The Effect of Inequality on Growth: Theory and Evidence from the Indian States

Sugata Ghosh and Sarmistha Pal*

Abstract
The paper examines the effect of inequality on growth among the subnational states in India. Theoretically, growth of the regional economy is driven by productive public investment in the provision of health and education services financed by a linear output tax, and the optimum tax rate is determined by the median voter. In contrast to existing results, the authors obtain an ambiguous relationship between initial inequality and subsequent economic growth. Analysis of the Indian state-level data suggests that rural inequality influences growth of total output more than urban inequality, and does so negatively. The indicator of intersectoral inequality is more important in explaining sectoral output growth.

1. Introduction
Existing studies for developing countries (e.g., Ravallion, 1995; Ravallion and Datt, 1996) have often emphasized the need to boost economic growth in an attempt to reduce poverty. The relationship between growth and poverty is complex and depends, to a large extent, on the relationship between growth and inequality (Datt and Ravallion, 1992). If there is a rise in inequality while the economy is growing, this may not only offset the poverty-reducing effects of growth, but may also retard subsequent growth through redistribution in favor of nonaccumulable factors.

The literature on the effect of inequality on growth has gained momentum since the influential papers by Alesina and Rodrik (1994), and Persson and Tabellini (1994) (hereafter we abbreviate these authors as AR and PT, respectively). In both papers there is a redistributive role for the government to combat inequality within a democratic setup. In AR, government investment in productive services financed primarily through taxation of capital will interact with the growth-enhancing policies. With a tax on capital, given the familiar incentive and disincentive effects on capital income, a “pure” capitalist (who has no labor income) prefers the growth-maximizing tax rate. Higher inequality (defined in terms of the distribution of labor endowment relative to capital in the cross-section of population) will, however, induce the median voter (who has some labor income) to prefer a tax rate that is greater than the growth-maximizing tax rate, thus lowering growth. Similar results are obtained in the overlapping-generations (two-period) model of PT, where taxes are used only for redistribution; thus a higher rate of capital tax unambiguously depresses the incentive for private investment and growth. Both papers analyze the effects of the political outcome (by

---

*Ghosh, Pal: Cardiff Business School, Colum Drive, Cardiff CF10 3EU, UK. Ghosh: Tel: 44-29-20875515; Fax: 44-29-20874419; E-mail: ghosh@cardiff.ac.uk. Pal: Tel: 44-29-20876164; Fax: 44-29-20874419; E-mail: pal@cardiff.ac.uk. We would like to thank Berk Ozler for providing us with the data, and Debasis Bandypadhyay, Parantap Basu, Jean Drieze, Gerry Makepeace, Andrew McKay, Patrick Minford, Dani Rodrik, Udayan Roy, participants of the Econometric Society 8th World Congress, Seattle, 2000, and an anonymous referee for helpful comments and suggestions. Any remaining errors are ours.

© Blackwell Publishing Ltd 2004, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA
assuming a voting process on the level of the tax rate) generated by a given income distribution, and suggest that countries with greater economic inequality experience lower future economic growth. Partridge (1997) empirically examines the nature of the inequality–growth relationship at a subnational level for the states within the United States and finds that there is a positive relationship between initial inequality and subsequent growth. Factors like free interregional mobility of physical and human capital were conjectured to be responsible for this positive relationship, though nothing was formally tested. Another interesting paper on the inequality–growth issue at a subnational level is by Ravallion (1998) for 131 counties of rural China for the period 1985–90, where he argues that aggregation affects conventional tests of whether inequality impedes growth.

The present paper examines the inequality–growth relationship for a developing country (India) and modifies the theoretical model developed by AR or PT to include features that resemble the subnational Indian states. Among the features of the theoretical model are: a model of overlapping families of infinitely-lived agents à la Weil (1989), and more importantly, an output tax (instead of a capital tax as in AR or PT). Output taxes closely resemble the taxes on sales and purchases which are a very important source of state revenues in India. We also consider the provision of public services in health and education at the state-level which is entirely under the jurisdiction of the Indian states, as enumerated in the “state list” of the Indian constitution. In section 2, we consider an endogenous growth model where growth is driven by productive investment in the provision of such public services financed by an output tax, and the optimum tax rate is determined by the political process (median voter rule) as in a democratic setup like India. As the tax is on output, it affects both labor and capital income. Since the median voter has both kinds of income, the effect of an output tax on the median individual’s utility is ambiguous. Consequently, the median voter will choose an optimum tax that may be higher or lower than the growth-maximizing tax rate depending on whether the redistribution raises or lowers his/her net labor income.

Using the state-level data for the period 1960–94, the second part of the paper empirically examines the nature of the inequality–growth relationship for the Indian states: section 3 describes the data while section 4 reports the econometric analyses. We perceive that there is some value-added in our empirical exercise: first, most existing empirical studies for developing countries examine the effect of growth on ensuing inequality (e.g., Ravallion, 1995; Ravallion and Datt, 1996). The only study that considers the two-way relationship between growth and inequality is that by Deininger and Squire (1998). Using data from a cross-section of countries, they find little support for the Kuznets hypothesis while there is a strong negative relationship between growth and inequality. There are obvious problems with the cross-country comparisons because of heterogeneity among national economies. The problem is minimized if one considers the subnational states instead.

India is an interesting case in point, as the importance of economic growth has long been emphasized in its fight against poverty. Within the federal setup, Indian states are also sufficiently diverse in terms of geographic, demographic, and economic characteristics. Existing empirical studies consider intercountry comparisons (see AR, PT, Deininger and Squire) or interregional comparison for a developed country (Partridge), while we focus on the interstate comparison for the developing economy of India. Also given the pronounced dichotomy between rural and urban sectors of an Indian state, we examine the sectoral nature of this relationship, something that has not been explored before.
2. Theoretical Framework

In this section we develop a theoretical model of growth with redistribution at a regional level. The national economy consists of \( n \) similar subnational regions (or “states” within a country). A single representative region is denoted by \( a \). Producers in region \( a \) comprise a large number of individuals. The output of individual \( i \) in region \( a \) is

\[
y'_a = A_i \left( k_{i, a}^{\gamma_1} (l_{i, a}^{\gamma_2} G_a) \right)^{1-\gamma_2} K_a \geq 0, \quad 0 < \gamma_1 + \gamma_2 < 1
\]

where \( k_{i, a} \) is his capital stock endowment and \( l_{i, a} \) is the number of units of labor supplied. \( G_a \) is the region-wide public investment in health and educational services, and \( K_a \) is the capital stock of the regional economy. The inclusion of \( G_a \) and \( K_a \) in the production function represents externality effects arising out of public investment in the region à la Barro (1990) and “knowledge” effects from the capital stock of all firms in the regional economy à la Romer (1986), respectively. Aggregating over all individuals in region \( a \), given CRS, we have:

\[
Y_a = A_i K_a^{1-\gamma_2} G_a^{\gamma_2}.
\]

In deriving (2), the aggregate labor endowment \( L_a (= L_0 e^{\eta t}) \) is normalized to unity by suitable choice of units. Here \( \eta \) is the rate of growth of the labor force.

The government balances its budget in every period, but unlike Alesina and Rodrik who consider taxation of capital only, we consider the taxation of overall factor income (or output), given by \( \tau_a \):

\[
G_a = T_a = \tau_a Y_a.
\]

Defining \( w_a \) and \( r_a \) as the competitively determined private marginal products of labor and capital, respectively, we have

\[
w_a = (1 - \tau_a) \frac{\partial y'_a}{\partial l_{i, a}} = (1 - \tau_a)(1 - \gamma_1)A_i\left( \frac{k_{i, a}}{l_{i, a}} \right)^{\gamma_1} \tau_a^{\gamma_2} Y_a^{\gamma_2} K_a^{1-\gamma_1-\gamma_2},
\]

\[
r_a = (1 - \tau_a) \frac{\partial y'_a}{\partial k_{i, a}} = (1 - \tau_a)\gamma_1 A_i \left( \frac{k_{i, a}}{l_{i, a}} \right)^{\gamma_1} \tau_a^{\gamma_2} Y_a^{\gamma_2} K_a^{1-\gamma_1-\gamma_2}.
\]

The consumption side of our model is a variant of the Weil (1989) model where the economy of region \( a \) consists of many infinitely-lived families which are not linked to preexisting cohorts through intergenerational transfers.

The utility function of an agent \( i \) (representative of the cohort born at time \( s \)) as of time \( t \) in region \( a \) is logarithmic:

\[
U_t = \int_s^t \ln \left[ c_i(s, v) e^{-\theta (v-s)} \right] dv.
\]

This is maximized subject to a dynamic budget constraint:

\[
k_{i, a} = (1 - \tau_a)y'_a - c'_a = \left[ r_a k_{i, a} + w_a l_{i, a} \right] - c'_a.
\]

The optimization for the representative individual then gives

\[
\frac{d}{dt}(c_i) = (r_a - \theta)c_i.
\]

Integrating (7) and (8) and appropriately combining them yields
where \( \omega_i^t \) is the present discounted value of future labor income accruing to those currently alive; i.e.,

\[
\omega_i(s, t) = \int_0^\infty (w_i l_i) e^{-\frac{1}{\theta}(\mu)} d\mu = \frac{w_i l_i}{r_a},
\]

We next consider the political process of majority voting through which the government chooses that tax rate which maximizes the welfare of the median voter. Substituting (8) and (9) into the utility function (6), and then integrating by parts, we have an expression for the indirect utility function for the \( i \)th individual (\( u_i^t \)) which can be written as follows:

\[
u_i^t = \left( \frac{1}{\theta} \right) \left[ \ln(\theta) + \ln(w_i l_i) - \ln r_a \right] + \frac{r_a - \theta}{\theta^2}.
\]

The optimal tax \( t_i^\ast \) satisfies \( du_i^t/dt_a = 0 \), and the second-order condition. Note that \( u_i^t \) is given by (11), and \( w_a \) and \( r_a \) in (11) are given by (4) and (5). This enables us to obtain expressions for \( \partial w_a/\partial t_a \) and \( \partial r_a/\partial t_a \).

We can thus obtain the optimal tax rate \( t_a^\ast \) which will depend on \( s_i \), where \( s_i \) is defined as

\[s_i = \frac{Y_{a,L}}{L_a} \leq \frac{Y_{a,K}}{K_a} = t_a^\ast/\frac{L_a}{K_a}.
\]

Here \( y^{L_i}_a \) and \( y^{K_i}_a \) are individual \( i \)'s income from labor and capital, respectively; and \( Y^{L,a}_a \), \( Y^{K,a}_a \) are aggregate incomes of region \( a \) from labor and capital, respectively. Thus the choice of \( t_a^\ast \) will depend on the value of \( \sigma' \). Given an initial distribution of labor and capital income (corresponding to certain value of \( t_a \)), a higher \( t_a \) ought to change the individual’s labor/capital income ratio and also that of the region (i.e., both the numerator and denominator of the above expression will change). It is therefore not clear how \( \sigma' \) will change with respect to \( t_a \), and consequently the median voter’s optimum tax rate \( t_a^\ast \) will be ambiguously related to \( \sigma' \).

The intuition for this comes from the fact that here the tax is on output, so there is an incentive as well as a disincentive effect on both labor and capital income. This is in contrast to AR where labor income responds positively to a higher tax rate (which is exclusively on capital). Consequently, in our case, the median voter—who has both labor and capital income—responds ambiguously to higher \( t_a \). The “pure” capitalist (as in AR) prefers the growth-maximizing tax rate (\( \hat{t}_a \)), but here the median voter’s ideal \( t_a \) (\( t_a^\ast \)) could be greater or less than \( \hat{t}_a \). Therefore, the effect of initial inequality on growth could be negative or positive.

In order to derive the growth-maximizing tax rate, we use the goods market equilibrium condition for region \( a \) in per capita terms:

\[
\varepsilon_a = \frac{k_a}{k_a} = \frac{y_a - c_a - g_a}{k_a} - n,
\]

where \( \varepsilon_a \) is the per capita (endogenous) growth rate for the region. Also, \( k_a = K_a/L_a \), \( y_a = Y_a/L_a, c_a = C_a/L_a, g_a = G_a/L_a \), \( \tau_a L_a \), and \( \tau_a = \tau_a L_a \).
The evolution of aggregate consumption can be shown to satisfy \( \dot{C}_a = (r_a + n - \theta) \times C_a - n\theta K_a \), from which it follows that

\[
\varepsilon_a = \frac{\dot{c}_a}{c_a} = r_a - \theta - n\theta \frac{k_a}{c_a}.
\]

(13)

In the steady state, \( \frac{dK_a}{dt} = 0 \), and also \( \frac{dc_a}{dt} = 0 \) by definition. Combining equations (12) and (13) by eliminating \( c_a \), and then using the balanced-budget constraint (3) in per capita terms, gives us an equation (not shown) linking \( \varepsilon_a \) and \( \tau_a \), from which the Barro (1990)-type inverse U-shaped relation between \( \varepsilon_a \) and \( \tau_a \) emerges—here with overlapping families of infinitely-lived agents. The growth-maximizing tax rate (\( \hat{\tau}_a \)) satisfies \( \frac{d\varepsilon_a}{d\tau_a} = 0 \).

We next consider the important case that is typical of many developing countries, where a proportion (\( \mu \)) of the total tax revenues \( T_a \) is used by the government as redistributive transfers to the labor component of individual income. This means that the proceeds of the tax revenue augment an individual’s labor income by the amount \( \mu \tau_a y_a \). Labor income after the transfer is therefore \( w_d t_a + \mu \tau_a y_a \), and capital income is \( r_a k_a \). Also, the government spends \( (1 - \mu) \tau_a y_a \) on health and education schemes, which means that we now have the government budget constraint as \( G_a = (1 - \mu) \tau_a Y_a \).

By retracing the steps as before, we can find the tax rate that maximizes the utility of the median individual. Once again, it is unclear whether the median voter prefers a tax rate (\( \tau_a^* \)) higher than the growth-maximizing tax rate (\( \hat{\tau}_a \)). What we can say, however, is that with redistributive transfers from the output tax revenue, the median voter’s utility-maximizing tax rate ought to be higher than without the redistribution scheme. In this sense the possibility of initial inequality having a negative impact on growth (as in AR) is increased, but it is still not certain that this would be the case.

3. A Case Study of Indian States

Sections 3 and 4 will examine the nature of the relationship between inequality and growth, using state-level data from India over 1960–94. The Indian constitution gave strong economic/financial powers to the national government (center) with respect to the core sectors, including industry, defence, railways, post and telegraph, and atomic energy; while states have primary control in health and education. The most elastic sources of tax revenue, namely income, excise and customs taxes, are levied by and accrue to the center, while taxes on properties, purchases, and sales (most important source of state tax revenue) are levied by the states. However, most expenditures are incurred at the state level. Given this imbalance in financial powers between the center and the states, there are different mechanisms of transfer of resources from the center to the states. Compared to states’ share in central taxes, proportions of state taxes are generally higher in total state revenues, which finance various developmental and nondevelopmental expenditure. The bulk of the state revenues is spent on the development account, which includes expenditure on health, education, and community services; an interstate variation is noted here.

Our theoretical model attempts to capture the important characteristics of the Indian federal states. For example, a proportional output tax closely resembles taxes on sales and purchases, which are one of the main sources of state taxes in India. Also, public development expenditure in the provision of social and economic services features prominently in the Indian states, and its inclusion in the production function thus captures an important characteristic of the Indian subnational economy. Finally, we
consider the political process of majority voting to determine the optimum tax rate, which reflects the democratic nature of the process in line with the Indian practice.

**Testable Hypotheses**

For given values of the rate of growth of population \((n)\) and the rate of time preference \((\theta)\)—taken to be identical for all regions under consideration—growth of output in any state \(a (= 1, 2, \ldots, j)\) over a period of time is a function of initial output, \(Y_{a0}\), capital, \(k_{a0}\), and some index of initial distribution, \(\lambda_{a0}\):

\[
\varepsilon_a = f(Y_{a0}, k_{a0}, \lambda_{a0}).
\]  

(14)

Inclusion of \(Y_{a0}\) as the level of output or income in the initial period allows us to test the validity of the hypothesis of growth convergence as advocated by Barro (1991). In endogenous growth models with externalities, convergence holds only if measures of initial human capital are held constant (Barro, 1991). Thus for a given level of initial capital, \(Y_{a0}\) is expected to have a negative influence on growth in the cross-sectional analysis. \(\kappa_{a0}\) refers to indices of both physical and human capital, both of which are expected to exert a positive influence on growth (also see discussion in section 4). The central hypothesis of our concern is that, other things remaining unchanged, economic growth \(\varepsilon_a\) in any state \(a\) depends on initial inequality \(\lambda_{a0}\) in the state: initial inequality may have a negative or positive effect on growth, depending on whether the median voter desires a tax rate that is higher or lower than the growth-maximizing rate.

In the Indian context, the median voter is clearly one who has more labor income than capital income, which follows from the fact that India is a labor-surplus economy. It is also true that popularly elected governments undertake redistributive transfers in favor of the poor (whose income is largely labor income). These two factors are more likely to result in a negative relation between inequality and growth, as the median voter’s preferred tax rate is more likely to be higher than the (regional) economy’s growth-maximizing tax rate.4

Indian states are often characterized by the pronounced dichotomy between rural and urban sectors within a state. Although in the theoretical section we do not model the rural and urban sectors explicitly, we can still obtain testable hypotheses regarding the link between initial inequality and growth from a sectoral perspective. If we now make the plausible and quite realistic assumptions that (1) the rural sector comprises individuals having mostly labor income, and (2) the government levies taxes to redistribute incomes in favor of the rural sector, then the median voter (who in the Indian context has more labor than capital income and thus can be assumed without much loss of generality to belong to the rural economy) would favor higher taxes. And given that the growth rate has an inverse U-shaped dependence on the tax rate, it is more likely that the median voter’s ideal tax rate will lie to the right of the growth-maximizing rate. In other words, rural inequality matters more in the Indian context, and it is also more likely that higher initial rural inequality would lower growth.

**Description of Data**

The data used for our purposes are obtained primarily from various government sources like the National Sample Survey (NSS), government accounts, and a dataset compiled by the World Bank (Ozler et al., 1996). This latter is a unique dataset that contains, among other things: information on net sown area (for all crops); net state
domestic product (SDP) including sectoral SDPs for agriculture, manufacturing, etc.; population, rural, and urban Gini coefficients in the distribution of consumer expenditure; and state-level developmental expenditure on the public provision of social and economic services, including health and education, for 16 major states in India. For the analysis of this paper, we study the period 1960–94 for which the dataset is complete. This basic dataset has been supplemented by the information on literacy (source: Reports of the Census and the Education Department, Government of India, various issues) and also that on state-level taxes and expenses (source: Reserve Bank of India) for these states.

An analysis of rural and urban inequality Gini indices in 1960 and output growth per capita over 1960–94 does not suggest any specific pattern (positive or negative) in the relationship between initial inequality and subsequent growth. There are states with low (high) initial inequality and high (low) subsequent growth, as well as those with low (high) initial inequality and low (high) subsequent growth. For example, among the six most initially unequal rural regions (Rajasthan, MP, Karnataka, Kerala, AP, and Gujarat), the annual rate of growth was low (less than 1%) in four (the exceptions being AP and Gujarat). Among the states with low rural inequality in 1960, Bihar, Jammu and Kashmir (J&K), and West Bengal had low growth (below 1%), while Maharashtra and Tamil Nadu had high annual growth rates around 4%. Punjab witnessed one of the highest rates of growth among the Indian states, and levels of initial rural and urban inequality have been seventh largest among the selected states. Table 1 shows the bivariate correlation coefficients between growth of output per capita and rural and urban Gini indices. Annual growth rate of per capita output is negatively and statistically significantly related to both rural and urban inequality Gini indices. There is a significant and positive correlation between per capita development expenditure and state-level tax rates, and between Gini indices and state-level tax rates, as assumed in our theoretical model.

The bivariate correlation analysis assumes that the states under consideration are identical in all respects other than growth and inequality. There is, however, significant interstate variation in the initial levels of income per capita, literacy rates, net sown area, and per capita state development expenditure. This, in turn, necessitates an assessment of the inequality–growth relationship in a multiple regression framework, after controlling for all possible factors affecting the relationship.

Table 1. Bivariate Correlation Coefficients

<table>
<thead>
<tr>
<th></th>
<th>RGINI60</th>
<th>UGINI60</th>
<th>PCDEXP60</th>
<th>TAXRT60</th>
<th>GRPC6094</th>
<th>GRAG6088</th>
<th>GRMF6088</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGINI60</td>
<td>1.000</td>
<td>0.741**</td>
<td>0.638**</td>
<td>0.247*</td>
<td>-0.654**</td>
<td>-0.475</td>
<td>-0.694**</td>
</tr>
<tr>
<td>UGINI60</td>
<td>0.741**</td>
<td>1.000</td>
<td>0.391*</td>
<td>0.304*</td>
<td>-0.501*</td>
<td>-0.298</td>
<td>-0.506**</td>
</tr>
<tr>
<td>PCDEXP60</td>
<td>0.638**</td>
<td>0.391</td>
<td>1.000</td>
<td>0.490*</td>
<td>-0.248</td>
<td>-0.231</td>
<td>-0.348</td>
</tr>
<tr>
<td>TAXRT60</td>
<td>0.247*</td>
<td>0.304*</td>
<td>0.490*</td>
<td>1.000</td>
<td>—</td>
<td>-0.286</td>
<td>-0.265</td>
</tr>
<tr>
<td>GRPC6094</td>
<td>-0.654**</td>
<td>-0.501*</td>
<td>-0.248</td>
<td>—</td>
<td>1.000</td>
<td>0.574*</td>
<td>0.641**</td>
</tr>
<tr>
<td>GRAG6088</td>
<td>-0.475</td>
<td>-0.298</td>
<td>-0.231</td>
<td>-0.286</td>
<td>0.574*</td>
<td>1.000</td>
<td>0.673**</td>
</tr>
<tr>
<td>GRMF6088</td>
<td>-0.694**</td>
<td>-0.506*</td>
<td>-0.348</td>
<td>-0.265</td>
<td>0.641*</td>
<td>0.673**</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Notes: * and ** denote the level of significance at 10% and 1%, respectively. RGINI60, UGINI60, PCDEXP60, TAXRT60: initial values of rural Gini, urban Gini, per capita development expenditure, and tax rate in 1960, respectively. GRPC6094: annual rate of growth of SDP per capita during 1960–94; GRAG6088, GRMF6088: annual rate of growth of agricultural and manufacturing SDP per capita during 1960–88.
4. Econometric Analysis

Based on a multiple regression framework, this section examines the effect of initial inequality on subsequent economic growth in the Indian states.

Single Cross-sectional Estimates

The simplest way to proceed is to make use of the single cross-sectional framework where the growth rate of output over the sample period 1960–94 is determined in terms of values of explanatory variables (including indices of inequality) prevailing in the initial year, 1960. The resultant sample is rather small as we have information on only 16 major Indian states observed during this period. Nevertheless, one may argue that a single cross-sectional framework is more suitable to analyze the effect of initial inequality on economic growth since it takes time for the political process to affect distribution and subsequent growth. Hence, we start with the single cross-sectional analysis of the complete model and then delete the insignificant variables to obtain a parsimonious model. The dependent variable of our interest is the rate of growth of per capita state domestic product (GRPCINC) over the period 1960–94.5

Choice of the explanatory variables is guided by equation (14). However, given the nature of the data at our disposal, we need to make some modifications to the growth equation we estimate. We express state domestic product, its growth, and also capital as a proportion of total population in per capita terms. For the Indian economy, the size distribution of income is not readily available; what we observe is the size distribution of per capita monthly consumer expenditure6 (available from the NSS reports) which in turn generates the estimates of inequality Gini indices used in our analysis. However, we observe separate indices only of the rural and urban inequality Gini indices, but not an index of total state-level inequality. Total inequality can, however, be thought of as an aggregate of rural inequality (RGINI), urban inequality (UGINI), and the inequality between rural and urban areas. In the absence of a better measure of the intersectoral inequality component, we construct the ratio of rural monthly mean consumption to urban monthly mean consumption and use this ratio (RRUMEAN) as an indicator of intersectoral inequality.7 The mean rural consumption is relatively higher than mean urban consumption if RRUMEAN is greater than one, and relatively lower if it is less than one. If the distribution of consumption is positively skewed, the median is always lower than the corresponding mean. Thus, we include UGINI, RGINI, and RRUMEAN as the three components of total inequality in the growth regression. One may also argue for the inclusion of the state tax rate in our regression. However, there is a significant statistical correlation between the tax rate and inequality indices. Accordingly, we include only inequality indices (and not the tax rate) in our regressions.

The set of explanatory variables includes per capita state domestic product (PCSDP60), rural Gini (RGINI60), and urban Gini (UGINI60) coefficients, and the intersectoral inequality component (RRUMEAN60) prevailing in the initial period, 1960. In view of the difficulty in obtaining an overall index of aggregate capital at the Indian state level (Loh, 1995), we include three indicators of capital—literacy rate (LITRT60) as an index of human capital, per capita sown area (PCAREA60), and per capita state expenditure in the provision of social and economic services (PCDEXP60)—in the year 1960 as indices of physical capital.

Means and standard deviations of explanatory variables are given in column 1 of Table 2. We first estimate the complete model (column 2) and then delete the insigni-
ficant variables to obtain a more parsimonious specification (column 3). These estimates use White’s correction for the presence of heteroskedasticity. $R^2$ and $F$ statistics describe the goodness of fit of each specification. A comparison of these two sets of estimates suggests that the parsimonious model yields slightly better results in terms of the $F$-statistic.

The main inferences from our estimates are worth noting. The coefficient of $PCSDP60$ is negative and significant, implying that states with a relatively higher level of initial output per capita have a significantly lower growth rate over the period 1960–94. This in turn lends support to Barro’s $\beta$ convergence hypothesis. It is also noteworthy that the rate of convergence among the Indian states has been rather low, as has been observed by Cashin and Sahay (1996). Secondly, the effects of $PCAREA60$, $PCDEXP60$, and $LITRT60$ on growth per capita reveal that the coefficients of all three types of capital are positive and significant: so both physical and human capital per capita significantly enhance growth of total output per capita, thus confirming our \textit{a priori} expectation. More importantly, the coefficient estimate of $RGINI60$ is negative and significant while that of $UGINI60$ is positive (but insignificant); thus initial rural inequality has a negative relationship with economic growth per capita in the ensuing period while initial urban inequality has a positive but \textit{insignificant} impact on growth. In addition, the intersectoral inequality index $RRUMEAN60$ is negative and significant. In terms of the testable hypotheses arising out of our theoretical model, the result on $RGINI$ seems plausible while that on $RRUMEAN60$ seems less plausible. As discussed in section 3, given our assumptions, we expect initial rural inequality (as indicated by the rural Gini index, $RGINI$) to have a negative impact on growth. With regard to the sign for $RRUMEAN60$, assuming that the median voter lives in the rural sector and that the popularly elected government uses an output tax to redistribute incomes in favor of this sector (see the discussion in section 3), the median voter will desire a relatively higher tax (which is expected to lead to lower growth) when mean

---

**Table 2. Single Cross-sectional OLS Estimates of Growth of Total Output**

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Mean (SD)</th>
<th>Coefficients (t-ratio)$^a$</th>
<th>Coefficients (t-ratio)$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Intercept</td>
<td>—</td>
<td>4.5 (21.694)**</td>
<td>4.5 (22.386)**</td>
</tr>
<tr>
<td>$PCSDP60$</td>
<td>308.1983 (82.22)</td>
<td>−0.003 (2.989)**</td>
<td>−0.002 (2.969)**</td>
</tr>
<tr>
<td>$RGINI60$</td>
<td>28.77 (8.22)</td>
<td>−0.05 (5.113)**</td>
<td>−0.05 (4.927)**</td>
</tr>
<tr>
<td>$UGINI60$</td>
<td>30.78</td>
<td>0.002 (0.827)</td>
<td>—</td>
</tr>
<tr>
<td>$RRUMEAN60$</td>
<td>0.80 (0.17)</td>
<td>−0.68 (2.304)*</td>
<td>−0.66 (2.273)*</td>
</tr>
<tr>
<td>$LITRT60$</td>
<td>0.23 (0.10)</td>
<td>2.27 (6.341)**</td>
<td>2.32 (7.202)**</td>
</tr>
<tr>
<td>$PCAREA60$</td>
<td>0.32 (0.19)</td>
<td>1.8 (7.452)**</td>
<td>1.8 (7.423)**</td>
</tr>
<tr>
<td>$PCDEXP60$</td>
<td>0.16 (0.09)</td>
<td>1.3 (2.414)*</td>
<td>1.17 (1.976)*</td>
</tr>
<tr>
<td>$R^2$</td>
<td>—</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>$F$-statistic</td>
<td>—</td>
<td>7.41**</td>
<td>9.52**</td>
</tr>
<tr>
<td>No. of observations</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>

$^a$These estimates use White’s correction for heteroskedasticity.

Notes: * denotes that the variable is significant at the 10% level, ** at the 1% level. $PCSDP60$, $LITRT60$, $PCAREA60$, $PCDEXP60$: initial values of total output per capita, literacy rate, net sown area, and state development expenditure per capita in 1960, respectively. $RGINI60$, $UGINI60$, $RRUMEAN60$: initial values of rural, urban and intersectoral inequality indices in 1960, respectively.
consumption in the rural sector is lower, and vice versa when the mean rural consumption is higher. In other words, we expect the coefficient of $RRUMEAN$ to be positive when regressed on growth.

One possible statistical reason for this inconsistency is the omitted-variable bias in the single cross-sectional regression that arises from ignoring the state-specific aspect of the aggregate production function. Islam (1995) argues that this omitted-variable bias can be redressed if one considers a panel-data approach instead, which takes us to the next section where we reexamine this inequality–growth relationship by considering both the state-specific effects as well as the variation over time.

**Panel-data Estimates**

Islam (1995) argued that the significant convergence result of cross-country regression analyses may be revised in a panel-data analysis when the differences in the aggregate production function across economies are allowed for. In our case, however, the use of panel data introduces individual state-specific effects as has been indicated by the striking diversity among the Indian states (see the discussion in section 3). The state-specific aspect of the aggregate production function that is ignored in the single cross-sectional regression is correlated with the included explanatory variables that creates omitted-variable bias. The panel approach as discussed below redresses this bias.

Let us now modify equation (14) to consider a simple fixed-effects model of the following form for the $s$th state, $s = 1, 2, \ldots, 16$:

$$e_{st} = \beta x_{st-1} + \phi_s + u_{st},$$

where $\phi_s$ is the fixed effect, $x_{st-1}$ is the $k \times 1$ vector of lagged exogenous regressors, and $u_{st}$ is the vector of random disturbance term. The fixed-effects $\phi_s$ of (15) will account for the unobserved differences in growth due to interstate differences in history and economic structure, and will capture the heterogeneity that causes the inconsistency in the OLS regression. We consider a fixed-effects rather than a random-effects model because the set of states selected in our dataset is unlikely to be random—it is rather a conscious choice largely determined by the availability of relevant information.

We hypothesize that the growth of output per capita for a given state $s$ over each subperiod $t$ depends on: output per capita ($PCSDP$); rural, urban, and intersectoral inequality indices ($RGINI$, $UGINI$, $RRUMEAN$); literacy rate ($LITRT$); per capita total sown area ($PCAREA$); and per capita state development expenditure ($PCDEXP$) at the beginning of that subperiod. We correct for the presence of heteroskedasticity in the error structure (using White’s correction) when state-level heterogeneity is considered in the fixed-effects model.

There are 35 annual observations (over 1960–94) for each of 16 major states in India, though there are missing observations on inequality Gini indices for the years when the NSS was not conducted. Since growth takes place over a period of time where the short-term disturbances for annual observations may feature prominently, we rearrange the annual data so that 35 years of annual observations for each state is divided into seven subperiods, namely 1960–63, 1964–68, 1969–73, 1974–78, 1979–83, 1984–88, and 1989–93. The choice of these subperiods has been dictated by the years for which NSS information on inequality Gini indices are available. Thus, the total number of observations for 16 states turns out to be 112. Using this five-yearly panel data, we estimate the fixed-effects least-squares dummy variable ($LSDV$) model (15). Means and standard deviations of the regression variables are shown in column 1 of Table 3, while the coefficient estimates with $t$-statistics (shown in parentheses) are

© Blackwell Publishing Ltd 2004
Table 3. Least-squares Dummy-variable Panel Estimates of Growth

<table>
<thead>
<tr>
<th>Variables</th>
<th>GRPCY</th>
<th></th>
<th>GRPCAGY</th>
<th></th>
<th>GRPCMFGY</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Coefficient (t-ratio)</td>
<td>Mean (SD)</td>
<td>Coefficient (t-ratio)</td>
<td>Coefficient (t-ratio)</td>
<td></td>
</tr>
<tr>
<td>PCSDP</td>
<td>363.4 (137.1)</td>
<td>-0.0002 (4.1)**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>PCAGY</td>
<td>—</td>
<td>—</td>
<td>156.1 (49.6)</td>
<td>-0.001 (2.1)*</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>PCMFGY</td>
<td>—</td>
<td>—</td>
<td>47.88 (30.4)</td>
<td>-0.15 (0.551)</td>
<td>−0.01 (1.65)*</td>
<td>—</td>
</tr>
<tr>
<td>RGINI</td>
<td>28.23 (6.26)</td>
<td>-0.003 (3.35)**</td>
<td>28.41 (6.68)</td>
<td>-0.15 (0.551)</td>
<td>28.41 (6.68)</td>
<td>-0.15 (0.551)</td>
</tr>
<tr>
<td>UGINI</td>
<td>31.78 (6.62)</td>
<td>—</td>
<td>31.63 (7.02)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>RRUMEAN</td>
<td>0.73 (0.16)</td>
<td>0.007 (1.688)*</td>
<td>0.73 (0.173)</td>
<td>0.2 (1.948)*</td>
<td>0.98 (1.917)*</td>
<td>0.98 (1.917)*</td>
</tr>
<tr>
<td>LITRT</td>
<td>0.37 (0.16)</td>
<td>0.23 (5.770)**</td>
<td>0.35 (0.15)</td>
<td>3.2 (2.227)*</td>
<td>16.5 (2.188)*</td>
<td>16.5 (2.188)*</td>
</tr>
<tr>
<td>PCAREA</td>
<td>0.27 (0.15)</td>
<td>—</td>
<td>0.28 (0.16)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>PCDEXP</td>
<td>0.39 (0.48)</td>
<td>—</td>
<td>0.34 (0.49)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>R²</td>
<td>—</td>
<td>0.36</td>
<td>—</td>
<td>0.42</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>F-statistic</td>
<td>—</td>
<td>2.77*</td>
<td>—</td>
<td>2.93*</td>
<td>2.67*</td>
<td>2.67*</td>
</tr>
<tr>
<td>No. of observations</td>
<td>112</td>
<td>112</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
</tr>
</tbody>
</table>

Notes: * denotes that the variable is significant at the 10% level, ** at the 1% level. RGINI, UGINI, RRUMEAN: rural, urban, and intersectoral inequality indices. LITRT, PCAREA, PCDEXP: literacy rate, net sown area (hectares) for all crops per capita, and state development expenditure (in Rs) per capita, respectively. GRPCY, GRPCAGY, GRPCMFGY: growth of total output, agricultural SDP, and manufacturing SDP per capita, respectively. PCSDP, PCAGY, PCMFGY: total output, agricultural and manufacturing SDP per capita (in Rs), respectively.
given in column 2 of the table. All estimates are corrected for the presence of heteroskedasticity.

We compare the LSDV generalized least-squares estimates (column 2 of Table 3) with corresponding single cross-sectional OLS estimates with the parsimonious specification shown in Table 2 (column 3). As before, there is evidence of convergence of total output among the Indian states as the coefficient of initial state-level output per capita (PCSDP) is negative and significant though the absolute value of the coefficient is now smaller. Among various capital variables, only LITRT still significantly enhances growth per capita, while the coefficients of PCDEXP and PCAREA are now insignificant (and hence dropped from the parsimonious specification). The insignificance of the two capital terms could be attributed to the limited sample size.

Significant differences arise with respect to the impact of inequality on growth: while in single cross-sectional estimates the coefficients of both RGINI and RRUMEAN were negative and significant, in the LSDV model the coefficient of RGINI is negative and that of RRUMEAN is positive, both coefficients being significant. Thus we obtain consistent LSDV estimates such that lower rural or intersectoral inequality enhances economic growth per capita; however, urban inequality still fails to have a significant impact on subsequent growth in the panel-data analysis.

In view of the dichotomous effect of rural and urban inequality on economic growth, finally we estimate the rate of growth of agricultural and manufacturing output per capita in terms of rural, urban, and intersectoral inequality, respectively, among other things. As before, the final specification as summarized in columns 4 and 5 of Table 3 is obtained by deleting the insignificant capital variables. These estimates are also corrected for the presence of heteroskedasticity. In this case, however, only the intersectoral inequality term is significant for both rural and urban sectors, and the coefficient is positive as in the case of total output growth per capita (column 2 of Table 3). While the coefficient of RGINI is still negative, this absolute rural inequality does not have any perceptible impact on growth in either sector. In other words, these results indicate the importance of sectoral interlinkages in explaining sectoral growth.

Taken together, our results indicate that higher initial rural inequality lowers subsequent growth of total output while urban inequality does not have any perceptible impact. One plausible explanation of this result is that about 70% of the Indian population still lives in rural areas so that a popular government is urged to respond more to rural (absolute and/or relative) inequality by undertaking various tax-financed redistributive programs. In addition, we find that the indicator of the intersectoral inequality component is more important than urban inequality, especially when we consider sectoral growth regressions. Thus compared to most existing studies like by AR (1994), PT (1994) or Deininger and Squire (1998), we derive a richer set of results. In addition to the decomposition of total inequality into sectoral inequality components and the use of the panel-data technique, our results may also be influenced by the fact that we use consumption inequality instead of income inequality. Since the extent of consumption inequality is generally less than that of income inequality, it is expected that the optimum tax of the median voter would be less than in the case of income inequality, thus lowering the disincentive effect of tax on investment and subsequent output growth.

5. Conclusion

In terms of an endogenous growth framework this paper has examined how initial inequality affects economic growth in the ensuing period, where growth of the regional
economy is driven by productive public investment financed by linear output taxation. It is suggested that initial inequality in the distribution of income leads to the optimum rate of taxation (determined by the median voter) being different from the rate that maximizes the economy’s growth rate. However, the precise relationship remains ambiguous and depends on the net effect of the output tax on labor and capital income of the median voter. Empirical estimates from the Indian states suggest that rural inequality is more important to explain growth of total output per capita and there is a negative relationship between the two. It is also clear that the indicator of intersectoral inequality is quite important in explaining sectoral output growth.

References
Notes

1. While according to Kuznets’ (1955) inverted-U hypothesis, growth causes higher or lower inequality depending on the level of development, the direction of this causality has been reversed in the recent endogenous growth literature.
2. There have also been studies that focus on growth with nonpolitical considerations of redistribution. For example, see Galor and Zeira (1993).
3. This is clear from simulations with different values of $\tau$, and using parameter values consistent with the Indian case.
4. Note from section 2 that for the case where a $\mu$-proportion of tax revenues is used as redistributive transfers to the labor component of income, this is what we would expect.
5. In the single cross-sectional analysis, initial inequality is considered to be predetermined relative to growth over the next 35 years; thus any direct causation from growth to inequality is ruled out.
6. Consumption is well insulated from transitory movements and can thus help us focus on the pure long-run component of income inequality (Blundell and Preston, 1998).
7. We thank a referee for suggesting this measure.