State Owned Enterprises and Redistribution: An Empirical Analysis*

Cem Karayalcin† and Mehmet Ali Ulubasoglu‡

Abstract

In the past decade many developing economies started to privatize their state owned enterprises. Recently, however, this process seems to have slowed down in some economies and have completely been stalled in others. Here we formalize the view that this is so because these enterprises are major instruments of income redistribution and, in economies with significant degrees of income inequality, segments of the population that benefit from this redistribution would use whatever political power they may have to oppose its abandonment. We find strong and robust empirical support for this hypothesis using cross-country data on the relative size of the state-owned-enterprise sector and different measures of inequality. We also find support for the propositions that dictatorships as well as democracies use this redistributive tool and that left-wing governments tend to redistribute more than right-wing governments through state owned enterprises.

*We would like to thank Devashish Mitra, Dimitrios Thomakos and Dimitrios Doucouliagos for helpful comments on earlier drafts.

† Department of Economics, Florida International University, Miami, FL 33156.

‡ Department of Economics, Deakin University, Australia

1 Introduction

Recent years have witnessed a widespread attempt in developing countries to dismantle one of their most entrenched institutions-the state owned enterprises (SOE's). Some of these attempts (for instance in Argentina) have been quite successful, others (for instance in Turkey) much less so. This paper tries to formalize, in the simplest analytical terms possible, a popular, but as yet informal, argument as to why some countries find it very difficult to privatize their SOE's and to test this argument empirically. Simply put, the argument is that the SOE sector has become a major instrument of income redistribution especially for countries undergoing a taxing structural adjustment. In the words of an acute observer writing about Turkey, "...the privatization drive...has lost its attractiveness to the extent that it would impede the state from using the SOE's to ease the pain of other components of the structural adjustment process."[1]

The SOE's owe their genesis to the adoption of strategies of import-substituting industrialization.[2] In some cases these policies had been implemented quite early. For instance, in Turkey the SOE sector dates back to the late 1920's, in Mexico to the 1930's, in India to the late 1940's. In most cases the common ostensible rationale for establishing SOE's was that the private sector in existence was weak, unable to compete with foreign goods or prone to the formation of alliances with foreign capital to the detriment of national interests. It soon became clear to economists and policy makers alike that the performance of the SOE sector left something to be desired: SOE's experienced chronic losses which resulted in rising domestic budget deficits and inflation. The response was attempts at rationalizing and streamlining the SOE sector. This soon proved impossible.

In order to explain why such attempts were doomed to fail from the start, we will focus in what follows on two main factors that contribute to the losses commonly registered in the SOE's:

[1] See Waterbury (1992, p.194).

[2] The genesis of the SOE sector may in addition reflect the need for income redistribution. Waterbury (1993, p. 263) stresses that "...the political logic that gave rise to the SOE sectors in the first place [was] the need to redistribute income..."

high wages to SOE employees and "surplus labor".[3]

First, a clarification. When we say the SOE's pay *high wages* we mean that the SOE's typically pay wages that are higher than those paid by private enterprises in a given country. It is frequently the case that SOE's are monopolies and labor unions negotiate high wages without fearing a depression in wages caused by attraction of labor to SOE's from competing private firms. The SOE's may also pay a compensating wage differential when they operate in locations where private firms may be reluctant to locate. The full compensation package of the SOE's may include superior leave privileges and retirement benefits. Furthermore, even if wage rates in the SOE's are similar to the ones offered by private firms, given the low productivity epidemic in the former, the ratio of wages to marginal productivity of labor is higher. Finally, there is strong empirical evidence from Latin America, Africa, and Southeast Asia that supports the observance that the SOE's pay high wages.[4]

Second, casual empiricism as well as careful empirical studies suggest that the SOE's carry "surplus labor", that is, they employ more workers than their operations would justify on strictly rational economic grounds. Thus, for instance, an official study found that though the output of the SOE's in the Western and Mid-Western states of Nigeria remained unchanged in the period 1963-1967, the wages and salaries' bill more than doubled.[5] Complaints by management of surplus workers in the SOE's in Sri Lanka, Trinidad and Tobago, India, and Britain, inter alia, are well documented.[6]

It is, therefore, not surprising that the SOE's suffer from chronic losses given the wages they pay and the surplus labor with which they operate. And, they do so because they are instruments of income redistribution. Furthermore, this redistributive tool seems not to be shunned by governments either on the left or on the right (though as we show below not to the same extent),

[3] The following discussion relies heavily on Ramanadham (1988).

[4] For the evidence see Ramanadham (1988), ch.2.

[5] See the Report of the Conference on Public Enterprises in Nigeria (1970), cited in Ramanadham (1988).

[6] See Ramanadham (1988).

by governments democratic or dictatorial. Thus, for instance in Bolivia ruled by the left-wing MNR (Movimiento Nacionalista Revolucionaria) "[b]y the early 1960s a form of state capitalism developed, controlled and exploited by various competing groups of the middle classes....[T]he state enterprises became a source of enrichment for these private factions, some civilian and some military."[7] Under the right-wing rule of General Hugo Banzer, who was installed as president of Bolivia following a coup d'etat in August 1971, "...the public enterprises served frequently as a mechanism to transfer state-owned (or state-guaranteed) resources to privileged groups in the private sector. Access to government officials and government contracts were considered the most important asset from the viewpoint of many private-sector businessmen."[8] Further, "[i]n fact, a non-negligible part of the support for the Banzer government and succeeding military regimes was the willingness to create employment in the public sector. The return to democracy in 1982 was also accompanied by a big spurt in the expansion of jobs in the most important public enterprises, particularly in COMIBOL."[9] When in 1970 the Mexican president Diaz Ordaz had to choose his successor unilaterally, the new president "...Echeverria faced the difficult task of creating his own supporting coalition after assuming office. The simplest method of shoring up the weakening political consensus was to spend on everyone's behalf: dole out subsidies to education and agriculture, increase government jobs for the middle classes, grant large wage increases to mollify organized labor, etc.....Between 1970 and 1976, the number of federal government employees doubled and the growth rate of general government employment averaged 10.8 percent. A series of large wage hikes after 1972 further inflated the government wage bill."[10]

In what follows we thus take it for granted that the SOE's are used to redistribute income.[11]

[7] See Morales and Sachs (1989), p.180.

[8] See Morales and Sachs (1989), pp.192-193.

[9] Ibid. p.197.

[10] See Buffie (1989), p. 420.

[11] Here we do not deal with the question as to why they are used as a tool of redistribution when there are more efficient tools. One possible answer to this question is that the lack of transparency in generating redistribution through nonmonetary transfers makes SOE's a politically efficient tool. See Coate and Morris (1995) for a formal model that shows that politicians would prefer to redistribute via public works rather than cash transfers when

The question that remains to be answered then is: Under what conditions are they an acceptable means of income redistribution? In what follows we set up a formal model that yields an answer to that question, namely, that as long as the median wealth is less than the average wealth the majority of the population will prefer to establish an SOE sector to redistribute income. A corollary to this theoretical finding is that the more unequal the distribution of wealth (as measured by the difference between median and average wealth) the more extensive will be the SOE sector. In a democratic setting with majority voting this implies, under certain conditions, that we can invoke the median-voter theorem to suggest that an SOE sector would be established for redistributive purposes.[12] Otherwise, we surmise that the preferences of the majority would, through channels we leave unspecified, find its expression in policies that favor redistribution through the SOE's. These channels may take the form of a desire on the part of dictators (as in the case of Banzer in Bolivia) or autocrats (as in the case of Echeverria in Mexico) to find popular support for their rule.

We also empirically test the hypothesis formulated in the theoretical model. We employ three measures of inequality to study its effect on the size of the SOE sector. One measure of inequality we use is the land gini. This is a measure that comes closest to the theoretical measure in that it is a measure of wealth inequality. A second measure we use is the share of the median quintile of the population in national income. This measure comes closest to representing the share of the median voter in that interpretation of our formulation. The third measure is the more frequently used measure of income inequality, namely the income gini. We measure the relative size of the SOE sector in three different ways as well: the share of the SOE's in GDP, their share in nonagricultural GDP, and their share in total employment. Across most measures we find strong empirical support in favor of our hypothesis that an increase in inequality is associated with a bigger SOE sector. We also show that this conclusion is extremely robust with respect to the controls used.

voters lack information.

[12] The median voter approach has been the target of empirical inquiry in a wide variety of contexts. For a recent study that finds strong empirical support for the median voter approach to trade policy determination, see Dutt and Mitra (2002a).

In addition, we test whether democracies are more prone than dictatorships to redistribute income through the use of the SOE sector. We find no evidence that this is the case and confirm the Alesina and Rodrik (1994) finding that when it comes to distributional issues even dictators bow to popular will. We are also interested in probing the role of political ideology in our context and test the hypothesis that left-wing governments redistribute more than the right-wing ones and find strong and robust support for it. Though both types of governments do use the SOE sector as a redistributive tool, the tendency of the left-wing governments to do so is more pronounced. This should come as no surprise since the constituencies of left-wing governments would favor such redistribution more strongly than those of right-wing governments.[13]

It is also worth noting that our empirical findings are also consistent with the recent literature that emphasizes the concept of common property and the attempt by different groups in societies to appropriate the common property.[14] These models are said to apply to societies where there is "extreme inequality". If the resources of the SOE's are viewed as common property by the "various competing groups of the middle classes" as suggested in a quote above then our findings should be interpreted as also giving empirical support to the common property notion and the models built to elucidate it.

To sum up the contribution of our paper to the literature is twofold. First it sets up a simple theoretical framework to establish formally the idea that SOE's will be used to redistribute income when inequality is sufficiently pronounced. Secondly, the paper contributes to the empirical literature on redistribution with and without democracy, on the role of political ideology and on state owned enterprises.

In section 2, we set up the theoretical model. Section 3 describes the empirical testing, while the last section provides some concluding remarks.

[13] Dutt and Mitra (2002b) also find strong empirical support for the hypothesis that left-wing governments tend to redistribute more *via* trade policy than right-wing governments.

[14] See Tornell and Velasco (1992) and Benhabib and Rustichini (1996) among others.

2 Model

In this section we formulate the simplest possible model that conveys the proposition that the SOE's may be used as a redistributive mechanism. To do so consider an economy populated by L households all endowed with a single unit of labor and varying amounts of capital. The economy may potentially produce three goods in three different sectors. For analytical simplicity, one could either imagine the economy under consideration to be closed-in which case we will assume that the goods are perfect substitutes in consumption with their relative price fixed at unity-or that it is a small open economy-in which case the parametrically given relative prices are again normalized to unity with an appropriate choice of units. We now turn to a detailed discussion of the production side of the model.

2.1 Production

The first of the three sectors will be called the "formal sector". This sector functions as the "modern", industrialized sector in the model. It uses capital and labor to produce a consumption good under a constant-returns-to-scale technology in a perfectly competitive market. The output of this good is given by

$$Q_p = F(K, L_p) \tag{1}$$

where the production function $F(\cdot, \cdot)$ possesses the usual neo-classical properties and K and L_p denote the capital (physical and/or human) and labor employed in the sector under consideration.[15]

The second sector is labeled the "informal sector". Agents employed in this perfectly competitive sector have access to a Ricardian technology with a constant input-output coefficient $1/\alpha$ and produce a consumption good using labor alone. The labelling of this sector is motivated by

[15] Since the formal sector is the only sector that employs capital (as will be seen below) K also denotes the total capital stock of the economy.

the observation that in developing countries (as well as some "developed" countries, such as Spain or Southern Italy) agents who are not employed in either the "modern" private sector or by the SOE's find employment (or are considered officially unemployed) in an informal sector, of which street peddling and Mariachi bands are the most picturesque examples.

Finally, the public sector may also employ labor, L_g, in SOE's to produce the same consumption good (or a perfectly substitutable good) with the same Ricardian technology that the informal sector uses.[16] The SOE's may pay a wage, w_g, higher than the marginal productivity of labor employed. If this is the case, the losses, $(w_g - \alpha)l_g$, (where lowercase letters denote per-capita variables) of the SOE's, are financed by the revenues of a proportional income tax, τ, imposed on the factors employed by the formal sector.[17] The government budget constraint is, thus, given by

$$(w_g - \alpha)l_g = \tau q_p \tag{2}$$

where the right hand-side of (2) denotes the outlays of the government and the left hand-side its tax revenue.

Now, competition from workers in the informal sector ensures that the after-tax wage rate across these private sectors is equalized.[18] Thus, given the proportional tax rate, profit maximization by firms in the formal sector implies that the rate of return on capital, r, and the level of employment in this sector depend negatively on the tax rate:

[16] In reality, of course, the SOE's may have (and typically, many of them do have) access to technologies potentially more capital-intensive than the "modern" private sector. Three considerations motivate the modeling choice made in the paper. First, as argued in the Introduction, we observe that SOE's normally employ "surplus labor" that renders their technology labor-intensive. Second, the actual pricing of the capital input in the SOE's is such that the shadow rentals to be attributed to the capital they own is typically below market value. Finally, the technology modeled in the paper brings out very sharply the potentially superfluous nature of the SOE's from a strictly economic (as opposed to political) point of view.

[17] The informal sector is in practice very difficult to tax, thus any revenue that governments actually raise by taxing this sector is small enough to be safely ignored in the present model.

[18] Though the formal sector (because of market imperfections) typically pays a wage above that which can be earned in the informal sector, that wage is depressed by competition from the pool of workers that find themselves in the latter sector. The result here should, therefore, be interpreted as a version (at the margin) of the observed outcome.

$$l_p = l(\tau), \quad l'(\tau) < 0, \tag{3}$$

$$r = r(\tau), \quad r'(\tau) < 0. \tag{4}$$

Intuitively, a higher tax rate lowers the employment level by reducing the after-tax marginal productivity of labor in the formal sector. Decreased employment, in turn, diminishes the marginal productivity of capital.[19]

2.1.1 Consumption

Turning to the consumption decisions of the households note that they supply labor and capital in competitive markets. Each household is assumed to supply inelastically the unit of labor with which it is endowed. However, households differ with respect to their capital endowments $k_i \geq 0$ $(i = 1, .., L)$.

Given the static nature of the model and the fact that there is, effectively, one composite good to be consumed, the consumption decisions of households are quite simple-each household, facing the parametric tax rate and factor prices uses its wage and rental income (received in exchange for the services of labor and capital supplied) to consume this composite good.

2.1.2 Political Economy

The environment within which we work is now endowed with sufficient structure to answer the principal question we are interested in: Under what conditions will this economy choose to operate a state owned enterprise described above?

To answer this question, however, we need to prescribe a rule which governs the political decision process. The simplest and most frequently used rule in the literature is the majority rule provided that decisions are made in a democracy, the agenda consists of a single item, and voters' preferences are single peaked. If this is the case, it is the preferences of the median voter we need

[19] For explicit expressions for the derivatives in these and following equations see Appendix II.

to consult to see if s/he would choose to operate the SOE under question. Even in the absence of democracy, the median voter's preference yields, at the very least, some measure of the strength of support for the SOE's that the rulers will need to pay attention to.[20]

Now, since the consumer consumes all of his income in this static world, the determination of the tax rate preferred by a household i can be formulated as the solution to the maximization problem faced by this household of (expected) indirect utility

$$U = l_g u[(l-\tau)rk_i + w_g] + (1-l_g)u[(l-\tau)rk_i + \alpha] \tag{5}$$

(where l_g, the ratio of employment in the SOE to total employment, also denotes, from the household's point of view, the probability of being employed by the SOE and, thus receiving w_g; the household will otherwise be employed by the (formal or informal) private sector and earn the wage α) subject to the government budget constraint (2) with

$$(w_g - \alpha)l_g = \tau q_p, \quad \tau q_p \geq 0 \tag{6}$$

The first-order conditions for the problem are

$$\frac{l_g u'(c_g) + (1-l_g)u'(c_p)}{u'(c_g) + \sigma} = \frac{\beta - \tau}{\beta(1-\tau)} \frac{k}{k_i} \tag{7}$$

$$u(c_g) - u(c_p) = u'(c_g)(w_g - \alpha) \tag{8}$$

$$\sigma \tau q_p = 0 \tag{9}$$

where σ is the multiplier associated with the inequality constraint and β denotes the income share of capital in the formal sector.

Proposition 1 *Agent i preferes no taxation ($\tau = 0$) and, thus, no state state owned enterprises if and only if his capital endowment k_i exceeds the average capital endowment k.*

[20] It is well understood that even dictators need some minimal popular support for their survival.

Proof. The proof proceeds in two steps. The first step shows that $\tau = 0$ implies $k_i > k$. The second step shows the reverse implication.

1. Suppose $\tau = 0$. This implies $c_g = c_p = c$, $l_g = 0$, and $\sigma > 0$. The first order condition then becomes $u'(c)/[u'(c) + \sigma] = k/k_i$. Since the lefthand-side of this equation is less than 1, we have $k < k_i$.

2. Suppose $k < k_i$. We will show that if $\tau > 0$ this would violate of the first-order condition. If $\tau > 0$ then $\sigma = 0$. Then the lefthand-side of the equation in the first-order condition becomes $[l_g u'(c_g) + (1 - l_g)u'(c_p)]/[u'(c_g)]$. It is straightforward to show that this expression is greater than 1, because in this case $c_g > c_p$ and $u''(\cdot) < 0$. However, with $k < k_i$ and $\beta < 1$ the expression on the righthand-side of the first-order condition is less than 1. Thus, the first-order condition is violated. ■

Proposition 2 *Ceteris paribus agent i benefits more from a higher tax and, thus, a larger state-owned enterprise sector the smaller is his capital endowment k_i relative to the average capital endowment k.*

Proof. It is straightforward to obtain the following

$$\frac{dU_i}{d\tau} > 0 \Longleftrightarrow \frac{(\beta - \tau)}{\beta(1 - \tau)} \frac{u'(c_g)}{l_g u'(c_g) + (1 - l_g)u'(c_p)} > \frac{k_i}{k} \tag{10}$$

The smaller is k_i relative to the average capital endowment k the more likely it is for this inequality to hold. Note that for $k_i = 0$ this inequality implies a welfare-maximizing level of tax $\tilde{\tau}$ such that $\tilde{\tau} = \beta$. this is the tax rate that maximizes the tax revenue of the government allowing for the maximum size of the state-owned enterprise sector. ■

This proposition can be taken to imply that the lower is the median agent's capital endowment relative to that of the average agent, the bigger will be the size of the SOE sector.[21] To see why start with some tax rate. Will the median agent have a higher utility if the tax rate is raised? Or, to put it differently, would the median agent prefer a higher tax rate? The answer

[21] The median agent is defined here to mean the agent that owns the median capital stock.

is more likely to be in the affirmative the smaller is the median agent's endowment relative to the endowment of the average agent. Intuitively speaking, this result belongs to a class of results obtained in the political economy literature that redistributive instruments will be preferred by a median voter whose wealth (or income) lies below that of the average agent. It differs from the existing literature in the form of the redistributive instrument (which is generally a lump-sum transfer payment[22]). But our result can also be interpreted to imply simply that (i) the majority of the population would prefer to have an SOE sector as long as the the median agent's capital endowment is smaller than that of the average agent, and (ii) this majority would be a larger fraction of the population the lower is median agent's capital endowment relative to that of the average agent. Under such an interpretation we would leave the nexus between the preferences of the majority and the ultimate political decision unspecified, and yet point out to the needs of even autocrats and dictators for popular support to perpetuate their rule. This need seems to have motivated rulers like Banzer of Bolivia and Echeverria of Mexico to use the SOE's to redistribute income to obtain the support required.

3 Econometric Specification and Methodology

The theory outlined above predicts that the more unequal the distribution of wealth is, the more likely it is for a country to operate a SOE sector as a redistributive tool. To test this prediction we use cross-country regressions of the following type

$$RSOE_i = \alpha_1 + \alpha_2 INEQ_i + \alpha_3 PCAP_i + \varepsilon_i \tag{11}$$

where $RSOE$ is an indicator of the relative size of the SOE's in economic activity, $INEQ$ denotes a measure of wealth inequality and $PCAP$ is the per-capita 1960 income of country i. This last variable is included in the equation as a control variable to check whether the initial level of development of an economy affects the use of SOE's as a redistributive instrument. Our expectation

[22] See, for example, Persson and Tabellini (1992) and Alesina and Rodrik (1992).

is that rich countries would tend to use the SOE's less as a tool for redistribution since they may have access to more efficient means of redistributing income. Thus, we expect the sign of the coefficient for per capita income to be negative.

It is possible to construct several indicators that measure the relative size of the SOE sector in overall economic activity in our sample. The share of the SOE's in GDP (denoted by $\frac{SOE}{GDP}$), their share in nonagricultural GDP (denoted by $\frac{SOE}{NONAG}$) and their share in total employment (denoted by $\frac{SOE\_EMP}{EMP}$) are three potential measures. We use the average of the period 1978-1991 for these variables.

In our main regressions we use both asset (wealth) distribution and income distribution variables on the right-hand side. Land gini in particular is used as the indicator of the wealth inequality. We expect a positive sign for its coefficient, that is, as the wealth distribution becomes more unequal (represented by an increase in land gini), we expect the share of SOE's in total GDP to increase. To check the robustness of our results, we also use indices of income inequality, namely the share of third quintile in income distribution (henceforth referred to as $QUIN$) and the income gini coefficient (denoted by $GINI$). It should be clear that for our purposes the share of third quintile is a better indicator of income inequality than the income gini, because the former is closer to the share of median voter, whereas the latter is a broader interpretation of income inequality. In those regressions where we use the share of the third quintile in income distribution we expect its coefficient to have a negative sign because this is a measure of reverse inequality. Finally, we of course expect the sign of the coefficient of the income gini variable to be positive.

The theory does not suggest any reverse causation running from the relative size of the SOE's in the economy to wealth distribution, since the wealth distribution is exogenous to the model. However, the size of the SOE sector may affect income inequality. Specifically, in countries where income inequality is more pronounced we expect the SOE's to be used to redistribute income, thereby reducing income inequality. Hence, both the $QUIN$ and the $GINI$ variables may be endogenous. It is also possible that the $PCAP$ variable may suffer from the endogeneity problem

to the extent that the availability of other politically acceptable means of redistribution that are more efficient for this purpose than the SOE's may itself lead to higher per-capita income. One can argue that the regression of SOE-related variables of the period 1978-1991 on per-capita incomes in 1960, i.e., a time lag between dependent and independent variables, may rule out the problem. However, if the variables exhibit stickiness, the problem may persist.

In view of these potential problems, we use weighted two-stages least squares (W2SLS) to address the heteroscedasticity and endogeneity problems. Furthermore, to check the endogeneity of the variables in question we use the test suggested by Hausman (1978), which involves comparing the estimates from the OLS and an instrumental variable (IV) method to see whether these are significantly different from each other. As for heteroscedasticity, we first do a visual inspection: we plot the residuals from preliminary (unweighted) OLS regressions against the per capita income variable, and check whether the dispersion of residual variances differ along per capita income, that is, if they are clustered into groups. The presence of this would suggest the existence of the heteroscedasticity problem. Formally, we test the heteroscedasticity by modeling the error variance as a function of the size of the economies. The details of these tests are given in the section 5.

As for the instruments, a small number of exogenous variables is sufficient to get sensible results. Our list includes the following variables: the land gini, an indicator of democracy, dummies for lower-middle and upper-middle income countries. All of these instruments are used consistently across our regressions.[23] In addition to this, we need to use an exogenous weight to take care of the heteroscedasticity problem. We choose Nehru-Dhareswar capital-labor ratio (henceforth referred to as ND) for this purpose since it can approximate the size of the economy.[24] ,[25]

[23] Since the number of variables in our instruments matrix is greater than the number of independent variables, our equations are over-identified. We also perform the corresponding over-identification test, and the null hypothesis is accepted in every case.

[24] As the variable that is causing the heteroskedasticity problem, $PCAP$ suggests itself as a weight. However, since we are using the two-stages least squares method to take care of the endogeneity problem, we cannot use this endogenous variable as the weight, which, otherwise, will be correlated with the error term.

[25] The correlation between per capita income and ND capital-labor data in our sample is 0.77. We prefer to

4 Data

In this section we provide a brief summary of the data and their sources. Our dependent variables that measure the relative size of SOE's in each country are the share of the SOE sector as percentage share of (1) GDP, (2) nonagricultural GDP and (3) total employment. This data comes from *Bureaucrats in Business: The Economics and Politics of Government Ownership, (*a World Bank Policy Research Report).

We obtain the data on land gini and income gini from Deininger and Squire (1996).[26] The data on the share of third quintile in income distribution averaged for 1980's were obtained from World Development Indicators (1999). Nehru-Dhareshwar capital-labor ratio data are from Nehru and Dhareshwar (1993).[27]

The data on democracy comes from the Easterly-Levine dataset. The scores on democracy in this dataset are very close to the political rights index of Gastil, but the country coverage of Easterly-Levine dataset is broader. We used the democracy scores for the year 1989. The Easterly-Levine index is such that more democratic countries are assigned a lower score than less democratic countries (the scores run from 1 to 7). We reversed the scores by substracting each score from 8 so that more democratic countries are assigned higher scores.

The data on per capita income and the dummies for lower-middle and upper-middle income countries are obtained from *Social Indicators and Fixed Factors* dataset of the World Bank. The per capita income variable is scaled by 1/1000.

proceed with ND dataset since its country coverage is broader than the other capital-labor ratio data sets on in the literature: that of Summers-Heston (SH) and that of Easterly-Levine (EL). We also exercise with these datasets. While our results do not change qualitatively when we use them, the coefficients become marginally insignificant due to the loss of observations.

[26] Specifically, we use the adjusted version of Deininger and Squire dataset by Dutt and Mitra (2002a). They add 6.6 points to the income gini values to adjust for the differences between income-based data and expenditure-based coefficients.

[27] The figures are in 1987 local prices. We first convert them into US dollars with 1987 exchange rate of each country, and then divide them by the number of population between the ages 15 and 64, the data of which were obtained again from Nehru and Dhareshwar (1993). Afterwards, for comparability with the Summers-Heston and Easterly-Levine capital-labor ratio data, which are in 1985 dollars, we adjusted the data to the 1985 prices by applying the formula $(K/L) * P85/P87$ where $P85$ and $P87$ are the GDP deflators of each country taken from World Development Indicators (1999). In proceeding, we deleted Mexico and Uruguay from the data set as they had extraordinarily high figures.

5 Results

In Table 1 we provide the estimation results when we use the land gini as a measure of wealth distribution.[28] When we use $\frac{SOE}{GDP}$ and $\frac{SOE}{NONAG}$ as our measures of the relative size of the SOE sector, our estimates for coefficients of the land gini are significant at 1% level with correct positive signs.[29] The coefficients are 0.0013 and 0.0017, respectively (to interpret these coefficients, it should be kept in mind that both land gini and our measures of the relative SOE size are in percentages, implying, for instance, that a 1% increase in wealth inequality leads to a 0.13% increase in the size of the SOE sector relative to GDP). For these regressions, $PCAP$ is negative and significant. The number of observations included in the regressions is around 42-43. The models are significant at 1% and 5% level, respectively, as shown by the F-statistic.[30] For the regression with dependent variable $\frac{SOE\_EMP}{EMP}$, we fail to obtain significant estimates, possibly due to a very small number (25 to be precise) of countries.

Table 2 presents the estimation results when we use the share of the third quintile $QUIN$ as an indicator of the inequality of income distribution. With the dependent variables $\frac{SOE}{GDP}$ and $\frac{SOE}{NONAG}$, the coefficients of this variable are negative and significant. These coefficients are -0.0218 (significant at 5%) and -0.0266 (significant at 5%), respectively. These results suggest that in countries where the share of the median quintile in income distribution is higher (more equal distribution), SOE sector is not used as extensively to redistribute income. In these regressions, the coefficient of the $PCAP$ variable is not significant while its sign is negative for $\frac{SOE}{GDP}$, and positive for $\frac{SOE}{NONAG}$. The number of observations are around 35-36. F-statistics show that both models are significant at 5%. As for the dependent variable SOE_EMP, the coefficient of $QUIN$

[28] The equation numbers that appear in the tables refer to equations listed in appendix 1.

[29] The estimation methodolgies indicated in the tables reflect the fact that if our tests indicated heteroscedasticity and endogeneity we run W2SLS, if the tests indicated only heteroscedasticity we run WLS, otherwise we run OLS.

[30] R^2s are purely descriptive in the generalized regression context and does not indicate the fit. The models become different once they are transformed (weighted). In other words, the dependent variable in the transformed regression $y_* = \mathbf{X}_*\boldsymbol{\beta} + \boldsymbol{\varepsilon}_*$ is different than the one in untransformed (original) regression, $y = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. For example, in our case, the new dependent variable is $y_* = y/(K/L)$.

is insignificant. However, that of $PCAP$ is negative and significant at 5%. This would imply that in countries where initial per capita income is higher, the shares of SOE's in employment is lower. This version of the model is significant overall at 10% level, with, again, a low number (25) of observations.

Table 2 further provides the estimation results when we use the income gini ($GINI$) as a measure of income distribution. The coefficient of this variable is also significant and displays the predicted (positive) sign. For the dependent variables $\frac{SOE}{GDP}$ and $\frac{SOE}{NONAG}$, the coefficients of $GINI$ are positive and significant at 1% level, 0.0084 and 0.0112, respectively. This suggests that the more unequal the income distribution in a country is, the more pronounced is the role of the SOE's as redistributive instruments. The $PCAP$ variable is negative and significant at 1% level for both of the models. With the number of observations around 36-37, the models are overall significant at 1%. We also have the SOE_EMP variable with a significant estimate for $GINI$. The estimate is 0.0047 and significant at 10%. The intuition is that as wealth becomes more evenly (unevenly) distributed, the share of SOEs in employment diminishes (increases).

5.1 Checking for Robustness

5.1.1 Hausman Tests

As mentioned above, we suspect that the variables $QUIN$, $GINI$ and $PCAP$ may be endogenous. For this reason we carried out the Hausman Test to check whether the regressors are correlated with the error term. The classical linear model assumes $E(X\varepsilon) = 0$, that is independence of the vector of regressors, X, from the vector of disturbances, ε. Violation of this assumption leads to biased and inconsistent OLS estimates. The results for the Hausmann tests we ran are given in Table 3.

For each of the suspected variables, we run auxiliary regressions in which we used land gini, the democracy score, dummies for lower-middle and upper-middle income countries as right-hand side variables. In each of these reduced form equations, we observed great significance levels for

the exogenous variables.

After saving the residuals from each auxiliary regression, we inserted them into the corresponding original regressions, and looked at their significance levels. The Hausman test predicts that if the coefficients of these residuals in the original regressions are significant, then the OLS estimates are inconsistent, and the use of an instrumental variable method is in order. However, if they are insignificant, OLS estimates are consistent and identical to IV estimates.

The series of tests for the suspected variables show that the $PCAP$ variable is endogenous when the dependent variables are $\frac{SOE}{GDP}$ and $\frac{SOE}{NONAG}$ while the measure of wealth inequality is the land gini. Therefore, the use of the IV method is appropriate. For the equation in which the dependent variable is $\frac{SOE\_EMP}{EMP}$, the Hausman test provides no evidence for endogeneity.[31]

Similarly, when the dependent variables are $\frac{SOE}{GDP}$ and $\frac{SOE}{NONAG}$ while the measure of inequality is $QUIN$, the $PCAP$ and $QUIN$ variables are jointly endogenous to the models in question.[32] The OLS estimates would have been inconsistent had we employed that method. When we use $\frac{SOE\_EMP}{EMP}$ as the dependent variable in this case, the variables $PCAP$ and $QUIN$ do not suffer from endogeneity jointly.

When $GINI$ is used as the measure of income distribution, the Hausman test does not provide any evidence for the joint significance of the auxiliary $PCAP$ and $GINI$ variables. Hence, OLS would provide consistent estimates.

5.1.2 Modeling Heteroscedasticity

As mentioned above, the $PCAP$ variable may create a heteroscedasticity problem due to clustering of developing and developed countries. Failure to correct this problem will lead to the violation of one of the assumptions of the classical linear model; that is, $E(\varepsilon\varepsilon') \neq \sigma^2 I$ any more, meaning

[31] For those regressions in which we cannot detect these problems, we report, without loss of generality, the results of OLS in our appendix.

[32] In this sort of multivariate Hausman test, it is the joint significance of the variables that matters for their endogeneity, rather than their individual significance. Therefore, we test if these first-step residuals in the original regressions are jointly significant to check their endogeneity. The results are given in the bottom part of Table 3.2

that individual disturbance terms do not have uniform variance with the consequence that our estimates would be inefficient.

The first step in the investigation of this problem is the plotting of the residuals from a set of preliminary regressions against per capita income with which we suspect the disturbance variance is related. In seven of the nine regressions, we see a clustering of the developing and developed countries in the scatter plots, with low level of $PCAP$ having high variances and vice versa. This suggests that the size of the economy matters in explaining the share of the SOE's in the economy.

In addition, we model the log of our error variance as a function of the log of capital-labor ratio, an exogenous variable that indicates the size of the economy.[33] ,[34] The choice of functional form is such that we first try different variance specifications to see what the error variances in the regressions are proportional to, and these trials indicate that they are proportional to the square of the capital-labor ratio. Formally, we find that, in a model $y_i = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i$ with assumptions $i = 1, 2, .., n,\ E[\varepsilon_i] = 0,$ the variance specification $E[\varepsilon_i^2] = \sigma_i^2 = \sigma^2 k_i^2$ is the most significant (where k_i denotes the K/L ratio). Whether there is heteroscedasticity or not is empirically answerable by looking at the estimate of θ, i.e., if it is significantly different from zero in the model $ln(\hat{\varepsilon}_i^2) = \delta + \theta \ln(k_i^2) + v_i$.

Estimating this model in our context is as follows: after running an unweighted regression for each equation, we saved the corresponding residuals. Then we regressed the log of these squared residuals on a constant and on the log of the squared capital-labor ratio. A finding of $\widehat{\theta} < 0$ and significant would confirm our hypothesis that the error variance is higher for the first group of countries than for the second group, bearing out the clustering conjecture. In fact, for eight of the

[33] Most heteroscedasticity tests such as Goldfeld-Quandt (1965) and White (1980) are quite general in that they do not take into account the relationship between the error variance and the variables suspected to determine that variance. As we are aware of the source of the problem (clustering of per capita income variable of developed and developing countries), we preferred to model our error variance. Additionally, Green (1997) states that tests based on modeling heteroscedasticity are likely to be more powerful than an omnibus test such as White's *in the specific context of its regression model.*

[34] Like any variable in economics, the endogeneity of the capital-labor ratio can be argued, but both capital and labor are taken fixed in our theoretical model.

nine regressions, we find a negative relationship between the residuals and the capital-labor ratio, showing that the higher the capital-labor ratio is (the more developed a country is), the lower is the residual (the lower is the variance). For the equations (R1), (R2) and (R8), these relationships were significant at around 5% level, for the equation (R9) at 10%, and for (R4), (R5) and (R7) at some higher levels. Therefore, for these regressions the heteroscedasticity problem is formally diagnosed. The details of the results are provided in Table 4.

Having determined the existence of the heteroscedasticity problem, next we take steps to correct it. For this purpose, we weight the regressions with the Nehru-Dheshwar capital-labor ratio to account for the size of the economies. All the variables in the model, including the instruments where appropriate, are multiplied by the reciprocal of the capital-labor ratio, rendering the variances uniform. In other words, the diagnosis of the problem allows us to form our final transformed regression model in the following way:

$$\frac{y_i}{k_i} = \frac{\mathbf{x}_i'\boldsymbol{\beta}}{k_i} + \frac{\varepsilon_i}{k_i} \tag{12}$$

The correction of this problem improves the significance level of the (unweighted) regressions, having improved their efficiency. We retain the OLS for those regressions for which we could not confirm the problem formally.

5.1.3 Using $PCAP$ as Weight

For the regressions in which $PCAP$ does not exhibit an endogeneity, as shown by the tests, it is legitimate to check the relevant equations by using this variable as the weight. We can apply this robustness check to the equations (R3), (R6)-(R9). For the equations (R3), (R6) and (R9), heteroscedastic error model with $PCAP$ variable (instead of K/L) shows the existence of heteroscedasticity weakly. The application of the test to the equations (R7) and (R8) supports the existence of the problem at 10% and 1%, respectively. Running the original regressions with $PCAP$ as the weight does not make any difference in the coefficients. Hence, the use of ND

capital-labor ratio proves to be a good option.

5.1.4 Ramsey's RESET Test

In order to check for the existence of non-linearity, we employ Ramsey's RESET test. The test results show that the equations (R1) - (R6) and (R9) do not have any specification error, while the equations (R7) and (R8) may necessiate a change of the functional form. Converting these models into reciprocal models (reciprocal in $PCAP$ only) removes the specification error problem. Additionally, the results do not change qualitatively, i.e., signs of $GINI$ remain exactly as before.[35] This may imply some non-linearity in these models in that as the per capita income in a country increases, the share of SOEs in GDP and in nonagricultural GDP decrases at a decreasing rate.

5.1.5 Dictatorship vs Democracy

In the introduction we have argued that policymakers tend to be responsive to the concerns of the majority of the population whether they are democratically elected or not. One could however plausibly counter that there certainly might be a significant difference in degree if not in kind between dictatorial and democratic policymakers. To investigate this claim in the context of the SOE sector as a means of redistribution, we ran a series of tests. First, we ran our regressions including our democracy variable and the interaction of this variable with our inequality variables. Second, we used a dummy variable for democracy (trying different threshold values to separate countries into democratic and dictatorial ones) and the interaction of this dummy variable with measures of inequality. Third, we generated residuals from our main regressions and regressed the absolute values of these residuals on our democracy variable. None of these tests yielded any statistically significant results, strongly suggesting that (as far as redistribution through the SOE sector is concerned) dictatorships are as responsive to majoritarian concerns as democracies: a conclusion in conformity with the anecdotal/historical evidence a limited selection of which we

[35] The results of Ramsey's RESET test are available upon request.

cited in our introduction.[36]

5.1.6 The Role of Political Ideology

The role of political ideology or partisan politics in the choice of policy instruments and on macroeconomic outcomes is an important issue and has been explored both theoretically and empirically.[37]

Here we investigate empirically the related question as to whether left-wing governments would be more likely to redistribute (using SOE's as their redistributive tool) than right-wing governments. The investigation is motivated by the observation that the constituency of left-wing parties consist of (appropriating the terminology of Hibbs (1987)) "down-scale classes" whereas the constituency of the right-wing parties consist of "up-scale" groups who tend to be richer. Though we would expect governments composed of either type of political party to redistribute, left-wing governments would be expected to be more responsive to the demands of their core constituency and to redistribute more than their right-wing counterparts. To test this hypothesis we ran regressions of the type

$$\left(\frac{SOE}{GDP}\right)_i = \xi_1 + \xi_2 INEQ_i + \xi_3 PCAP_i + \xi_4 WING_i + \xi_5 (WING)(INEQ) + \varepsilon_i$$

where $INEQ$ is a measure of inequality and $WING$ is a variable indicating whether the government in question left-wing, center, or right-wing, taking respectively the values 1, 2, or 3.[38] For the $INEQ$ variable we used the land gini, as well as the $GINI$ and $QUIN$ variables employed above. Our results are not statistically significant in the case of land gini. However, when we use $GINI$ or $QUIN$ as our measure of inequality our results are both significant and the coefficients have the signs we expect them to have *a priori*.

[36] Alesina and Rodrik (1994) also conclude that "..even in a dictatorship, distributional issues affecting the majority of the population will influence policy outcomes." More recently, Dutt and Mitra (2002a) reach a similar conclusion in their empirical investigation of the political economy of trade policy.

[37] The basic models here are that of Hibbs (1977, 1987, 1994) and of Alesina (1987). For an extensive discussion of the issues involved, see Drazen (2000). Outside the macroeconomic literature there has not been much work done. For an exception, see Dutt and Mitra (2002b) who discuss the effects of political ideology on trade policy.

[38] The data we have on the political ideology of governments is qualitative. We converted this data to the following numerical codes: left-wing governments are assigned number 1, center governments number 2, and right-wing governments number 3. The data covers the years 1978 to 1991. Most countries in our data set had governments with that were either right- or left-wing for the entire period. In the case of those countries which had governments of different ideologies for different arts of the period covered we assigned the ideology that was dominant.

In the case where we measure inequality using $GINI$ we have $\frac{\partial SHSOE}{\partial GINI} = 0.03 - 0.01(WING)$, $\frac{\partial SHSOE}{\partial WING} = 0.33 - 0.01(GINI)$, $\frac{\partial SHSOE}{\partial GINI \partial WING} = -0.01$. *A priori*, we expect $\frac{\partial SHSOE}{\partial GINI} > 0$ since as $GINI$ and, thus, inequality increases, our theory predicts that the share of SOE's in GDP should rise. Given our numerical codification of the political ideology of governments, regression results above confirm this most importantly for left-wing and center governments. Further, we expect $\frac{\partial SHSOE}{\partial WING} < 0$ because as $WING$ rises, governments become more right-wing which we don't expect to redistribute as much leading to a decline in the share of SOE's. Our regression results above support this conclusion. Finally, we expect the left-wing governments to reinforce the tendency to redistribute when inequality is high. That is we expect $\frac{\partial SHSOE}{\partial GINI \partial WING} = \xi_5 < 0$. As the results above show this expectation is also borne out by our regressions.

When we use $QUIN$ as our mesaure of inequality we obtain $\frac{\partial SHSOE}{\partial QUIN} = -0.08 + 0.02(WING)$, $\frac{\partial SHSOE}{\partial WING} = -0.43 + 0.02(QUIN)$, $\frac{\partial SHSOE}{\partial QUIN \partial WING} = 0.02$. *A priori*, we expect $\frac{\partial SHSOE}{\partial QUIN} < 0$ since as $QUIN$ falls, this being an inverse measure of inequality, inequality rises, our theory predicts that the share of SOE's in GDP should rise. Given our numerical codification of the political ideology of governments, regression results above confirm this for all types of governments. Further, we expect $\frac{\partial SHSOE}{\partial WING} < 0$ because as $WING$ rises, governments become more right-wing which we don't expect to redistribute as much leading to a decline in the share of SOE's. Our regression results support this conclusion as well. Finally, we expect the left-wing governments to reinforce the tendency to redistribute when inequality is high. That is we expect $\frac{\partial SHSOE}{\partial QUIN \partial WING} = \xi_5 > 0$. Our results once again bear out this expectation.

6 Conclusion

The paper formalizes a popular, but informal, argument to explain the persistence of the SOE sector in many less developed countries (as well as transitional countries like Russia). In its broadest outlines the argument is that the SOE sector is used as a redistributive device and cannot be easily given up especially given the pains of other reforms that form a package of

structural adjustment. It is shown that as long as the wealth of the median voter is less than that of the average agent, the former prefers to establish or maintain an SOE sector that pays higher wages than the private sector and carries surplus labor. The paper then finds strong and robust empirical support for this hypothesis.

In addition, we use our setup to test two questions that generally come up in the political economy literature when median voter is invoked: whether (i) democracies are more likely to be responsive to popular pressures when it comes to redistribution, and (ii) whether and to what extent the political ideology of governing parties affects redistribution. We find first that we cannot find any support for the hypothesis that democracies would be more likely to redistribute through SOE's than dictatorships. Second, we do find strong and robust support for the hypothesis that left-wing governments are more prone to use the SOE' sector as a redistributive device.

The next natural question to ask, given the result obtained, is how to explain the successful privatization experiments such as Argentina's. To answer the question one can point out that some factors that are not taken into account in the present model drive the process of privatization. For example, Waterbury argues that this process is driven by fiscal crises of varying intensity coupled with inflation, reduced international credit-worthiness, and impediments to export promotion. Since, to keep the model as analytically simple as possible, we have abstracted from such considerations, the model will not help us explore these factors.

References

Alesina, A., 1987, "Macroeconomic policy in a two-party system as a repeated game," Quarterly Journal of Economics 102-651-678.

Alesina, A. and D. Rodrik, 1992, "Distribution, political conflict, and economic growth: A simple theory and some empirical evidence," in Cukierman, A., Z. Hercowitz, and L. Leiderman (eds.,) Political Economy, Growth, and Business Cycles, Cambridge, MIT Press.

Alesina, A. and D. Rodrik, 1994, "Distributive politics and economic growth," Quarterly Journal of Economics 109:465-490.

Benhabib, J. and A. Rusticchini, 1996, "Social conflict and growth," Journal of Economic Growth 1:125-142.

Buffie, E., 1989, "Economic policy and foreign debt in Mexico", in Jeffrey Sachs, ed., Developing Country Debt and Economic Performance, vol. 2. Chicago: University of Chicago Press.

Coate S. and S. Morris, 1995, "On the form of transfers to special interests" Journal of Political Economy 103, 1210-1235.

Davidson, R. and J. McKinnon, 1993, *Estimation and Inference in Econometrics,* Oxford University Press, Oxford.

Deininger, K. and L. Squire, 1996, "A New Data Set Measuring Income Inequality", *World Bank Economic Review* 10: 565-591.

Deininger, K. and L. Squire, 1998, "New Ways of Looking at Old Issues", Journal of Development Economics, 57: 259-87

Drazen, A. Political Economy in Macroeconomics, Princeton, NJ, Princeton University Press.

Dutt, P. and D. Mitra, 2002a, "Endogenous Trade Policy Through Majority Voting: An Empirical Investigation" Journal of International Economics 58: 107-134.

Dutt, P. and D. Mitra, 2002b, "Political ideology and endogenous trade policy," mimeo.

Greene, Willian H.1997, *Econometric Analysis,* Macmillan Publishing Company.

Hibbs, D. 1977, "Political parties and macroeconomic policy," American Political Science Review 71: 1467-1487.

Hibbs, D. 1987, *The American Political Economy: Macroeconomics and Electoral Politics in the United States*, Cambridge, MA, Harvard University Press.

Hibbs, D. 1994, "The partisan model of macroeconomic cycles: More theory and evidence for the United States," Economics and Politics 6: 1-24.

Judge, George, W.E.Griffiths, R.C. Hill, H.Lütkepohl and T. Lee, 1984, *The Theory and Practice of Econometrics,* John Wiley and Sons.

Kennedy, P. 1988, *A Guide to Econometrics*, Cambridge, MA, The MIT Press.

Morales, J.A. and J. Sachs, 1989, "Bolivia's economic crisis", in Jeffrey Sachs, ed., Developing Country Debt and Economic Performance, vol. 2. Chicago: University of Chicago Press.

Nehru, V., and A. Dhareshwar, 1993, "A New Database on Physical Capital Stock: Sources, Methodology and Results'. Revista de Análisís Económico. Vol. 8, No. 1: 37-59.

Persson, T. and G. Tabellini, 1992, "Growth, distribution, and politics," in Cukierman, A., Z. Hercowitz, and L. Leiderman (eds.,) Political Economy, Growth, and Business Cycles, Cambridge, MIT Press.

Ramanadham, V., 1988, *Public Enterprise and Income Distribution*, London, Routledge.

Tornell, A. and A. Velasco, 1992, "The tragedy of commons and economic growth: Why does capital flow from poor to rich countries?," Journal of Political Economy 100, 1208-1231.

Waterbury, J., 1992, "The heart of the matter? Public enterprise and the adjustment process," in Haggard, S. and R. Kauffman (eds.,) The Politics of Economic Adjustment, Princeton, Princeton University Press.

Waterbury, J., 1993, *Public Enterprise and State Power in Egypt, India, and Turkey*, Cambridge, Cambridge University Press.

World Bank, 1995, *Bureaucrats in Business: The Economics and Politics of Government Ownership,* Policy Research Report.

World Development Indicators, 1999.

6.1 Appendix I

$$\left(\frac{SOE}{GDP}\right)_i = \alpha_1 + \alpha_2 LAND_i + \alpha_3 PCAP_i + \varepsilon_i \tag{R1}$$

$$\left(\frac{SOE}{NONAG}\right)_i = \beta_1 + \beta_2 LAND_i + \beta_3 PCAP_i + \xi_i \tag{R2}$$

$$\left(\frac{SOE\_EMP}{EMP}\right)_i = \gamma_1 + \gamma_2 LAND_i + \gamma_3 PCAP_i + \epsilon_i \tag{R3}$$

$$\left(\frac{SOE}{GDP}\right)_i = \delta_1 + \delta_2 QUIN_i + \delta_3 PCAP_i + \omega_i \tag{R4}$$

$$\left(\frac{SOE}{NONAG}\right)_i = \phi_1 + \phi_2 QUIN + \phi_3 PCAP_i + \nu_i \tag{R5}$$

$$\left(\frac{SOE\_EMP}{EMP}\right)_i = \eta_1 + \eta_2 QUIN + \eta_3 PCAP_i + u_i \tag{R6}$$

$$\left(\frac{SOE}{GDP}\right)_i = \theta_1 + \theta_2 GINI_i + \theta_3 PCAP_i + \zeta_i \tag{R7}$$

$$\left(\frac{SOE}{NONAG}\right)_i = \lambda_1 + \lambda_2 GINI_i + \lambda_3 PCAP_i + \varsigma \tag{R8}$$

$$\left(\frac{SOE\_EMP}{EMP}\right)_i = \mu_1 + \mu_2 GINI + \mu_3 PCAP_i + \tau_i \tag{R9}$$

Table 1. Estimation Results for Wealth Distribution, Equations (R1) to (R3)

	R1	R2	R3
	$\frac{SOE}{GDP}$	$\frac{SOE}{NONAG}$	$\frac{SOE\_EMP}{EMP}$
$Constant$	0.0387	0.0204	0.1470
	0.0455)	(0.0587)	(0.0601)
$LAND$	0.0013**	0.0017**	-0.0006
	(0.0006)	(0.0008)	(0.0010)
$PCAP$	-0.0122***	-0.0126***	-0.0171
	(0.0038)	(0.0049)	(0.0129)
$Est.Methodology$	W2SLS	W2SLS	OLS
N	43	42	25
$F-statistic$	6.09***	4.46**	1.69

Table 2. Estimation Results for Income Distribution, Equations (R4) to (R9)

	R4	R5	R6	R7	R8	R9
	$\frac{SOE}{GDP}$	$\frac{SOE}{NONAG}$	$\frac{SOE\_EMP}{EMP}$	$\frac{SOE}{GDP}$	$\frac{SOE}{NONAG}$	$\frac{SOE\_EMP}{EMP}$
$Constant$	0.4343***	0.5082***	0.2507	-0.0870	-0.1924**	-0.1329
	(0.1568)	(0.1622)	(0.2355)	(0.0910)	(0.0958)	(0.1145)
$QUIN$	-0.0218**	-0.0266**	0.0009			
	(0.0105)	(0.0108)	(0.0150)			
$GINI$				0.0084***	0.0112***	0.0047*
				(0.0024)	(0.0026)	(0.0025)
$PCAP$	-0.0006	0.0023	-0.1046**	-0.0210***	-0.0171***	-0.0005
	(0.0087)	(0.0090)	(0.0479)	(0.0054)	(0.0057)	(0.0045)
$Est.Meth.$	W2SLS	W2SLS	OLS	WLS	WLS	WLS
N	36	35	25	37	36	17
$F-stat.$	3.31**	3.97**	2.86*	14.12***	14.560***	2.27

Note: 1. Standard errors in parantheses. *** denotes the significance level at 1%, ** at 5%, and * at 10%.

Table 3.1: Hausman Tests, Equations (R1) to (R3)

	R1	R2	R3
	$\frac{SOE}{GDP}$	$\frac{SOE}{NONAG}$	$\frac{SOE\_EMP}{EMP}$
$Constant$	0.1779***	0.2636***	0.1487**
	(0.0612)	(0.0823)	(0.0615)
$LAND$	-4.27E-05	-0.0004	-0.0006
	(0.0009)	(0.0012)	(0.0010)
$PCAP$	-0.0243***	-0.0383***	-0.0197
	(0.0094)	(0.0127)	(0.0151)
RES_PCAP	0.0230*	0.0353**	0.0051
	(0.0129)	(0.0174)	(0.0143)
R^2	0.1348	0.1864	0.1387
N	50	49	25
$F-stat.$	2.39*	3.44**	1.13

Note: 1. Standard errors in parantheses. *** denotes the significance level at 1%, ** at 5%, and * at 10%.

2. RES_PCAP is the coefficient of the residuals of the per capita income from the auxiliary regression in which the $PCAP$ is regressed on exogenous variables such as land gini, the indicator of democracy and the dummies of lower-middle and upper-middle income countries.

Table 3.2. Hausman Tests, Equations (R4) to (R9)

	R4	R5	R6	R7	R8	R9
	$\frac{SOE}{GDP}$	$\frac{SOE}{NONAG}$	$\frac{SOE\_EMP}{EMP}$	$\frac{SOE}{GDP}$	$\frac{SOE}{NONAG}$	$\frac{SOE\_EMP}{EMP}$
$Constant$	0.2949**	0.3166**	0.0695	-0.0341	0.0201	0.2630
	(0.1118)	(0.1461)	(0.2778)	(0.1040)	(0.1203)	(0.1768)
$QUIN$	-0.0107	-0.0080	0.0061			
	(0.0083)	(0.0110)	(0.0183)			
$GINI$				0.0031	0.0026	-0.0038
				(0.0021)	(0.0024)	(0.0038)
$PCAP$	-0.0134	-0.0265**	-0.0485	-0.0005	-0.0065	-0.0131
	(0.0096)	(0.0126)	(0.0349)	(0.0099)	(0.0114)	(0.0153)
RES_QUIN	0.0236**	0.0213	-0.0067			
	(0.0107)	(0.0140)	(0.0183)			
RES_GINI				-0.0029	-0.0017	0.0069
				(0.0032)	(0.0037)	(0.0056)
RES_PCAP	0.0131	0.0237	0.0070	-0.0018	0.0030	0.0307
	(0.0125)	(0.0164)	(0.0203)	(0.0127)	(0.0147)	(0.0190)
R^2	0.2198	0.2267	0.2393	0.1080	0.1105	0.2315
N	41	40	20	33	32	15
$F-stat.$	2.54*	2.57**	1.18	0.85	0.84	0.75
$F-Hausman$	3.83**	3.01*	0.08	0.42	0.17	1.44

Note: 1. Standard errors in parantheses. *** denotes the significance level at 1%, ** at 5%, and * at 10%.

2. RES_QUIN and RES_GINI are the coefficients of the residuals of $QUIN$ and $GINI$ from the auxiliary regressions.

3. F-Hausman shows the joint significance of the residuals of $QUIN$ and $PCAP$, and $GINI$ and $PCAP$. It is the significance of these tests that leads to the conclusion of endogeneity.

Table 4. Modeling Heteroscedasticity, Eqs. (R1) to (R9)

$log(\widehat{\varepsilon}_i^2) from$	$\widehat{\delta}$	$\widehat{\theta}$	N
$Eq.(R1)$	-1.8983	-0.2391**	43
	(1.9536)	(0.2287)	
$(R2)$	-1.0955	-0.2600**	42
	(2.0953)	(0.1222)	
$(R3)$	-5.9178	-0.0141	21
	(2.9404)	(0.1872)	
$(R4)$	-3.4140	2.1800§	36
	(-0.1487)	(0.1283)	
$(R5)$	-2.1340	-0.2341§	35
	(3.0324)	(0.1778)	
$(R6)$	-6.4102	0.0333	20
	(4.2451)	(0.2762)	
$(R7)$	-2.5914	-0.2530§§	29
	(3.0746)	(0.2385)	
$(R8)$	0.2633	-0.4028**	28
	(3.3293)	(0.1873)	
$(R9)$	2.4766	-0.5474*	17
	(4.9640)	(0.2998)	

Note: 1. Standard errors in parantheses. *** denotes significance level at 1%, ** at 5%, and * at 10%. For equations R4, R5 and R7, §§ denotes significance level at 15% and § at 25%. Weighting these equations with capital-labor ratio improves the efficiency of the estimates.

6.2 Appendix II

$$r_\tau = \frac{-(1-\beta)Q}{(1-\tau)K} = \frac{-(1-\beta)q_p}{(1-\tau)k} < 0 \tag{1}$$

$$dl_p/d\tau = \frac{-(1-\beta)q_p}{\alpha\beta} < 0 \tag{2}$$

$$\frac{d(\tau q_p)}{d\tau} = \frac{(\beta-\tau)q_p}{\beta(1-\tau)} \tag{3}$$

$$q_p = k^\beta l_p^{1-\beta} \tag{4}$$