used for these
distributions were the proportion of pixels with an NDVI > 0.2 (NDP) and the mean NDVI of those areas with an NDVI exceeding 0.2. The product of these two measures was also constructed as a measure of overall biomass attributable to forests (NDT).

Figure 2 displays on a district-level map of India the districts in which the sample villages are located and the average change in the proportion of the village areas that was forested according to the spectral-image satellite data. Those areas with the greatest amount of deforestation are indicated by the darkest shades and those with the largest increases in forest coverage are indicated by the lightest shades of grey (white areas are districts that do not contain sample villages). The constructed NDVI’s indicate that across all villages, weighted by village land area, the proportional drop in forested area, measured by NDP, was 20%. The areas with the largest decreases in forest include those located on the Eastern coast, in the states of Andhra Pradesh and Tamil Nadu, those on the Western coast, in Maharashtra and Karnataka and Gujurat, as well as those in the Punjab and in districts around Delhi, such as in Uttar Pradesh and Haryana, consistent with the state-level Richards and Flint data. But not all villages in the high-deforestation states lost forest area - there is substantial heterogeneity within states. For example, in Karnataka, a relatively densely populated state that ranked second overall in deforestation in Richard and Flint between 1970 and 1980, the 17 villages represented in our data experienced on average an 83% decline in forested area between 1972 and 1979. However, the satellite images indicate that six of the seventeen villages actually experienced an increase in forested area during that period.

The total number of villages for which we could obtain values of all of the variables at the two survey periods was 192. Table 2 provides descriptive statistics for the sample of villages at the two sample periods. While the villages were losing forest area during the 1970's, average household size in the villages evidently grew by 3.7% and the number of households grew by 44%, so that on average overall population size grew by 50% in the villages during the 11-year interval between surveys. The combination of population increase and forest area and tree losses resulted in falls in forest area and forest density per
<table>
<thead>
<tr>
<th>Variable</th>
<th>1970-71</th>
<th>1981-82</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forested area per household (hectares) (NDP)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.22</td>
<td>2.75</td>
</tr>
<tr>
<td>Mean forest biomass per household (NDT)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.04</td>
<td>.663</td>
</tr>
<tr>
<td>Mean household size</td>
<td>5.91</td>
<td>6.13</td>
</tr>
<tr>
<td>Mean number of households</td>
<td>353.8</td>
<td>512.0</td>
</tr>
<tr>
<td>Mean HYV-crop productivity (1971 rupees per acre)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>572.0</td>
<td>721.1</td>
</tr>
<tr>
<td>Factory present in village</td>
<td>.101</td>
<td>.538</td>
</tr>
<tr>
<td>Mean rural male wage (rupees per day)</td>
<td>6.15</td>
<td>7.90</td>
</tr>
<tr>
<td>Mean land price (rupees per acre)</td>
<td>8556.6</td>
<td>9124.9</td>
</tr>
<tr>
<td>Total rainfall in year (mm)</td>
<td>1345.8</td>
<td>1187.4</td>
</tr>
<tr>
<td>Proportion of households poor</td>
<td>.449</td>
<td>.434</td>
</tr>
<tr>
<td>Mean income of poor households</td>
<td>4140.1</td>
<td>4222.8</td>
</tr>
<tr>
<td>Mean income of non-poor households</td>
<td>9198.8</td>
<td>10379</td>
</tr>
</tbody>
</table>

<sup>a</sup>All monetary values in 1981 rupees based on all-India average rural cpi index, except for HYV productivity.

<sup>b</sup>Weighted by village land area.

<sup>c</sup>Standard deviation in parentheses.
household, the dependent variables used in the specification of the forest equilibrium equations, of 35% and 36%, respectively. However, HYV crop productivity also increased, by 26.1%, during the period, and there was considerable rural industrialization, as measured by the growth in the proportion of villages with a factory present, from 10% to more than 50%. The villagers experienced a real increase in wage rates of 28.4% and a 6.6% increase in real land prices. Average incomes also rose in real terms in the villages by 7.8%.

IV. Estimation procedures

We estimate log-linear approximations to (i) the reduced-form equations (4) relating the variation in population size, agricultural productivity and rural industry to the equilibrium values of the village wage, agricultural land price, average household income, and the measures of forest coverage, and (ii) the forest area equilibrium relationship (5), which provides a test of whether incomplete markets affects the use of forest lands. Estimates from the reduced forms enable an assessment of the validity of the framework and provide quantitative estimates of the effects of population growth, technical change and rural industrialization on deforestation and the paths through which these factors affect forests. The equilibrium equation estimates provide a test of whether markets are complete and land management is first-best efficient.

The reduced-form wage, land price, income and forest area estimating equations are given by

\[ z_t = b_{zt} + b_{zA} \theta_t + b_{zL} l_t + b_{z(1/\phi)} (1/\phi)_t + b_{zn} n_t + b_{zv} v + \epsilon_{zt} \]  

(16)

were for \( z=r \), the log of the average price of land in the village, \( w \), the log of the village male agricultural

\[ \text{For the two household-level variables used in the analysis - household size and household income - we computed the village mean of the log of each variable. In section V we also assess the robustness of the significance of the population variables household size and number in the forest-area equilibrium equation to functional form by estimating that equation using a translog (second-order) specification. For that test, the means of the (log) household-level income and household size squared and interaction terms were computed at the village level as well.} \]
wage rate, $y$, average log of household income in the village, and $a_p$, village forest area, measured both by per-household NDT and NDP. $\theta$ is measured by the four-crop productivity index, $l$ is the average log of household size in the village, $\lambda$ is measured by the log of the number of households in the village divided by village land area, and $\eta$, industrial infrastructure, is measured by the presence of a factory in the village.

The forest-equilibrium estimating equation is given by

$$a_{ft} = d_{ft} + d_{r_{ft}} + d_{w_{ft}} + d_{l_{ft}} + d_{(1/a_{ft})(1/a_{ft})} + d_{y_{ft}} + d_{\eta_{ft}} + \zeta_{ft}$$  \hspace{1cm} (17)

where the test of complete markets (efficient land management) is that $d_{l}=d_{(1/a)}=0$.

Estimation of (16) or (17) by least squares is not likely to yield consistent estimates because the environmental variable $v$, representing permanent attributes of the weather and soil as well as proximity to urban areas and markets, influences prices and incomes, and is likely to be correlated with the population variables, agricultural productivity, the presence of industry and the density and size of forests. For example, the productivity of green revolution high-yielding seed varieties is sensitive to water availability as well as to soil properties, and population density and perhaps tree growth, net of population effects, is likely to be higher where climate conditions are conducive to higher agricultural productivity. We can exploit the fact that we have data from two points in time to eliminate all fixed effects by differencing both equations across the two time periods, yielding

$$Dz_{t} = Dh_{zt} + b_{z0}D\theta_{t} + b_{zl}DL_{t} + b_{z(1/a)}D(1/a) + b_{z\eta}D\eta_{t} + D\varepsilon_{zt}$$  \hspace{1cm} (18)

and

$$Da_{ft} = Dd_{ft} + d_{fr_{ft}} + d_{w_{ft}} + d_{l_{ft}} + d_{(1/a_{ft})(1/a_{ft})} + d_{y_{ft}} + D\zeta_{ft}$$  \hspace{1cm} (19)

where $D$ is difference operator.

The differenced forms (18) and (19) still contain the time-varying errors $\varepsilon_{zt}$ and $\zeta_{ft}$, representing, for
example, weather realizations in the two time periods. These may also affect forest biomass, population growth, land productivity and industrialization. For example, a favorable weather outcome in the first period may increase incomes and thus increase subsequent fertility (e.g., by improving nutritional intake and thus fecundity). Weather shocks also affect contemporaneously crop productivity and also future crop productivity by affecting agricultural investments. In addition, shocks to forest area in (19) may directly affect the equilibrium price of land - a mandated increase in forest area in a given year, for example, which reduces the amount of crop land would raise the price of land. This would lead to a spurious positive relationship between land prices and forests.

We deal with the possibility of weather shocks biasing the differenced estimates by including a measure of weather for each time period - the rainfall in the relevant year in the district in which the village is located. In addition, to take into account the fact that rainfall does not fully capture weather and to avoid biases caused by other shocks, we use instruments to predict the over-time differences in the right-hand-side variables in both (18) and (19). Valid instruments include variables uninfluenced by the first-period or second-period errors and that affect the changes in the RHS variables over time. In particular, variables characterizing periods prior to the realization of shocks in the first survey year are valid instruments for the differenced forms of the reduced-form price and income and forest-equilibrium equations. We use village-level measures of the population and household size in 1970, variables measuring the distance of the village from the regional market (mandi) and the closest railway and bus station, and the presence of a factory and small-scale industry in the village in 1970. In addition, we make use of the inheritance information for land, farm equipment and other nonagricultural assets (which predict nonagricultural income variation) in the 1982 data to construct measures of farm and nonagricultural income.

\footnote{The relevant years for the reduced-form price and income equations are the NCAER ARIS and REDS survey years. For the forest area equations, the relevant years are those for which we constructed the satellite-based forest measures.}
wealth in the village for time periods preceding the initial survey. Finally, we use our village-specific estimates of technical change over the period 1971-82, which is stripped of the influence of endogenous investment, as an instrument for crop productivity change at the village level in equations (18). Use of all of these instruments, which include variables measuring proximity to transportation nodes and markets, presurvey initial agricultural and non-agricultural assets, initial demographic composition and the technical change measure, are required to identify the equilibrium reduced-form effects as well as the parameters of the equilibrium equations in differenced form in the presence of both unobserved shocks jointly affecting the growth of all variables in the specifications and feedback effects of forest area on equilibrium prices, incomes and population change.

V. Estimates

The instrumental-variables fixed-effects (FE-IV) estimates of the reduced-form price and income equations (18) are provided in Table 3. Estimates from two specifications of the reduced forms are shown, one in which the log of total village population size is substituted for the two (log) household size and household number variables for comparability to other estimates in the literature. The coefficients in Table 3 are in general estimated with precision and conform to those of the standard perfect markets model and the simulation effects in Table 1 for specific parameterizations of both that model and one in which forest labor cannot be monitored - increases in population size increase land prices and decrease wages, exogenous increases in crop productivity increase wages, land prices and incomes, and the presence of a factory in the village increases wages and incomes. The one exception is the effect of factory presence on village land prices. The estimates suggest that rural industrialization raises land prices, while the simulations indicate that increases in the demand for labor from industrialization draw labor from the agricultural sector and thus lower land prices. This discrepancy is likely due to the fact that the model assumes that rural land is not a factor of production in the non-agricultural, non-forest employment activity. In practice, of course, rural factories use land, and thus increase rural land values.
### Table 3
Effects of Population Size, Agricultural Productivity and Presence of Rural Industry on Log Land Prices, Log Wages and Log Household Income

<table>
<thead>
<tr>
<th>Variable</th>
<th>Log Land Price</th>
<th>Log Agricultural Wage</th>
<th>Log Household Income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>Log population size(^a)</td>
<td>.285</td>
<td>-</td>
<td>-.249</td>
</tr>
<tr>
<td></td>
<td>(2.14)</td>
<td>(3.15)</td>
<td>(1.16)</td>
</tr>
<tr>
<td>Log number of households(^a)</td>
<td>-</td>
<td>.289</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(2.13)</td>
<td>(2.95)</td>
<td>(0.52)</td>
</tr>
<tr>
<td>Log household size(^a)</td>
<td>-</td>
<td>.0069</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(3.30)</td>
<td>(2.76)</td>
</tr>
<tr>
<td>Log HYV productivity (rupees)(^a)</td>
<td>.401</td>
<td>.406</td>
<td>.115</td>
</tr>
<tr>
<td></td>
<td>(3.58)</td>
<td>(3.55)</td>
<td>(1.72)</td>
</tr>
<tr>
<td>Presence of factory in village(^a)</td>
<td>.462</td>
<td>.520</td>
<td>.456</td>
</tr>
<tr>
<td></td>
<td>(3.58)</td>
<td>(2.68)</td>
<td>(4.22)</td>
</tr>
<tr>
<td>Rainfall (mm x 10(^{-3}))</td>
<td>.164</td>
<td>.165</td>
<td>.364</td>
</tr>
<tr>
<td></td>
<td>(0.49)</td>
<td>(0.68)</td>
<td>(2.57)</td>
</tr>
<tr>
<td>t (1971-82)</td>
<td>.703</td>
<td>.957</td>
<td>.981</td>
</tr>
<tr>
<td></td>
<td>(6.83)</td>
<td>(13.1)</td>
<td>(16.0)</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>412</td>
<td>412</td>
<td>412</td>
</tr>
</tbody>
</table>

\(^a\)Endogenous variable. Instruments include pre-1970 (inherited) assets, average household size and total village population, and presence of factory, distances of the village from the nearest town, regional market (mandi), and railway and bus stations in 1970 and technical change estimate for 1971-1982.

\(^b\)Absolute value of t-ratio in parentheses.
The quantitative estimates of population growth, agricultural productivity change and rural industrialization on prices and incomes are relatively large in magnitude, as in the simulations. A doubling in population size increases land prices by 29% and lowers rural wages by 25%. Our estimates also suggest that doubling household size, for given number of households, lowers the rural wage by 67% and raises household income by only 50%, so that per-capita income declines by 25%. A doubling of crop productivity, however, increases land rentals by 40%, rural wages by 12%, and rural household incomes by 17-18%. Introducing a factory in the rural area increases rural land prices and wages by approximately 50% and rural incomes by from 14 to 21%.19

Table 4 reports the reduced-form FE-IV estimates for the two per-household forest-area measures. These indicate that population change, agricultural productivity growth and industrialization, along with rainfall variation, have powerful (and statistically significant) effects on forest change. The point estimates suggest that a doubling of population size would decrease forest area per household by 13% and average per-household forest density by 18%. Similarly, doubling crop productivity would reduce forest area and density, by 6% and 8%, respectively. Thus while technological change in agriculture evidently forestalls the Malthusian income trap (Table 3), it reinforces the destructive effects of population growth on forests. Rural industrialization, however, appears to save forests - introducing a factory in a village results in a 19% increase in forest area and a 25% increase in average forest density. To assess whether the latter effects were possibly due to an association between the presence of rural industry and the availability of alternative (to firewood) fuel sources, we included in another specification of the reduced-form forest

19Our reduced-form estimates of population and agricultural technology effects on wages are comparable to those obtained by Evenson (1993) for North India based on district-level data over a comparable period - his results indicate that population doubling decreases wage rates by 33% while an increase in productivity raises wages by 19%. Evenson's North Indian elasticity estimates also indicate that doubling population has a large positive effect on land rents - a doubling of population increases land rents by 84%, but in contrast to our findings, which appear to conform to both the model and the nature of the green revolution technology, his estimates suggest that agricultural productivity growth lower land rents.
Table 4
Effects of Population Size, Agricultural Productivity and Presence of Rural Industry on Forested Area (NDP) and Forest Biomass (NDT) per Household

<table>
<thead>
<tr>
<th>Variable</th>
<th>NDP</th>
<th></th>
<th>NDT</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Log population size(^a)</td>
<td>-.565</td>
<td>-</td>
<td>-.188</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(2.96)</td>
<td></td>
<td>(2.93)</td>
<td></td>
</tr>
<tr>
<td>Log number of households(^a)</td>
<td>-</td>
<td>-.584</td>
<td>-</td>
<td>-.194</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.09)</td>
<td></td>
<td>(3.07)</td>
</tr>
<tr>
<td>Log household size(^a)</td>
<td>-</td>
<td>-1.08</td>
<td>-</td>
<td>-3.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.60)</td>
<td></td>
<td>(3.62)</td>
</tr>
<tr>
<td>Log HYV productivity (rupees)(^a)</td>
<td>-.243</td>
<td>-.234</td>
<td>-.0831</td>
<td>-.0798</td>
</tr>
<tr>
<td></td>
<td>(3.55)</td>
<td>(3.45)</td>
<td>(3.62)</td>
<td>(3.52)</td>
</tr>
<tr>
<td>Presence of factory in village(^a)</td>
<td>.784</td>
<td>.774</td>
<td>.258</td>
<td>.255</td>
</tr>
<tr>
<td></td>
<td>(5.25)</td>
<td>(5.24)</td>
<td>(5.16)</td>
<td>(5.16)</td>
</tr>
<tr>
<td>Rainfall (mm x 10(^3))</td>
<td>.369</td>
<td>.275</td>
<td>.122</td>
<td>.0895</td>
</tr>
<tr>
<td></td>
<td>(2.54)</td>
<td>(1.84)</td>
<td>(2.50)</td>
<td>(1.79)</td>
</tr>
<tr>
<td>t (1971-82)</td>
<td>-.0801</td>
<td>-.0573</td>
<td>-.0254</td>
<td>-.0176</td>
</tr>
<tr>
<td></td>
<td>(0.81)</td>
<td>(0.58)</td>
<td>(0.76)</td>
<td>(0.54)</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>384</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
</tbody>
</table>

\(^a\)Endogenous variable. Instruments include pre-1970 (inherited) assets, average household size and total village population, and presence of factory in 1970, distances of the village from the nearest town, regional market (mandi), and railway and bus stations, and technical change estimate for 1971-1982.

\(^b\)Absolute value of t-ratio in parentheses.
equations (not shown) a variable indicating whether or not the village had been electrified. While this variable is positively correlated with village factory presence, it did not have a significant effect on forest area.

Table 5 reports the estimates of the forest-area equilibrium relationships obtained for the purpose of testing the exclusion restriction associated with the perfect markets assumption. To assess the robustness of the test results to estimation procedure and specification, we ran the test using alternative estimation procedures and specifications. In particular, we estimated the equilibrium equations (i) using both fixed-effects and fixed effects with instruments, (ii) for specifications with household income and household consumption expenditure, (iii) and for both log-linear and generalized second-order (translog) specifications. The results based on the log-linear specification for the two estimation procedures using both household income and consumption expenditure are reported in Table 5.20

The estimates in Table 5 indicate that net of equilibrium wages, land prices, and incomes or consumption expenditure, variation in average household size or number of households still has a statistically significant and negative effect on forest area and forest density per household. This result does not change with estimation procedure, nor with whether household income or expenditure is included in the specification. Moreover, tests of the joint significance of the household size coefficients in the translog specifications of the forest area and density equilibrium equations (estimated with instruments) also lead to rejection of perfect-markets exclusion restriction \($F(8, 151)=2.95 \text{ and } 3.01,$ respectively\). Thus, the estimates resoundingly reject the hypothesis derived from the perfect markets model that the effects of household size on forests arise solely from their effects on equilibrium prices and incomes.

Finally, the estimates of the effects of land prices on forest area and density are, in contrast to

20To achieve identification of the parameters of the translog specification we also added squared and interaction terms for the instruments to the first-stage equations. No squared terms or interactions are part of the instrument set for the log-linear specifications. The 33 parameter estimates from the translog specification of the equilibrium equations are available from the authors upon request.
Table 5
Equilibrium Equations: Effects of Land Prices, Wage Rates, Population Size and Income on Forested Area (NDP) and Forest Biomass (NDT) per Household

<table>
<thead>
<tr>
<th>Variable</th>
<th>NDP</th>
<th>NDT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FE</td>
<td>FE-IV</td>
</tr>
<tr>
<td>Log of land price*</td>
<td>-.0753</td>
<td>-.584</td>
</tr>
<tr>
<td></td>
<td>(0.79)</td>
<td>(2.22)</td>
</tr>
<tr>
<td>Log of wage rate*</td>
<td>-.202</td>
<td>.290</td>
</tr>
<tr>
<td></td>
<td>(1.51)</td>
<td>(1.00)</td>
</tr>
<tr>
<td>Log of household size*</td>
<td>-.602</td>
<td>-.924</td>
</tr>
<tr>
<td></td>
<td>(2.53)</td>
<td>(1.97)</td>
</tr>
<tr>
<td>Log of number of households*</td>
<td>-.336</td>
<td>-.473</td>
</tr>
<tr>
<td></td>
<td>(2.79)</td>
<td>(1.84)</td>
</tr>
<tr>
<td>Log household income*</td>
<td>-.223</td>
<td>-.358</td>
</tr>
<tr>
<td></td>
<td>(1.76)</td>
<td>(1.56)</td>
</tr>
<tr>
<td>Log household consumption*</td>
<td>-</td>
<td>-1.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.84)</td>
</tr>
<tr>
<td>Rainfall (mm x 10^4)</td>
<td>.110</td>
<td>.932</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.58)</td>
</tr>
<tr>
<td>t (1971-82)</td>
<td>.762</td>
<td>.841</td>
</tr>
<tr>
<td></td>
<td>(4.25)</td>
<td>(2.37)</td>
</tr>
</tbody>
</table>

*Endogenous variable. Instruments include pre-1970 (inherited) assets, average household size and total village population, presence of factory in the village, distances of the village from the nearest town, regional market (mandi), and railway and bus stations in 1970 and technical change estimate for 1971-1982.

bAbsolute value of t-ratio in parentheses.
those pertaining to the household variables, sensitive to estimation procedure. Neglect of the possibility that shifts in the amount of land devoted to forest area directly affect land prices results in an expected upward bias in the estimated land price effect on forest area. Use of instruments (specification (2)) leads to a statistically significant negative coefficient for land price on forest area and density, in conformity to at least one parameterization of the unmonitorable forest labor setting (Table 1). The FE-IV coefficient is approximately 8 times the magnitude in absolute value of that estimated without instruments.

VI. Conclusion

The issue of environmental degradation in rural areas of developing countries has received substantial attention in recent years and has been cited as a possible motivation for a wide variety of programs and policies. One important area in particular concerns the efficiency with which land is managed, particularly land that is communally owned. One important example of communal land use is forests. Despite the importance of forest management in terms of global externalities, however, little is known about the magnitudes and, in some cases, the signs of the consequences of agricultural technical change, population growth, and rural industrial growth for forest cover and deforestation. Even less is known about the mechanisms underlying these relationships and, in particular, the relative efficiency of alternative mechanisms of forest-resource management. Unfortunately, existing evidence on the efficiency of local management of forest resources is limited to case studies that cannot easily be generalized.

In this paper, we have assembled a village-level data set based on longitudinal household survey data combined with satellite images that cover a wide area of rural India over a twelve-year period to partially address limitations to knowledge about population and income effects on land management in general and forest exploitation in particular. We use the data not only to obtain estimates of population and income effects on forest use and to identify the mechanisms by which these factors affect land use but to address the question of whether forest areas are first-best locally managed. We show, based on the solution to a three-sector general-equilibrium representative agent model, that the proposition that forest area is
first-best efficiently managed, as would be the case given complete markets for land and labor, implies an exclusion restriction, namely that population growth does not have any effects on forest area or forest resource extraction net of its effects on equilibrium wages, land rents and incomes. Using a specific parameterization of the model we also established that this test has power against the alternative hypothesis that forest labor is unmonitorable but forest land allocations are second-best efficient given this limitation.

Our reduced-form econometric results confirm that higher population densities do indeed lead to reductions in forest area and density, but indicate that the effect of income growth on forests is importantly conditioned by the specific mechanisms causing incomes to increase. In particular, we find that both agricultural technical change and rural industrial growth increase equilibrium land rents, wages, and incomes but, due to differentials in the relative magnitudes of these effects, the former results in decreased forest area while the latter has the opposite effect. These results suggest that agricultural technical change, while forestalling the negative income consequences of a Malthusian equilibrium, exacerbate the negative consequences of population growth on forest exploitation. Our estimates also reject the restriction implied by the first-best model of land allocation for a variety of estimation procedures and specifications, and appear to conform reasonably well to the second-best model in which common forest land is optimally managed subject to limitations on controlling labor allocations, so that forests are overexploited. However, this second-best model is not easily distinguished empirically from a broad class of alternative models including ones in which forest land allocations are not even second-best efficient.

There are some significant shortcomings of these analyses, some of which will be addressed in future work. First, although the econometric analysis in this paper has allowed for the possibility that population change is endogenous with respect to changes in forest area, as well as price changes, the theoretical model does not explicitly allow for feedback between population and forests. The importance of examining this feedback may be seen, for example, by noting that if local institutions can limit fertility it is at least possible that any externality to fertility generated by the failure of the forest-labor market has
already been internalized at the village level thus limiting the need for further intervention. Second, while this work has been carried out for India as a whole, although at the village level, it is likely that there is substantial variation across India in the efficacy of institutions that manage forest resources. Analysis of the relationship between the degree of sub-optimality in land use and the organization of local governments may help to provide insight into the conditions under which interventions from outside the village are most likely to be productive. Third, this analysis considers only local market failures with respect to forest cover. No attempt has been made to address the issue of forest externalities associated with the process of global warming and diminishing biodiversity. While the methods used in this paper are clearly not appropriate for evaluating the extent of these effects, the estimates provided might be used to examine the extent and nature of subsidies likely to most effectively address this problem.
Appendix I

Derivation of Forest-Area Equilibrium Equation for Parametric

Model: Unmonitorable Forest Labor

The production functions are assumed to be \( f(a, l) = a_f^{1/2} l_f^{1/2} \), \( g(\theta, a, l) = \theta a_g^{1/2} l_g^{1/2} \), and \( m(\eta, l_f) = \eta l_m^{1/2} \) and the utility function is assumed to be \( u(c_f, c_g) = c_f^{\alpha} c_g^{1-\alpha} \). Income maximization with respect to labor in the three sectors and agricultural land, given forest land allocations, wages and prices and assuming unmonitorable forest labor yields \( l_f = \frac{f_r}{w} \), \( l_m = \frac{1}{4} \), \( l_g = \frac{1}{4} \), and \( a_g = \frac{1}{r} \). These combined with the labor-market clearing condition yield expressions for wages and rentals, \( w = \frac{q}{2l} \) and \( r = \frac{\theta^2}{2q} \), where \( q = \sqrt{\theta^2(a-a_f)+\eta^2} \), as functions of forest area and the endowments. Substitution into the expression for income yields \( y = q - \omega \). Inversion of the wage, rental and income equations with respect to \( \theta, \eta \), and \( \omega \) yields \( \theta = \sqrt{r w}, \eta = \sqrt{b w^2 - r w (a-a_f)}, \) and \( \omega = y - 2 w l \).

Optimal forest good consumption given income \( y \) and the price of forest goods \( p_f, c_f = \alpha \frac{y}{p_f} \), is obtained by maximizing utility subject to the budget constraint \( p_f c_f + c_g = y \). Substitution of the expressions for income and forest labor into the forest-products equilibrium equation, \( c_f = f(a, l) \), yields an equation that may be solved for the equilibrium price of forest goods, \( p_f = \frac{\sqrt{awy}}{\sqrt{a_f}} = \frac{\sqrt{aq(q+\omega)}}{\sqrt{2a_f}} \). Thus, indirect utility

\[
V(a_f) = (q + \omega)^{1-\alpha} \left( \frac{2\alpha a_f}{q} \right)^{\frac{\alpha}{2}} (1-\alpha)^{1-\alpha}.
\]

The social planner chooses \( a_f \) to optimize the welfare of the representative agent. Differentiating \( V(a_f) \) with respect to \( a_f \) yields a first-order condition that may be solved to obtain optimal forest area as a function of the exogenous variables. Alternatively, by substituting the expressions for \( \theta, \eta \), and \( V \) and then solving one obtains the equilibrium condition

\[
a_f = \left( \frac{2}{\alpha} - \frac{V}{2wl} - 1 \right) y
\]
References


National Laboratory, Environmental Sciences Division Publication No. 4174.


Vanneman, Reeve and Douglas Barnes, 1993, "Indian District Data, Release 3" (November).
