Darmstadt Discussion Papers in Economics

Endogenous Policy and Cross-Country Growth Empirics

Günther Rehme

Nr. 182

Arbeitspapiere
des Instituts für Volkswirtschaftslehre
Technische Universität Darmstadt

ISSN: 1438-2733

Economic Theory

Endogenous Policy and Cross-Country Growth Empirics

Günther Rehme*

Technische Universität Darmstadt†

April 13, 2007

Abstract

In this paper it is shown that it matters a lot for empirical research whether policy is taken to be exogenously set or to be endogenous. In the model investment depends on policy which depends on economically important fundamentals and is, thus, endogenous. Conditioning on factor accumulation in growth regressions that also include endogenous policy variables may then be problematic. When policy is endogenous the measured effects of policy on growth will generally be biased. Based on the model and OECD data, the signs of the biases for tax variables related to the tax base and for redistribution are derived. Based on these signed biases the paper discusses some empirical results that seem puzzling from a theoretical viewpoint. The paper argues that regressing growth on policy may still yield important information if policy endogeneity is taken account of.

KEYWORDS: Growth, Policy, Cross-Sectional Models

JEL classification: O4, D3, C2

*I would like to thank Tony Atkinson, James Mirrlees, Robert Waldmann, and Robert K. von Weizsäcker for helpful advice and useful suggestions. I have also benefitted from comments by Volker Caspari, Sabine Eschenhof, Berthold Herrendorf, Michael Pflüger, Jens Rubart, Jan-Christoph Rülke, Hannes Spengler, and Matti Viren as well as from presentations at the Econometric Society in Providence, Rhode Island, and Madrid, the German Economic Association in Dresden, and Saarland University in 2004. Of course, all errors are my own. Financial support by the Deutscher Akademischer Austauschdienst (DAAD), grant no. 522-012-516-3, is gratefully acknowledged.

†*Correspondence*: TU Darmstadt, FB 1/VWL 1, Schloss, D-64283 Darmstadt, Germany. *phone*: +49-6151-162219 ; *fax*: +49-6151-165553; *E-mail*: rehme@hrzpub.tu-darmstadt.de

1 Introduction

Most researchers acknowledge that empirical work on the effects of policy on economic growth may be riddled by endogeneity problems. This paper concentrates on that issue. It attributes any discrepancy of results to the fact that policy is economically endogenous and that treating it as exogenous provides one with a misleading picture of the empirical relationship between policy and growth.

For instance, many authors have investigated the effects of taxation on long-run growth. Although employing similar theoretical frameworks[1], their conclusions differ widely. See, for example, King and Rebelo (1990), Lucas (1990), Rebelo (1991), Jones, Manuelli, and Rossi (1993), Pecorino (1993), or Stokey and Rebelo (1995).

The link between (re-)distribution and growth has e.g. been analyzed by Bertola (1993), Alesina and Rodrik (1994), Persson and Tabellini (1994), or Perotti (1996). These studies often provide theoretical arguments that redistribution of resources from the accumulated towards the non-accumulated factor of production should be expected to affect growth negatively.

To test these theoretical predictions a large number of contributions has used cross-country growth regressions[2] This paper relates to that research. It first provides a theoretical model that is based on a simplified version of Alesina and Rodrik (1994). This model provides the theoretical "lens" through which we will look at some empirical findings. This is because it features some commonly agreed upon properties. For instance, policy affects factor accumulation, which in turn bears on output growth, and redistribution lowers long-run growth in the model. Importantly, policy is derived from optimizing behaviour which takes account of fundamental economic variables. Thus, policy is endogenous in the

[1]For instance, Koester and Kormendi (1989), Barro (1991), Barro (1997), Levine and Renelt (1992), Easterly and Rebelo (1993), or Sala-i-Martin (1997) have empirically analyzed the effects of fiscal policy on growth. Most of them find that tax rates or other, tax financed fiscal variables have a negative, but - when controlling for initial income - statistically insignificant effect on growth.

[2]In the paper I contemplate 'simple' cross-country growth regressions which are meant to reflect the procedure of relating time averages of or intial period data to each other. Of course, 'simple' does not mean simplistic, since the availability of data may not allow for another or a 'better' method of analysis. Some authors have advocated the use of *dynamic* panel data methods to pay explicit attention to the time series dimension. See e.g. Caselli, Esquivel, and Lefort (1996). But the latter methods seem to have their own problems as e.g. argued by Barro (1997), p. 37, Temple (1999), p. 132, and e.g. analyzed by Banerjee, Marcellino, and Osbat (2000).

paper.

It is generally agreed that productivity differences due to, for example, cultural, institutional or technological heterogeneity across countries play a major role in explanations of differences in growth rates. See, for example, Prescott (1998). Therefore, the paper focuses on the effects of differences in productivity and takes productivity as *the* economic fundamental. The paper derives the theoretical signs of covariances between policy, the growth rate and that fundamental economic variable.

The signs of the covariances are then checked empirically. Based on the approach of Hall and Jones (1999) I calculate productivity data and use tax data from the OECD and other well-known sources to find - at least for the OECD countries - that the theoretical signs of the empirical covariances for the theoretical predictions are borne out by the data.

I then conduct experiments with these data, which focus on estimation problems only.[3] This allows me to use as much empirical information as possible. I show how productivity bears on policy and growth. Based on the data I then relate the theoretical predictions to common empirical set-ups. In line with most studies it is assumed that productivity differences are usually unobservable. In this case estimation would suffer from omitted variable bias. The theoretical model is then transformed into a linear regression model. The derived signs of the covariances allows one to *sign* the biases. These signs are confirmed when calculating the biases using the paper's data.

For more complex growth regressions in the spirit of Mankiw, Romer, and Weil (1992) (MRW) it is found that signing the biases for the policy variables must be done by the data at hand. The same holds true for simplified variants of the set-ups used in the model uncertainty literature. See, for instance, Levine and Renelt (1992) or Sala-i-Martin (1997).

It is shown that growth regressions would generally yield misleading results if output growth is conditioned on variables relating to factor accumulation *and* policy. This is because, according to the theory, policy bears on investment which in turn affects output growth. Thus, including variables for policy and factor accumulation would lead to misspecified models.

[3] In terms of methodology I will only focus on estimation issues in the empirical exercises that follow. We will not address issues of statistical significance, which rely on estimation anyway. For a justification see, for example, McCloskey (1985) and McCloskey and Ziliak (1996).

Interestingly, the signs of the biases for the policy variables are the same for all the different empirical models studied. Furthermore, they are equal to what the simplest linear empirical formulation of the theory implies. In particular, the empirical exercise would seem to confirm the following:

The estimates for the effects on growth of tax rate variables related to the tax base are generally biased upwards and so overestimated. Thus, any reported negative effect of taxes on growth is understated, if measured by these variables.[4]

Under the assumption of endogenous policy, the estimated coefficients of the effect on growth of redistribution are generically biased downwards in the model. That would render the hypothesis that redistribution is bad for growth untestable. This is because the prediction of the theoretical model is that redistributive transfers are bad for growth. However, in the model an increase in efficiency makes an optimizing, redistributing government grant less transfers to the non-accumulated factor of production. This last effect is ignored in growth regressions when one assumes that public policy is exogenous.

For theoretical reasons many researchers expect a negative coefficient for the effect of redistributive transfers on growth. However, many people find positive coefficients.[5] As any downward bias of the estimated coefficients may be as large as minus infinity, a reported negative coefficient cannot corroborate the hypothesis that redistribution is bad for growth. On the other hand, any downward bias is perfectly consistent with many empirical findings and the alternative hypothesis that redistribution is not bad for growth.

Recently, Rodrik (2005) has argued that we "learn nothing from regressing economic growth on policy". In this paper a more positive stance is taken. The analysis reveals that acknowledging that policy is endogenous may indeed provide an avenue to gain some understanding of how policy and growth are associated across countries. For this one may simply have to analyze potential bias problems in more depth. The present paper moves in that direction by presenting experiments and empirical exercises to show how this may be accomplished.

[4]Effect is not meant to be causal. In this paper effect means that some underlying economic fundamental influences policy which in turn bears on growth. Then the "true" effect of policy on growth is spurious really, but can only be picked up by linear operationalizations of the model using data for policy variables. The paper then conducts thought experiments by interpreting empirical results under the assumption that policy is exogenous or endogenous.

[5]For example, Sala-i-Martin (1996) finds this for the effect of social security contributions on growth.

The main insights to be drawn from the paper are the following. The disentanglement of the interplay of economic fundamentals and policy one the one hand and policy and growth on the other should provide an interesting area for research. Conditioning on factor accumulation in growth regressions that also include policy variables may be problematic. Furthermore, analyzing biases in growth empirics should not be too difficult and would base some findings on a sounder footing.

The paper is organized as follows: Section 2 presents the theoretical model and derives the signs of the covariances. Section 3 presents empirical evidence for the covariances. Section 4 analyzes the bias problem theoretically for simple set-ups related to the literature. Section 7 presents empirical findings for the biases. Section 8 provides concluding remarks.

2 Theory

Consider a private ownership economy that is populated by two types of price-taking, infinitely lived individuals who are all equally patient. One group of agents, the capitalists (k), owns wealth equally and does not work. The other group of agents, the workers (W), owns (raw) labour equally, but no capital.[6] Population is stationary and each group of agents derives logarithmic utility from the consumption of a homogeneous, malleable good.[7]

Aggregate output is produced according to

$$Y_t = B\, K_t^{\alpha} L_t^{1-\alpha} \quad , \quad 0 < \alpha < 1 \tag{1}$$

[6]The assumption may be justified by various arguments, especially for the long run. See e.g. Kaldor (1956), Pasinetti (1962), Schlicht (1975), Bourguignon (1981), or Bertola (1993).

[7]This assumption is invoked for two reasons. First, suppose that ex ante, under a veil of ignorance (see, for instance, Harsanyi (1955)) people face a positive probability of becoming a worker or a capital owner. This risk must be evaluated by agents who have to make decisions for their and their offspring's lifetime income, given that they end up in some class. For such a scenario Sinn (2003), Robson (2001), Robson (1996) and Sinn and Weichenrieder (1993) have shown that only those people do best (in a biological selection process) that evaluate such risky choices by logarithmic utility functions. Thus, the model concentrates on "surviving" individuals in a world with risk and uses their ("fittest") preferences in a world with certainty. This also justifies why agents may have the same rate of time preference. The second reason is empirical. Recent evidence indicates that the intertemporal elasticity of substitution is in fact close to one. See, for example, Beaudry and van Wincoop (1996). Thus, these two arguments may justify a set-up with logarithmic instantaneous utility.

where Y_t denotes aggregate output, K_t is the real capital stock, and L_t is (unskilled) labour. Capital is broadly defined and includes human capital.[8] Labour is inelastically supplied and normalized so that the total labour endowment equals unity, $L_t = 1$. The model abstracts from the depreciation of capital so that output and factor returns are really defined in net terms.

The index B reflects multifactor productivity. Following Barro (1990) we assume that the latter depends on public inputs in production, G_t.[9] In particular, productivity is taken to depend on public inputs per worker as follows

$$B = A \left(\frac{G_t}{L_t} \right)^{\delta} \quad , \quad 0 < \delta < 1. \tag{2}$$

Thus, more public resources per worker channelled into production raise aggregate productivity and create a positive externality in production. For a detailed discussion of such external effects see Barro (1990) or Barro and Sala–i–Martin (1995), chpt. 4.4. In turn, the index A is constant and represents the economy's state of technology which depends on cultural, institutional and technological development and captures long-run exogenous factors that play a role in the production process.

The public inputs in production are financed by a wealth tax. Following Alesina and Rodrik (1994) we use a wealth tax scheme as a metaphor to represent a broad class of redistributive tax arrangements, which distort the investors' incentive to accumulate. In particular, we assume that the government taxes the accumulable factor of production at the constant rate τ, redistributes a constant share λ of its tax revenues to the (unskilled) workers[10] and runs a balanced budget: $\tau K_t = G_t + \lambda \tau K_t$. The LHS depicts the tax revenues and the RHS public expenditures. The workers receive $\lambda \tau K_t$ as transfers and G_t is spent on public inputs to production. Thus, λ will denote the extent of (unproductive) redistribution from the accumulated factor of production (capital) to the non-

[8]This eliminates a separate treatment of how human capital is accumulated. By assumption the economies are perfectly competitive and the return on human capital services equals that of physical capital services.

[9]G_t is a flow by assuming that the government purchases a portion of the private output and then uses it to provide free public services to private producers. See Barro and Sala–i–Martin (1995), p. 152.

[10]Alesina and Rodrik (1994) show that the optimal policies are constant over time and, thus, time-consistent. For convenience constancy of policy is assumed from the beginning in this paper.

accumulated factor of production (unskilled labour) in this paper.[11]

There are many identical, profit-maximizing firms which operate in a perfectly competitive environment. They are owned by the capital owners who rent capital to and demand shares of the firms. The shares are collateralized one-to-one by capital. The markets for assets and capital are assumed to clear at each point in time. The firms take B as given and rent capital and labour in spot markets in each period. The price of output y_t serves as numéraire and is set equal to one. Profit maximization entails that firms pay each factor of production its marginal product,

$$r_t = \alpha B K_t^{\alpha-1} \quad \text{and} \quad w_t = (1-\alpha) B K_t^{\alpha} \tag{3}$$

where we have used the normalization $L_t = 1$, and denote the rental rate for capital by r_t and the wage rate by w_t. Thus, the returns to each factor are decreasing for each firm at a *given* level of productivity. Below we will see that more public inputs in production have a positive, counteracting bearing on the marginal products through their positive externality effect.

The (unskilled) workers derive utility from consuming their entire income. They do not invest and are not taxed. Their intertemporal welfare is given by

$$\int_0^{\infty} \ln C_t^W \, e^{-\rho t} dt \quad \text{where} \quad C_t^W = w_t + \lambda \tau K_t. \tag{4}$$

The capitalists choose how much to consume or invest, and they have perfect foresight about the prices and tax rates, which they take as given. They maximize their intertemporal utility according to

$$\max_{C_t^k} \int_0^{\infty} \ln C_t^k \, e^{-\rho t} dt \tag{5}$$

$$s.t. \qquad \dot{K}_t = (r_t - \tau) K_t - C_t^k \tag{6}$$

$$K(0) = \overline{K_0}, \quad K(\infty) = \text{free}, \tag{7}$$

[11]The model would have the same qualitative features if redistribution was instead defined as, for example, tax rates higher than those which are optimal for growth, or after-tax incomes of workers relative to that of the capital owners (under some preferred policy) relative to that ratio under a growth maximizing policy. See, for example, Rehme (2006b) and Rehme (2006a).

where equation (6) is the capitalists' dynamic budget constraint which depends on their after-tax income $(r_t - \tau)K_t$. In appendix A it is shown that their consumption optimally grows at

$$\gamma \equiv \frac{\dot{C}_t^k}{C_t^k} = (r_t - \tau) - \rho \tag{8}$$

which is increasing in the after-tax return on capital.

2.1 Equilibrium

In equilibrium the spillovers resulting from public inputs in production influence the factor rewards. As Barro (1990) and Barro and Sala–i–Martin (1995), chpt. 4.4, we will concentrate on situations where there is endogenous growth.[12] To this end assume $\delta = 1 - \alpha$ so that $B = A\left(\frac{G}{L}\right)^{1-\alpha}$. Since $G = (1-\lambda)\tau K$ the factor rewards in (3) then become

$$r = \alpha A[(1-\lambda)\tau]^{1-\alpha} \tag{9}$$

$$w_t \equiv \eta(\tau, \lambda)K_t = (1-\alpha)A[(1-\lambda)\tau]^{1-\alpha}K_t. \tag{10}$$

Thus, the productive role of government services makes policy have a bearing on the marginal products. The return on capital is constant over time while the wages grow with the capital stock. Notice that more redistribution lowers r and η, while higher taxes raise them.

Given the constancy of the return to capital, $r_t = r$, it is not hard to verify that $\gamma = \frac{\dot{C}_t^k}{C_t^k} = \frac{\dot{K}_t}{K_t} = (r - \tau) - \rho$ in equilibrium. (See appendix A.) In fact, all relevant variables grow at this rate. Thus, in steady state the economy is characterized by *balanced growth* at the rate γ, which is first increasing and then decreasing in τ for given λ. Thus, first the positive effect of public inputs for the after-tax return on capital dominates and that raises growth. But eventually the taxes necessary to finance public inputs put a brake on this and reduce the after-tax return on capital, which lowers growth. Growth is maximized when $\tau = [\alpha(1-\alpha)A]^{\frac{1}{\alpha}} \equiv \hat{\tau}$ and $\lambda = 0$. If taxes higher than $\hat{\tau}$ are levied, then growth is traded off against redistribution when $\lambda > 0$ and $\check{\tau} > \hat{\tau}$.

[12] Thus, the setup builds on Frankel (1962) and Romer (1986) who provide a detailed discussion of the spillover effects that may generate endogenous growth.

Furthermore, $r - \tau = \alpha A[(1-\lambda)\tau]^{1-\alpha} - \tau$ so that for *given* policy an increase in efficiency A raises growth.

2.2 Policy

To capture distributional conflicts we will for simplicity focus on governments with opposing preferences.[13] Integrating the agents' welfare functions (5) and (4) under the condition that the growth rate is constant in equilibrium yields the intertemporal welfare of an entirely pro-capital, V^r, resp. entirely pro-labour government[14], V^l,

$$V^r(C_t^k) = \frac{\ln(\rho K_0)}{\rho} + \frac{\gamma}{\rho^2} \quad \text{and} \quad V^l(C_t^W) = \frac{\ln\left[(\eta(\tau,\lambda) + \lambda\tau)K_0\right]}{\rho} + \frac{\gamma}{\rho^2}. \quad (11)$$

The governments respect the right of private property[15] and maximize the welfare of their clientele under the condition $\lambda \geq 0$. That restricts the governments in that even an entirely pro-capital government does not tax the (unskilled) workers.

The optimal pro-labour policy is derived in appendix B and given by

$$\text{If} \quad \rho \geq [(1-\alpha)A]^{\frac{1}{\alpha}} \quad \text{then:} \qquad \tau = \rho, \qquad \lambda = 1 - \frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{\rho}\,. \quad (12)$$

$$\text{If} \quad \rho < [(1-\alpha)A]^{\frac{1}{\alpha}} \quad \text{then:} \qquad \tau[1 - \alpha(1-\alpha)A\tau^{-\alpha}] = \rho(1-\alpha), \qquad \lambda = 0\,. \quad (13)$$

Denote the optimal pro-labour tax rate by $\check{\tau}$ and notice that for a wide range

[13] It is possible to analyze political preferences by weighting the agents' welfare by some constant. Except for introducing a new parameter (see appendix B) that is hard to measure the model's features would not change with respect to the fundamentals considered of prime importance for policy here.

[14] This assumption allows to place governments on the traditional 'left' (pro-worker) - 'right' (pro-capital) spectrum. The analysis may capture democratic and non-democratic political regimes. As regards democratic regimes the assumption has the advantage of transcending the more conventional Downsian approach of political decision making that relies on a median voter. The Downsian approach can be criticized on various grounds. The present set-up avoids these criticisms by assuming that political parties, once in power, adhere to their party platforms. On Downsian approaches and justifications of why the present set-up is useful see, for instance, Roemer (2001).

[15] Although the command optimum in the model would call for expropriation of capital even for an entirely pro-capital government, it is ruled out as a policy option since it is not very common in the real world.

of parameter values there is no redistribution. In particular, if the agents are sufficiently patient or the economy is very efficient, the owners of the non-accumulated factor of production (workers) prefer to have higher growth instead of direct redistribution. This is because high growth may be better for their income stream and so long-run welfare than direct (unproductive) transfers. In that way the model distinguishes between redistributing and non-redistributing (pro-labour) governments.

In contrast, the pro-capital government chooses $\tau = \hat{\tau}$, does not redistribute, grants the maximum after-tax return on capital, and acts growth maximizing in this model.[16]

All optimal policies depend on A, α, and ρ. In that sense policy is endogenous. The rate of time preference will not be considered any further because it is considered a variable, which is very hard to measure. Furthermore, most researchers find that there is not much variability in the capital share α over time and across countries. See, for example, Barro and Sala–i–Martin (1995), p. 380; Mankiw, Romer, and Weil (1992), p. 341; Sachs (1979), Table 3, or more recently Gollin (2002). For that reason it is commonly ignored in growth regressions. However, A is usually considered a very important variable for which I find the following:[17]

Table 1: Growth and Policy Effects

	PC		$PL_{\lambda=0}$		$PL_{\lambda\geq 0}$		
	$\hat{\tau}$	$\hat{\gamma}$	$\check{\tau}$	$\check{\gamma}$	$\check{\tau}$	λ	$\check{\gamma}$
A	+	+	+	+	0	-	+

PC - pro-capital, PL - pro-labour
Sign: (+) - positive, (−) - negative

[16]The optimal pro-capital and growth maximizing policies may not always coincide. But that feature of the model is not essential as long as the optimal pro-capital policy is taken to lead to higher growth than a policy that is optimal for the non-accumulated factor of production (labour), and if the pro-capital policy has the properties derived from the model. Although there is no necessity for these points to be true, see Rehme (2002a), the assumptions appear to capture what is conventionally argued to be the case .

[17]The signs of these effects are derived in appendix C. In Rehme (1998), chpt. 1, I also analyze the effects of changes in the capital share on policy and growth and find that they have ambiguous effects on policy but generally raise growth under endogenous policy.

Thus, an increase in efficiency raises growth, does not imply lower tax rates but calls for lower redistribution under all policies considered. This is due to the positive externality of public inputs. In the optimum higher efficiency calls for more tax revenues for productive services channelled into production and calls for less direct (unproductive) redistribution. That in turn raises the return on capital and so growth.[18] These are the theoretical predictions of the signs of correlations and covariances one should expect in empirical research.

In the appendix it is also shown that the growth rate is convex in A under all policies. Furthermore, the tax rates are also convex in A under the policies. For the redistributing policy I find that λ is concave in A.

3 Empirical Evidence

In this section empirical checks are provided for the theoretical predictions in Table 1. The empirical evidence should be viewed as suggestive only, because reliable data on tax rates and redistribution are in general not easy to obtain for a large set of countries.[19] Therefore, I concentrate on the subset of OECD countries for which data are more readily available.[20]

In terms of methodology I will only focus on estimation issues in the empirical exercises that follow. We will not address issues of statistical significance, which rely on estimation anyway. For a justification of such an approach see, for example, McCloskey (1985) and McCloskey and Ziliak (1996).

3.1 The Data

Countries differ widely in the level of development, but reliable data capturing the *level* of development are not easily available. In order to calculate the model's state of technology, A, I have proceeded as follows: Following Hall and Jones

[18]For example, Demetriades and Mamuneas (2000) provide empirical evidence for the positive effect of public infrastructure on intertemporal output and how close various countries are to the growth maximizing public investment rate.

[19]For a detailed description of the data and the methods used in this paper confer http://www.tu-darmstadt.de/~rehme/endopol/data04-07.htm. The sample is presented in Table 4 on p. 42.

[20]An advantage of this, though, is that they form a homogeneous group so that one does not need to control for regional disparities by means of dummy variables as has so often been done in cross-country research.

(1999) I have used a *levels accounting* framework based on a production function approach. Under the assumptions made the aggregate production function is given by $Y = BK^{\alpha}L^{1-\alpha} = A\left(\frac{G}{L}\right)^{1-\alpha}K^{\alpha}L^{1-\alpha} = AG^{1-\alpha}K^{\alpha}$. From this the level of technology can be calculated as

$$A = \left(\frac{G}{Y}\right)^{\alpha-1}\left(\frac{K}{Y}\right)^{-\alpha}.$$

Using data from Barro and Lee (1994), Summers and Heston (1991) and the OECD, I have calculated long-run averages for G/Y and K/Y for the period 1970-2000.[21] Following Mankiw, Romer, and Weil (1992) I have set the share of broad capital α equal to 0.67. The resulting variable measuring multi-factor productivity is called GA. The cross-country productivity differences are expressed relative to the United States for which the value is set at unity.[22] For checks on the second derivatives of our variables of interest we also look at the squared values of GA, called GA^2.

Redistribution is measured following Milanovic (2000). Based on household data from the Luxembourg Income Study (LIS) redistribution is calculated from Gini indices for the distribution of households' factor income. These they are compared to the Gini indices for the distribution of disposable income. The difference between them is then related to the Gini coefficient for factor incomes.[23]

[21] Notice that we have assumed that G is the *flow* of public inputs in production. Barro and Lee (1994) provide data for G/Y by their variable INVPUB which measures the ratio of nominal public domestic investment (fixed capital formation) to nominal GDP. For the period 1970-90 K/Y was calculated from Summers and Heston (1991), PWT (Mark 5.6), who provide data of capital per worker (KAPW) and real GDP per worker. For the period 1990-2000 I have used OECD data for government fixed capital formation and real GDP. Notice that according to the model $K/Y = A^{-1}\left(\frac{K}{G}\right)^{1-\alpha} = \tau^{\alpha-1}/A$ which is constant, no matter whether capital is broadly defined or not. This is what I assume to be the case here. Thus, even though capital may only include physical capital in KAPW, it would still be a proxy for the models capital-output ratio. See the online appendix for more details.

[22] An earlier version of the paper used the productivity data from Hall and Jones (1999). The subsequent results are qualitatively and almost quantitatively the same when using their data. Thus, their productivity data are indeed a good proxy for the model's A as argued in the earlier version.

[23] Factor income is defined as pre-transfer *and* pre-tax income, and includes wages income from self-employment, income from ownership of physical and financial capital and gifts. Factor income also includes public pensions. Gross income, in turn, equals factor income *plus* social insurance transfers, which includes sick pay, disability pay, social retirement benefits, child or family allowances, maternity pay, military or veterans benefits and near-cash benefits. Gross income *minus* mandatory employee contributions *minus* income tax equals disposable income. See the Luxembourg Income Study for the variable definitions at

Milanovic provides data for 24, mostly OECD countries with a total of 79 observations. As there are not enough observation for the initial year 1970, I take averages for these differences in Gini coefficients in terms of the Gini coefficient for factor incomes over the sample period 1970-2000. The resulting variable is called $RRED$ and is taken to proxy the model's λ. Thus, $RRED = \frac{Gini^{FI} - Gini^{DI}}{Gini^{FI}}$ where FI denotes factor income and DI represents disposable income.

To proxy for the tax rate τ I have used data from the OECD Revenue Statistics 1965-2001, OECD (2002). For these years the ratio of tax revenue to GDP is provided for 30 countries. Here we focus on the initial tax rate in 1970 and call that variable $TAX70$.

In order to link up with studies that have identified robust regressors in cross-country work I have followed Sala-i-Martin (1997) and used male primary school attainment in 1960, called $MSCHOOL60$, and life expectancy in 1960, $LIFEEXP60$, as (robust) control variables. Both are taken from Table 10.1, Barro and Sala–i–Martin (1995).

Finally, long-run growth rates for the period 1970-2000, called GR, and the logarithm of initial income in 1970, $LNY70$, were calculated using the Penn World Table (Mark 6.1) provided by Heston, Summers, and Aten (2002).

The following table provides information on some descriptive statistics for the resulting sample of 29 countries.

Table 2: Descriptive Statistics

Variable	Mean	Std. Dev.	Minimum	Maximum	No. of Obs.
GA	0.713	0.166	0.44	1.01	29
GA^2	0.535	0.240	0.20	1.02	29
GR	0.023	0.010	0.01	0.06	29
$TAX70$	0.290	0.066	0.16	0.39	23
$RRED$	0.337	0.096	0.16	0.51	21
$LNY70$	9.243	0.452	7.93	9.92	27
$MSCHOOL60$	6.445	1.988	2.41	9.76	26
$LIFEEXP60$	68.760	4.460	54.20	73.40	25

For the variables of interest, τ, λ and A, this means that the average tax-revenues-to-GDP ratio in 1970, $TAX70$, is about 30 percent with a standard deviation of 7 percentage points. Redistribution $RRED$ is such that government intervention by means of taxes and transfers amounts to 34 percent of the pre-tax inequality. Finally, in the sample the average country features a level of

http://www.lisproject.org/techdoc/summary.pdf.

productivity, GA, that reaches roughly 71 percent of the level that pertains to the United States of America.[24]

The next table presents the pairwise correlations between the variables.[25]

Table 3: Pairwise Correlations

	GA	GA^2	GR	$TAX70$	$RRED$	$LN70$	$MSCHOOL60$
GA^2	0.994 (29)	1					
GR	0.108 (29)	0.097 (29)	1				
$TAX70$	0.096 (23)	0.058 (23)	−0.122 (23)	1			
$RRED$	−0.412 (21)	−0.427 (21)	−0.060 (21)	0.609 (17)	1		
$LN70$	0.133 (27)	0.140 (27)	−0.589 (27)	0.309 (23)	−0.245 (19)	1	
$MSCHOOL60$	−0.206 (26)	−0.198 (26)	−0.341 (26)	0.296 (22)	0.178 (18)	0.514 (26)	1
$LIFEEXP60$	−0.004 (25)	0.017 (25)	−0.545 (25)	0.607 (21)	0.082 (18)	0.797 (25)	0.536 (25)

Number of observations in brackets.

For the variables of interest the theoretically derived signs for the correlations are generally borne out by the data.[26] For example, the correlation between $TAX70$ and GA is positive (0.096) and that between $RRED$ and GA is negative (-0.412). Furthermore, the correlations show a positive (simple, uncontrolled) relationship between taxes, $TAX70$, and redistribution, $RRED$. In addition, the correlations between initial income and $TAX70$, $RRED$ and GA all show the expected signs.

Furthermore, the correlation between the tax rates, redistribution and growth,

[24]Thus, the data confirm findings that show that the United States is among the most productive countries in the world. See e.g. Trefler (1995) and many others. In the sample countries like France (0.82), Germany (0.75), Japan (0.52) or the United Kingdom (0.90) feature lower values for multi-factor productivity - after controlling for public inputs in production. In that sense the U.S. is an exception. Interestingly, Iceland features a value greater than unity, suggesting that it is more productive than the U.S. Of course, values of A lower than unity also indicate that public inputs in production and the other factor inputs contribute more to explanations of output in these countries than in the U.S. That is because A can also be taken as an indicator of our "ignorance" about other, exogenous factors bearing on aggregate production. Thus, the percentage of exogenous, "unexplained" factors describing aggregate output is larger in the U.S. and Iceland than, for example, in Germany or Japan.

[25]The associated pairwise covariances are reported in table 5 at the end of the paper, where one also finds plots of $TAX70$, $RRED$ and the growth rate against GA.

[26]Notice that the simple correlations between taxes as well as redistribution and growth is negative. The latter would support the claim that policy is exogenous, as will be shown below.

and the squared GAs, called GA^2 support the model's prediction that τ and γ are convex, whereas λ is concave in A. See appendix E.

Therefore, the signs of the pairwise correlations seem to support the model's theoretical predictions.

3.2 Experiments

For the empirical checks in the previous section I have concentrated on pairwise correlations as they use as much information from the data as possible. Thus, we get correlations for a maximum of 29 countries and a minimum of 17 countries for the correlation between tax rates and redistribution. This is due to missing observations.

For estimation, which is the focus in the next sections, we would then only have 17 observations when including tax rates and redistribution in a sample. However, in order to continue to use the additional information that is contained in the rest of the sample, I will now make the counterfactual assumption that the covariance matrix is based on an equal number of observations for every entry in the matrix. I, thus, assume for the thought experiments below that the pairwise correlations correspond to the correlations of every element in a sample with thirty observations with no missing values. With this assumption I now turn to the implication of the model and the data for cross-country research. I will call the subsequent experiments "empirical exercises".

4 Growth Empirics

According to the theory the empirical, long-run relationship between growth and (endogenous) policy for a country i is of the form

$$\gamma_i = f(\tau_i(A_i), \lambda_i(A_i), A_i) = h(A_i) \tag{14}$$

where $f(\cdot)$ is a non-linear function of A_i which is assumed to be country-specific, that is, independent and thus uncorrelated across countries.[27] Notice that (14)

[27] Under the assumption of exogenous policy $\gamma_i = g(\tau_i, \lambda_i, A_i)$ where τ_i and λ_i are independent of the other variables included in $g(\cdot)$. Notice that $h(\cdot)$ and $g(\cdot)$ may be observationally equivalent when particular A_is lead to the same growth rate under either assumption. Thus, assume that empirical and theoretical researchers agree that the Data Generating Mechanism

represents an equilibrium relationship, because output grows at the same rate as the capital stock. The growth rate of the latter, in turn, depends on policy. Thus, the theoretical model implies that policy bears on factor accumulation, which, in turn bears on output growth.

Equation (14) represents a possibly highly non-linear relationship between A and γ. However, the model implies that this non-linearity can be separated out by means of the variables for tax rates τ and redistribution λ. Thus, instead of trying to find the appropriate empirical form of the non-linear association between growth and productivity directly, one can use the information provided by tax rates and redistribution in order to disentangle the direct and indirect channels through which A non-linearly influences the growth rate.

A simple form to use this information in cross-country research is to take a linear approximation of the growth rate γ_i in (14) for small positive changes in A_i around a (sample) mean $\overline{A}$, that is, $dA \equiv A_i - \overline{A} > 0$ and small, as follows

$$\tilde{d}\gamma_i = \frac{\partial f}{\partial \tau_i}\frac{\partial \tau_i}{\partial A_i}\, dA_i + \frac{\partial f}{\partial \lambda_i}\frac{\partial \lambda_i}{\partial A_i}\, dA_i + \frac{\partial f}{\partial A_i}\, dA_i$$

where $\tilde{d}\gamma_i = \gamma_i - \gamma(\overline{A})$. Then $\tilde{d}\tau_i = \frac{\partial \tau_i}{\partial A_i}\, dA_i$ and $\tilde{d}\lambda_i = \frac{\partial \lambda_i}{\partial A_i}\, dA_i$. Notice that $\tilde{d}\tau_i \geq 0$ and $\tilde{d}\lambda_i < 0$ since $\frac{\partial \tau_i}{\partial A_i} \geq 0$ and $\frac{\partial \lambda_i}{\partial A_i} < 0$. We also have $\frac{\partial f}{\partial \tau_i} \leq 0$, where the inequality is strict if the government does not choose a growth maximizing policy. Furthermore, $\frac{\partial f}{\partial \lambda_i} \leq 0$ and $\frac{\partial f}{\partial A_i} > 0$. These signs follow from the results presented in table 1. Thus, we can express the approximation as

$$\tilde{d}\gamma_i = \frac{\partial f}{\partial \tau_i}\, \tilde{d}\tau_i + \frac{\partial f}{\partial \lambda_i}\, \tilde{d}\lambda_i + \frac{\partial f}{\partial A_i}\, dA_i.$$

This looks very similar to a regression equation of a linear estimable model. However, the differentials are taken around A and so around $\gamma(\overline{A}), \tau(\overline{A})$ and $\lambda(\overline{A})$. But a regression passes through the point of means, that is, $\overline{\gamma}, \overline{\tau}$ and $\overline{\lambda}$. These means may not correspond to the variables as functions of the mean of A. To reflect the qualitative information of the linear approximation of the theoretical model in a linear estimable model implies that we must use the theoretical properties of the model to establish a qualitative congruence between the theo-

(DGP) is given by the joint probability distribution $D(\gamma, \tau, \lambda, A)$, which is expressed in terms of steady state variables and, thus, ignores any time dependence. That reflects the procedure to take time-averages of data which are considered of interest.

retical and the estimable model. In appendix (D) it is shown that the resulting regression for our model with endogenous policy would look like

$$\gamma_i = \beta'_c + \beta'_\tau \; \tau_i(A_i) + \beta'_\lambda \; \lambda_i(A_i) + \beta'_A \; A_i + \epsilon'_i \tag{15}$$

where ϵ_i is a country-specific disturbance term.[28] In the appendix it is also shown that from theory we would then expect a non-positive sign for β'_τ, a *positive* sign for β'_A and for β'_λ. The latter property is due to the fact that policy is endogenous. Redistribution lowers growth in the model, but governments optimally redistribute less when their economies are more efficient. Thus, the combined effect is such that the association between growth and redistribution is predicted to be positive in a cross-section.

The latter prediction is markedly different from what one obtains when one assumes that policy is exogenous. In this case the model implies a non-linear relation

$$\gamma_i = g(\tau_i, \lambda_i, A_i) \tag{16}$$

where τ_i, λ_i and A_i are all assumed to be statistically independent of each other.[29] Taking a linear approximation of (16) around the mean values of τ_i, λ_i and A_i yields

$$d\gamma_i = \frac{\partial g}{\partial \tau_i} \; d\tau_i + \frac{\partial g}{\partial \lambda_i} \; d\lambda_i + \frac{\partial g}{\partial A_i} \; dA_i.$$

The theoretical model implies $\frac{\partial g}{\partial \tau_i} \leq 0$, $\frac{\partial g}{\partial \lambda_i} < 0$ and $\frac{\partial g}{\partial A_i} > 0$. If we consider small positive changes in the exogenous variables we have $d\tau_i > 0$, $d\lambda_i > 0$ and $dA_i > 0$. Thus, the linear approximation could be estimated by OLS according to

$$\gamma_i = \beta''_c + \beta''_\tau \; \tau_i + \beta''_\lambda \; \lambda_i + \beta''_A \; A_i + \epsilon''_i.$$

We would then expect a non-positive sign for β''_τ and a *negative* sign for β''_λ and a positive sign for β''_A. Thus, under exogenous policy tax rates and redistribution

[28] This disturbance term would in general be a complicated, non-linear function of some underlying, normally distributed error term, which one would have to know for hypothesis testing.

[29] Of course, that is rarely the case in reality. In fact, the paper's data suggest that these variables are all linked to some extent.

should be negatively associated with long-run growth.

Hence, the question whether policy is endogenous or exogenous may lead to a sign reversal of the (theoretically) "expected" coefficients in a linear empirical model.

5 Empirical Exercise I

We have seen that it may matter for a growth regression whether one assumes policy is endogenous or exogenous. In order to get an empirical picture of the model's predictions I have used the paper's pairwise covariances, now taken to be based on 30 observations with no missing values, and run the regression

$$gr = \beta_1^0 \; tax70 + \beta_2^0 \; rred + \beta_3^0 \; ga + \nu^0,$$

where the lower case letters refer to our variables expressed in *mean deviation* form.[30] The signs of the estimated coefficients were[31]

$$\begin{bmatrix} \hat{\beta}_1^0 & \hat{\beta}_2^0 & \hat{\beta}_3^0 \end{bmatrix} = \begin{bmatrix} -0.0255 & 0.0123 & 0.0102 \end{bmatrix}. \tag{17}$$

Concentrating on these point estimates suggests that they seem to lend some support to the model's prediction that one may expect a positive coefficient for the effect of redistributive transfers on growth under the assumption that policy is endogenous. In view of the paper's arguments this estimated coefficient may support the hypothesis that policy is indeed endogenous.

6 Relation to standard research

An analysis of exogenous, once-and-for-all changes in A_i is similar in spirit to models with exogenous technological change, which is commonly thought to be unobservable.[32] Thus, in line with most studies A_i is now taken to be unob-

[30] For example $gr = GR - \overline{GR}$, where $\overline{GR}$ represents the arithmetic mean of GR.

[31] A "$\hat{\ }$" denotes the OLS estimate for the coefficient β_i^j. Further properties of the estimated coefficients, including inferential statistics, which are *not* the focus in this paper, are reported and discussed in more detail in the online appendix. The superscript on the βs indicates which model is being contemplated.

[32] The discussion about the *Solow-Residual* reflects these difficulties. See, for instance, Barro and Sala–i–Martin (1995), chpt. 10.4.

servable, implying that information on A_i would be contained in the disturbance term and would not feature separately in the regressions.

In most cross-country growth research policy variables are included in regressions as separate regressors, implying that they are assumed to be exogenous explanatory variables. A standard justification for treating policy as exogenous is a randomization argument. For example, Barro (1989) argues that in a large sample public policies may be treated as randomly generated. That comes close to saying that policies are exogenous. But in light of this paper's analysis the argument would not hold. Even if all countries had different governments with different welfare functions so that policies looked randomly chosen, the model predicts that *all* policies would be influenced by the same fundamental economic variables included or not included in the regressions. The paper concentrates on exactly that problem.

For that reason we will now concentrate on situations where one is aware of the fact that policy is endogenous and depends on A, but A is treated as unobservable. Furthermore, we will point out differences in interpreting estimated coefficients when one assumes that policy is exogenous.

When A is taken as unobservable the second-best, but operationally viable model for the regression in (15) would be

$$\gamma_i = \beta_c + \beta_\tau \ \tau_i(A_i) + \beta_\lambda \ \lambda_i(A_i) + v_i \tag{18}$$

where $v_i = v_i(A_i, \epsilon_i)$ is a country-specific disturbance term which depends positively on A_i and also on ϵ_i. The latter is assumed to be uncorrelated with A_i as well as with each of the regressors, and $E(\epsilon_i) = 0$.

If that model is estimated by OLS, multicollinearity and the omission of a relevant variable will be a problem. Thus, reported $t-$statistics do not report the true significance levels and statistical inferences are not really possible. However, here the focus is on estimation and the problem caused by assuming that policy is exogenous.

In order to see what endogenous policy and biases due to the omission of A_i imply, assume that for the estimation of (18) we use OLS and transform our data to *mean deviation form*. This allows us to calculate expressions for the estimators $\hat{\beta}_j$ where $j = \tau, \lambda$. Then the *expected bias*, called b_j, which bears on $\hat{\beta}_j$ obeys $\hat{\beta}_j = \beta_j + b_j$ where β_j is the true estimator. Thus, $\hat{\beta}_j - \beta_j = b_j$ is the expected bias. It is then pretty straightforward to show that the biases are given by

$$\begin{bmatrix} b_\tau \\ b_\lambda \end{bmatrix} = \begin{bmatrix} a_{\tau\tau} & a_{\tau\lambda} \\ a_{\tau\lambda} & a_{\lambda\lambda} \end{bmatrix}^{-1} \begin{bmatrix} a_{\tau\upsilon} \\ a_{\lambda\upsilon} \end{bmatrix} \tag{19}$$

where $a_{\tau\tau}$, $a_{\lambda\lambda}$ denote the variances, and $a_{\tau\lambda}, a_{\tau\upsilon}, a_{\lambda\upsilon}$ represent the covariances of τ, λ and υ.[33]

One easily establishes that the covariance matrix above is positive definite and has a positive determinant denoted by D. Using Cramer's Rule we can calculate the expected biases as

$$b_\tau = D^{-1}\left[a_{\lambda\lambda}\ a_{\tau\upsilon} - a_{\tau\lambda}\ a_{\lambda\upsilon}\right] > 0 \quad \text{and} \quad b_\lambda = D^{-1}\left[-a_{\tau\lambda}\ a_{\tau\upsilon} + a_{\tau\tau}\ a_{\lambda\upsilon}\right] < 0$$

since $a_{\lambda\lambda}, a_{\tau\tau}, a_{\tau\lambda}, a_{\tau\upsilon} > 0$, and $a_{\lambda\upsilon} < 0$. This establishes the following: If (18) is estimated, but (15) is the linear form of the 'true' model (14), then we get an upward bias, and thus an overestimation of the effect of taxes on growth. If we assume that in a large sample countries are not all acting growth maximizing then - according to the model - we should be on the downward sloping branch of the inverted U-shaped relationship between taxes and growth. Thus, we would expect $\beta_\tau < 0$, that is, a negative point estimate for the effect of taxes on growth, under exogenous or endogenous policy. If we find a positive one, it would not invalidate the theory as the true estimate may still be negative. Notice that the model implies this may apply to all tax rate variables that are related to the tax base.

In turn, we get a downward bias, and thus, an underestimation of the effect of redistribution on growth. As the model lets us expect a negative point estimate, $\beta_\lambda < 0$, when assuming that policy is exogenous, the underestimation implies that the hypothesis that redistribution is bad for growth is inherently untestable, when assuming that policy is exogenous, although it is endogenous. This is because the estimate can in principle be biased towards minus infinity. Thus, when we find a negative point estimate this cannot be taken to corroborate the theory that redistribution is bad for growth.[34] In turn, under endogenous policy we expect

[33] Thus, e.g., $a_{\tau\tau} = \frac{1}{N}\sum_i^N(\tau_i - \bar{\tau})^2$ and $a_{\tau\lambda} = \frac{1}{N}\sum_i^N(\tau_i - \bar{\tau})(\lambda_i - \bar{\lambda})$ where bars over the variables denote sample means.

[34] Further evidence that one may get a negative point estimate, when dropping GA from

$\beta_\lambda > 0$ so that a negative bias would not invalidate a "true" positive effect, even if the estimated coefficient was found to be negative.

Instead of pursuing the implications of the 'true' model any further, we now concentrate on the predicted covariances as implied by the model and relate to standard procedures in the growth empirics literature.[35]

6.1 Relation to Barro-Regressions

A typical cross-country growth regression that analyzes the effect of policy on output growth takes the form

$$\gamma_i = P\beta + X\delta + \zeta y_0 + v'(A_i, \epsilon_i). \tag{20}$$

Here P is a row vector of policy variables, X a row vector of control variables and y_0 is the logarithm of initial income. This formulation goes back to e.g. Barro (1991), Mankiw, Romer, and Weil (1992), Islam (1995), or Caselli, Esquivel, and Lefort (1996). The inclusion of initial income is due to the hypothesis that conditional on their steady states initially poorer countries have higher subsequent growth. Often initial income is found to be a robust regressor, that is, it is found to be statistically significant in many different model variations. See, for example, Levine and Renelt (1992) and Sala-i-Martin (1997).

First notice that conditioning on variables capturing accumulation like investment or investment in terms of GDP in a regression that includes policy variables will find that policy variables have no explanatory power, if policy is taken to bear primarily on factor accumulation and factor accumulation determines output growth.[36]

Next, suppose we wanted to test the predictions of the theoretical model using

the regression, is provided by the results of the coefficients for redistribution in model (23) in comparison to those for model (17) below. The former is positive in this paper but a lot lower and closer to zero than when including GA as a regressor.

[35] The theoretical model is supposed to capture essential features of the relationship between fiscal policy and growth. Thus, based on the theory one may derive the signs of covariances with A and the implications for possible biases for other fiscal variables like the ratio of redistributive transfers to GDP or public investment to GDP. This has been done in Rehme (2002b).

[36] The same applies to growth accounting equations where output growth is regressed on capital and labour growth, and other variables. See, for instance, Benhabib and Spiegel (1994). If these other variables include policy variables like τ or λ which bear on accumulation, then policy variables will be found to have no explanatory power in regressions for a cross-section.

the Barro set-up. For simplicity, assume that the variables in X are uncorrelated with the policy variables in P and with (the logarithm of) initial income y_0. This would be a rather desired property for estimation. Furthermore, for data in mean deviation form we will look at the following two model variants as examples that capture the essential features what is being done in the literature.

$$d\gamma_i = \beta^2_\tau \; d\tau_i + \beta^2_{x_i} \; dx_i + \beta^2_y \; dy_0 + v'(A_i, \epsilon_i) \tag{21}$$

$$d\gamma_i = \beta^3_\tau \; d\tau_i + \beta^3_\lambda \; d\lambda_i + \beta^3_y \; dy_0 + v''(A_i, \epsilon_i) \tag{22}$$

The first model in equation (21) represents in an abbreviated form an example of what authors like Levine and Renelt (1992) or Sala-i-Martin (1997) do when they study model uncertainty and assume that policy is exogenous. The typical procedure is to identify robust regressors like initial income y_0 and some other control variables like x (e.g. life expectancy or an indicator of human capital) to check whether adding variables of interest (here τ) are associated with growth in a statistically significant way.

The second model in equation (22) draws on a simple version of regressions that are derived from the Solow model. This approach has primarily been popularized by Mankiw, Romer, and Weil (1992) (MRW). Here the coefficients on τ and λ measure the effect on the steady state growth rate, whereas the coefficient on y_0 measures (conditional) $\beta-$convergence, that is, how far countries are from their long-run positions. The expectation is that β_y is negative, that is, initially poorer countries should exhibit higher subsequent growth.

One may then derive expressions for the expected biases for both of these model variants, when allowing for endogenous policy. It turns out, however, that one cannot unambiguously sign these biases from the theoretical model. Thus, a theoretical model that incorporates many features found empirically allows one to sign the biases only up to a certain point. For more complicated empirical models, however, one has to check numerically the signs of biases using the data at hand.

7 Empirical Exercise II

In this section empirical evidence is provided that serves to show the *direction* of the biases. For simplicity it is assumed that the covariances in Table 5 are all based on the same number of observations. That allows one to ignore problems of missing values that may have an impact on estimation. Against this assumption I will now analyze various empirical models that are all linear.

Let us start with the simplest model, called model 1,

$$gr = \beta_1^1 \; tax70 + \beta_2^1 \; rred + \nu^1$$

where ν is the country specific error term that depends on A. Again lower case letters indicate that the variables are in mean deviation form. Running the regression yields the following estimates

$$\begin{bmatrix} \hat{\beta}_1^1 & \hat{\beta}_2^1 \end{bmatrix} = \begin{bmatrix} -0.0145 & 0.0008 \end{bmatrix}. \tag{23}$$

Comparing the estimated coefficients for model (17), with GA included, with those of model (23) yields that

$$0 > \hat{\beta}_1^1 > \hat{\beta}_1^0 \quad \text{and} \quad 0 < \hat{\beta}_2^1 < \hat{\beta}_2^0.$$

Hence, the measured effect of taxes is more negative, and the measured effect of redistribution is more positive in model (17) than in model (23) where GA is missing as a regressor.

Many cross-country studies have found that fiscal policy variables do not affect growth in a statistically significant way. This really means that the estimators for, for instance, the coefficients β_1^1 and β_2^1, called $\hat{\beta}_1^1$ and $\hat{\beta}_2^1$, are assuming values that are statistically close to zero. From this it is then often concluded that the "true" coefficients β_1^1 and β_2^1 are likely to be zero. Clearly for this conclusion to hold it is assumed that the estimated coefficients are unbiased.

However, biases due to the omission of A may render such a conclusion invalid. To see what the biases imply assume that the estimated coefficients are close to zero. For simplicity assume that $\hat{\beta}_1^1 = 0$ and $\hat{\beta}_2^1 = 0$.

This assumption captures the following thought experiment: Under the assumption that policy is exogenous and the estimated coefficients for the effect of policy on growth are unbiased it is often found that they are statistically insignificant. Thus, one cannot rule out the possibility that the "true" coefficient might be zero under that assumption. Now we argue that the estimated coefficients are biased und that policy is endogenous. That entails that the "true" coefficients are related to the biases and, in general, are *not* equal to zero. In order to see clearly what the biases imply in relation to any "true" coefficient, the estimated coefficient is taken to be zero. This would then be similar to the finding that the estimated coefficient is statistically insignificant, when assuming unbiasedness and exogenous policy, and when taking the "true" coefficient to be zero.[37]

Again denote the biases by $b^i_j = \hat{\beta}^i_j - \beta^i_j$, where i represents the model under study and j indicates which regressor the coefficient pertains to.

When $\hat{\beta}^1_1 = 0$ and $\hat{\beta}^1_2 = 0$ the biases are given by $b^1_1 = -\beta^1_1$ and $b^1_2 = -\beta^1_2$. Thus, the "true" coefficients are directly related to the biases and usually not equal to zero. Using the covariances from Table 5 allows one to calculate the biases. They are given by

$$\begin{bmatrix} b^1_1 \\ b^1_2 \end{bmatrix} = \begin{bmatrix} 0.00431 & 0.00324 \\ 0.00324 & 0.00922 \end{bmatrix}^{-1} \begin{bmatrix} 0.00099 \\ -0.00690 \end{bmatrix} = \begin{bmatrix} 1.07671 \\ -1.12674 \end{bmatrix} \tag{24}$$

The first thing to notice is that the biases for both variables are quite large.

For instance, there is a huge overestimation of the effect of taxes on growth in this model. Given the positive bias and a point estimate that is taken to be zero implies that the taxes really co-vary *negatively* with growth. Quantitatively this means that an increase of one standard deviation in the ratio of tax revenues to GDP (approx. $\sqrt{0.00431} = 0.066$, that is, 6.6 percentage points) would lower growth by roughly $1.08 * \frac{0.066}{0.010} \approx 7$ standard deviations of the growth rate (0.01), that is, by approximately 0.07 so 7(!) percentage points under the assumption that the estimated coefficient is zero.[38] Thus, the presence of the bias seems to

[37]Thus, in the following thought exercise we assume that the "true" coefficient is *not* equal to zero, but related to the bias, under the assumption of endogenous policy.

[38]Here we use the concept of a $\beta-$coefficient by which $\beta^* = \hat{\beta}_{yx}\left(\frac{\sigma_x}{\sigma_y}\right)$ where $\hat{\beta}_{yx}$ denotes the OLS estimate in a regression of y on x, and σ_i denotes the standard deviations of $i = x, y$. Then β^* tells one how a 1 standard deviation change in x leads to a β^* deviation change in y. On this see, for example, Pindyck and Rubinfeld (1998), p. 98.

be a non-trivial problem for model 1.

For redistribution we find a negative bias. Thus, we can expect that redistribution really co-varies *positively* with growth under the maintained assumption that the estimated coefficient is zero. The quantitative implication is that according to model 1 a one-standard-deviation change in $RRED$ (approx. $\sqrt{0.00922} \approx 0.096$) would change growth by roughly $-1.13 * \frac{0.096}{0.010} \approx -10.1$ standard deviations of the growth rate (0.010), that is, it would really raise growth by 0.10 so 10(!) percentage points under the maintained assumption.

Of course, the quantitative effects are only so strong because they hold under the assumption that the estimated coefficients are (close to) zero. The magnitude of the biases seems very high, however. But what is of main interest in this context is that the *direction* of both biases confirms what was predicted theoretically.[39]

Next, we contemplate a simple model in the spirit of Mankiw, Romer, and Weil (1992) by adding initial income as an additional regressor. The expectation is that the biases will be reduced by adding more regressors. In particular, our model 2 is given by

$$gr = \beta_1^2\ tax70 + \beta_2^2\ rred + \beta_3^2\ lny70 + \nu^2$$

With the covariances from Table 5 I have calculated the biases using MATLAB. They are given by

$$\begin{bmatrix} b_1^2 & b_2^2 & b_3^2 \end{bmatrix}' = \begin{bmatrix} 1.146 & -1.173 & -0.027 \end{bmatrix}' \tag{25}$$

Taking account of the biases implies that $TAX70$ is really expected to co-vary *negatively* with growth and the association between redistribution $RRED$ and growth is *positive*. The quantitative effects are as follows.

A one-standard-deviation change in the tax rate (6.6 percentage points) reduces growth by 7.6 percentage points when one assumes that the estimated coefficients are zero. This effect is not negligible in the long run. Similarly, changing redistribution by one standard deviation (0.096) would raise growth by 11.2 percentage points, which is clearly a strong effect.

Comparing the models 1 and 2 yields that the addition of initial income as

[39]Recall it was found that the coefficients on τ should be biased upwards and that for λ should be biased downwards.

a further regressor reduces the biases and, therefore, mitigates their quantitative impact on any measured effect of the two fiscal policy variables on economic growth.

The last model is related to Levine and Renelt (1992) and Sala-i-Martin (1997). These studies identify robust regressors like male school attainment, MSCHOOL60, initial income, LNY70, and life expectancy, LIFEEXP60, and then add variables of interest to see if they are statistically significant. Their robustness checks are an important step forward in identifying relevant determinants of growth. These authors usually find that policy variables are non-robust regressors that are associated with long-run growth in a statistically insignificant way. To relate to these works we now contemplate model 3.

$$gr = \beta_1^3\ tax70 + \beta_2^3\ rred + \beta_3^3\ lny70 + \beta_4^3\ mschool60 + \beta_5^3\ lifeexp60 + \nu^3$$

The associated biases are now given by

$$\begin{bmatrix} b_1^3 & b_2^3 & b_3^3 & b_4^3 & b_5^3 \end{bmatrix}' = \begin{bmatrix} 1.267 & -1.066 & 0.070 & -0.024 & -0.005 \end{bmatrix}' \qquad (26)$$

Again the biases are not small. For instance, a one-standard-deviation change in $TAX70$ (6.6 percentage points) reduces growth by 8.4 percentage points. In turn, changing redistribution, $RRED$, by one standard deviation (0.096) increases growth by roughly 10.2 percentage points. Both of these effects hold when one assumes that the estimated coefficients are zero.

Moving from model 2 to 3 one might expect that the bias problem is reduced by adding more regressors. But, as the example shows, that is not necessarily the case, because the bias for $TAX70$ is larger in model 3 than in model 2. This suggests that a criterion for selecting empirical models on policy and growth is to look for models that minimize the bias problem.

From all this one may conclude that the presence of biases appears to make the estimated coefficients quite imprecise. This has, of course, implications for any t-statistic so that arguments based on them may surely be problematic.

Furthermore, this paper's data, when used for different model specifications, provide suggestive evidence that the coefficients on tax rate variables that are related to the tax base (the model's τ) appear to be biased upwards. That means that one should expect tax rate variables such as the ratio of tax revenue to

GDP to co-vary negatively with growth. On the other hand, the data provide suggestive evidence that redistributional variables might be biased downwards. This means that the hypothesis that redistribution slows down growth may be inherently untestable. This finding would call models into question that argue that across countries redistribution is bad for long-run growth.

8 Conclusion

Within a common theoretical framework it is shown how policy affects investment which in turn affects output growth. In the model optimizing governments take account of fundamental economic variables relating to institutions, technology and cultural features when making their decisions. This makes public policy economically endogenous and has interesting effects on long-run growth. Several findings of the paper are noteworthy.

First, growth regressions which study the effect of policy on growth should not be conditioned on variables for factor accumulation. This is because policy may work through investment so that including policy and investment variables would yield misleading results.

Second, when policy is endogenous and an important economic fundamental like productivity is omitted in growth regressions, the estimated coefficients on policy variables are generically biased. This has important implications for arguments based on statistical significance.

Third, for different empirical models the signs of the biases are analyzed theoretically and empirically. It is found that the coefficients on tax rate variables related to the tax base are generally biased upwards and those for redistribution are generally biased downwards.

For the latter this implies the following: If policy is economically endogenous, the effect of redistributive transfer variables on growth are generally underestimated so that the hypothesis that redistribution is bad for growth may not be testable. The downward bias is, however, perfectly consistent with empirical findings in the literature which find a positive association between redistributive transfers and growth. It may also represent corroboration of the hypothesis that redistribution is not bad for growth.

The paper argues that more work is needed for the disentanglement of the interplay of long-run economic fundamentals and policy on the one hand, and policy and growth on the other. Furthermore, paying more attention to the bias problem in growth empirics may be worthwhile. This should not be too difficult and would base some findings on a sounder footing. This paper has moved in that direction and the analysis would imply that we may still learn "something" from "regressing economic growth on policy".

A The capitalists' optimum

The current value Hamiltonian for the problem (5) - (7) is given by

$$H = \ln C_t^k + \mu_t((r_t - \tau)K_t - C_t^k). \tag{A1}$$

The necessary first order conditions for its maximization are given by (6), (7) and

$$\frac{1}{C_t^k} - \mu_t = 0 \tag{A2a}$$

$$\dot{\mu_t} = \mu_t\rho - \mu_t\,(r_t - \tau) \tag{A2b}$$

$$\lim_{t\to\infty}\ K_t\mu_t e^{-\rho t} = 0, \tag{A2c}$$

where the positive co-state variable μ_t represents the instantaneous shadow price of one more unit of investment at date t.

From (A2a) and (A2b) consumption grows at $\gamma \equiv \frac{\dot{C}_t^k}{C_t^k} = R_t - \rho$ where $R_t \equiv (r_t - \tau)$. In equilibrium R_t is constant. Thus, $C_t = C_0\ e^{(R-\rho)t}$ where C_0 remains to be determined. Substituting for C_t in (6) implies $\dot{K}_t = R\ K_t - C_0\ e^{(R-\rho)t}$ which is a first order, linear differential equation in K_t and solved as follows

$$\begin{aligned} \dot{K}_t - R\ K_t &= -C_0\ e^{\gamma t} \\ e^{-Rt}\left(\dot{K}_t - R\ K_t\right) &= -e^{-Rt}\ C_0\ e^{\gamma t} \\ \int e^{-Rt}\left(\dot{K}_t - R\ K_t\right) dt &= -\int C_0\ e^{-\rho t} dt. \end{aligned}$$

The last equation is an exact differential equation with integrating factor e^{-Rt}. The LHS is solved by $K_t\ e^{-Rt} + b_0$ and the RHS is solved by $\frac{C_0}{\rho}\ e^{-\rho t} + b_1$, where b_0, b_1 are arbitrary constants. Thus, $K_t = \frac{C_0}{\rho}\ e^{(R-\rho)t} + b\ e^{Rt}$ where $b = b_1 - b_0$. Substituting this into the transversality condition implies

$$\frac{1}{C_0}\lim_{t\to\infty}\left(\frac{C_0}{\rho}\ e^{(R-\rho)t} + b\ e^{Rt}\right)\ e^{-Rt} = \lim_{t\to\infty}\left(\frac{1}{\rho}\ e^{-\rho t} + \frac{b}{C_0}\right) = 0$$

which holds if the arbitrary constant b is set equal to zero. Then $K_t = \frac{C_0}{\rho}\ e^{\gamma t} \ \Rightarrow\ \gamma_k = \gamma = R - \rho$ so that consumption and wealth grow at the same constant rate in the optimum. Furthermore, the optimal level of consumption at each date is given by $C_t = \rho K_t$.

B Optimal Policies

The government solves: $\max\limits_{\tau,\lambda} \ (1-\beta)\,V^r + \beta\,V^l$ s.t. $\lambda \geq 0$ where β is the social weight attached to the workers' welfare. The FOCs are

$$\beta\,\frac{\eta_\tau + \lambda}{(\eta + \lambda\tau)\rho} + \frac{\gamma_\tau}{\rho^2} = 0 \quad , \quad \lambda\left(\beta\,\frac{\eta_\lambda + \tau}{(\eta+\lambda\tau)\rho} + \frac{\gamma_\lambda}{\rho^2}\right) = 0.$$

Notice that γ_τ must be negative for the first equation to hold, so in the optimum $\tau > \hat{\tau}$. Concentrating on an interior solution for λ, simplifying, rearranging and division of the resulting two equations by one another yields

$$\frac{\eta_\tau + \lambda}{\eta_\lambda + \tau} = \frac{\gamma_\tau}{\gamma_\lambda}. \tag{B1}$$

Then $\gamma_\lambda = r_\lambda$ and $\gamma_\tau = r_\tau - 1$ imply $(\eta_\tau + \lambda) r_\lambda = (\eta_\lambda + \tau)(r_\tau - 1)$ which upon multiplying out becomes $\eta_\tau r_\lambda + \lambda r_\lambda = r_\tau \eta_\lambda + r_\tau \tau - \eta_\lambda - \tau$. Notice $r_\lambda \eta_\tau = r_\tau \eta_\lambda$ and $\eta = \frac{1-\alpha}{\alpha} r$. Then $\lambda r_\lambda = r_\tau \tau - \frac{1-\alpha}{\alpha} r_\lambda - \tau$ and so

$$\left(\lambda + \frac{1-\alpha}{\alpha}\right) r_\lambda = \tau r_\tau - \tau \quad \Leftrightarrow \quad \left(\lambda + \frac{1-\alpha}{\alpha}\right) = \frac{\tau r_\tau}{r_\lambda} - \frac{\tau}{r_\lambda}.$$

Recall $r_\tau = \alpha E(1-\lambda), r_\lambda = \alpha E(-\tau)$ where $E = (1-\alpha)A[(1-\lambda)\tau]^{-\alpha}$. Thus, $\frac{\tau r_\tau}{r_\lambda} = -\frac{\tau \alpha E(1-\lambda)}{\alpha E \tau} = -(1-\lambda)$ and $\lambda + (1-\lambda) + \frac{1-\alpha}{\alpha} = -\frac{\tau}{r_\lambda} \ \Leftrightarrow \ \frac{r_\lambda}{\alpha} = -\tau$ and so

$$\tau = \frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{1-\lambda}. \tag{B2}$$

Notice that for this τ we have $E = 1$. For the first order condition for τ we note that $\eta = (1-\alpha)A[(1-\lambda)\tau]^{1-\alpha} = E[(1-\lambda)\tau] = [(1-\alpha)A]^{\frac{1}{\alpha}}$. Furthermore, $\eta_\tau = (1-\alpha)(1-\lambda),\ r_\tau = \alpha(1-\lambda)$. Eqn. (B2) implies $\lambda = 1 - \frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{\tau}$ so that

$$\eta + \lambda\tau = [(1-\alpha)A]^{\frac{1}{\alpha}} + \tau\left(1 - \frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{\tau}\right) = \tau.$$

Then the first order condition for τ becomes

$$\beta\,\frac{\eta_\tau + \lambda}{(\eta + \lambda\tau)} = -\frac{\gamma_\tau}{\rho} \quad \Leftrightarrow \quad \frac{\eta_\tau + \lambda}{\tau} = -\frac{\gamma_\tau}{\beta\rho} \quad \Leftrightarrow \quad \frac{\eta_\tau + \lambda}{\gamma_\tau} = -\frac{\tau}{\beta\rho}.$$

But from above $\frac{\eta_\tau+\lambda}{\gamma_\tau} = \frac{(1-\alpha)(1-\lambda)+\lambda}{\alpha(1-\lambda)-1} = -1$ so that $\tau = \beta\rho$. Thus,

$$\tau = \beta\rho \quad \text{and} \quad \lambda = 1 - \frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{\beta\rho}. \tag{B3}$$

which is equation (12) when $\beta = 1$. Recall that these equations hold for $\lambda \geq 0$, thus for $\beta\rho \geq [(1-\alpha)A]^{\frac{1}{\alpha}}$.

Suppose $\lambda = 0$, then the first order condition becomes

$$\frac{\eta_\tau}{\eta} = -\frac{r_\tau - 1}{\beta\rho} \quad \Leftrightarrow \quad \frac{(1-\alpha)E}{\tau E} = -\frac{\alpha E - 1}{\beta\rho} \quad \Leftrightarrow \quad (1-\alpha)\beta\rho = \tau - \alpha\tau E$$

so that the solution with $\lambda = 0$ is given by

$$(1-\alpha)\beta\rho = \tau\left[1 - \alpha(1-\alpha)A\tau^{-\alpha}\right] \tag{B4}$$

which holds only if $\beta\rho < [(1-\alpha)A]^{\frac{1}{\alpha}}$. For $\beta = 1$ this is equation (13) in the text.

If $\beta = 0$, then $\tau = \hat{\tau}$. Thus, the pro-capital policy maximizes growth.

C Reactions under Endogenous Policy

Pro-Capital. $\hat{\gamma} = \frac{\alpha}{1-\alpha}\frac{\hat{\tau}}{} - \rho$ and $\hat{\tau} = [\alpha(1-\alpha)A]^{\frac{1}{\alpha}}$. Clearly, $\frac{d\hat{\tau}}{dA} > 0$, $\frac{d\hat{\gamma}}{dA} > 0$, $\frac{d^2\hat{\tau}}{dA^2} > 0$ and $\frac{d^2\hat{\gamma}}{dA^2} > 0$. Thus, the tax rate and growth are both (strictly) increasing and convex in A under the pro-capital policy.

Redistributing, Pro-Labour. By equation (12) $\check{\tau} = \rho$ so that $\frac{d\check{\tau}}{dA} = 0$. As $\lambda = 1 - \frac{[(1-\alpha)A]^{\frac{1}{\alpha}}}{\rho}$ it follows that $\frac{d\lambda}{dA} < 0$ and $\frac{d^2\lambda}{dA^2} < 0$. Thus, redistribution is (strictly) decreasing and concave in A.

From equation (12) $r = \alpha A[(1-\lambda)\tau]^{1-\alpha} = \frac{\alpha}{1-\alpha}[(1-\alpha)A]^{\frac{1}{\alpha}}$ so that $\frac{d\check{\gamma}}{dA} > 0$ and $\frac{d^2\check{\gamma}}{dA^2} > 0$, that is, the growth rate is (strictly) increasing and convex in A under that policy.

Non-Redistributing, Pro-Labour. For $\lambda = 0$ the optimal tax rate $\check{\tau}$ solves equation (13), that is,

$$z = \frac{\tau}{1-\alpha} - \alpha A\tau^{1-\alpha} - \rho = 0.$$

Let $\tau' \equiv \frac{d\check{\tau}}{dA}$, $\tau'' \equiv \frac{d^2\check{\tau}}{dA^2}$ and let z_{ij}, where $i, j = \tau, A$, denote the partial derivatives of z with respect to τ and A. Then implicit differentiation yields $z_A + z_\tau \cdot \tau' = 0$. We have

$$z_\tau = \frac{1}{1-\alpha} - (1-\alpha)\alpha A \tau^{-\alpha} \quad \text{and} \quad z_A = -\alpha\tau^{1-\alpha}$$

The expression for z_τ is positive for all $\tau > \hat{\tau}$. As $z_A < 0$ it follows that

$$\tau' = \frac{d\check{\tau}}{dA} = -\frac{z_A}{z_\tau} = \frac{\alpha\tau^{1-\alpha}}{\frac{1}{1-\alpha} - (1-\alpha)\alpha A\tau^{-\alpha}} > 0.$$

Differentiating $z_A + z_\tau \cdot \tau' = 0$ again yields

$$z_{AA} + 2 \cdot z_{A\tau} \cdot \tau' + z_{\tau\tau} \cdot (\tau')^2 + z_\tau \cdot \tau'' = 0.$$

For our case and after simplification and rearrangement we have

$$-2(1-\alpha)\alpha\tau^{-\alpha}\tau' + (1-\alpha)\alpha^2 A\tau^{1-\alpha}(\tau')^2 + \frac{\tau''}{1-\alpha} - (1-\alpha)\alpha A\tau^{-\alpha} \cdot \tau'' = 0$$

Simplification then yields

$$\tau'' = \frac{d^2\check{\tau}}{dA^2} = \frac{2(1-\alpha)\alpha\tau^{-\alpha} \cdot \tau' - (1-\alpha)\alpha^2 A\tau^{-1-\alpha} \cdot (\tau')^2}{\frac{1}{1-\alpha} - \alpha(1-\alpha)A\tau^{-\alpha}}. \tag{C1}$$

The sign of τ'' depends on $2 - \alpha A\tau^{-1} \cdot \tau'$. Substituting for τ' from above yields that τ'' is positive for the following reason:

$$\begin{aligned} 2 &> \alpha A\tau^{-1} \cdot \tau' = \alpha A\tau^{-1} \cdot \left(\frac{\alpha\tau^{1-\alpha}}{\frac{1}{1-\alpha} - \alpha(1-\alpha)A\tau^{-\alpha}} \right) \\ 2 &> 2\alpha(1-\alpha)A\tau^{-\alpha} - \alpha^2 A\tau^{-\alpha} \end{aligned}$$

because $\alpha(1-\alpha)A\tau^{-\alpha} < 1$ with $\check{\tau} > \hat{\tau}$. Hence, $\check{\tau}$ is (strictly) increasing and convex under this policy.

For the growth rate one finds $\frac{d\gamma}{dA} = r_A + (r_\tau - 1)\frac{d\tau}{dA}$, that is,

$$\alpha\tau^{1-\alpha} - \tau' + (1-\alpha)\alpha A\tau^{-\alpha} \cdot \tau'. \tag{C2}$$

This is positive if

$$\begin{aligned} \alpha\tau^{1-\alpha} &> \left(1 - \alpha(1-\alpha)A\tau^{-\alpha}\right) \left[\alpha(1-\alpha)\tau\left(\tau^\alpha - \alpha(1-\alpha)^2 A\right)^{-1}\right] \\ \tau^\alpha - \alpha^2(1-\alpha)^2 A &> (1-\alpha)\tau^\alpha - \alpha^2(1-\alpha)^2 A \end{aligned}$$

which is equivalent to $1 > 1-\alpha$ and true. Thus, $\frac{d\tilde{\gamma}}{dA} > 0$ if $\lambda = 0$ in (13).

Differentiating (C2) again yields

$$2\alpha(1-\alpha)A\tau^{-\alpha}\cdot\tau' - \alpha^2(1-\alpha)A\tau^{-1-\alpha}\cdot(\tau')^2 - \tau'' + \alpha(1-\alpha)A\tau^{-\alpha}\cdot\tau''. \tag{C3}$$

Let $\Delta \equiv 2\alpha(1-\alpha)A\tau^{-\alpha}\cdot\tau' - \alpha^2(1-\alpha)A\tau^{-1-\alpha}\cdot(\tau')^2$. Using (C1) the sign in (C3) depends on

$$\Delta - \left[\frac{1-\alpha(1-\alpha)A\tau^{-\alpha}}{\frac{1}{1-\alpha} - \alpha(1-\alpha)A\tau^{-\alpha}}\right]\Delta$$

which is positive because the expression in square brackets is less than one. Hence, $\frac{d\tilde{\gamma}}{dA} > 0$ and $\frac{d^2\tilde{\gamma}}{dA^2} > 0$ so that the growth rate is (strictly) increasing and convex in A under that policy.

D Expected Coefficients under Exogenous and Endogenous Policy

In this appendix I show that, when policy is endogenous, the (theoretically) expected[40], so (theoretically) 'true' regression coefficients may have opposite signs compared to the (theoretically) expected 'true' coefficients under the assumption of exogenous policy, even though the models in question are observationally equivalent.

D.1 Exogenous Policy

Under the assumption that policy is exogenous we have

$$\gamma_i^1 = g(\tau_i, \lambda_i, A_i) \tag{D1}$$

where τ, λ and A are treated as exogenous variables and i indexes country i. The model implies

$$\frac{\partial g}{\partial \tau_i}_{|\tau_i,\lambda_i,A_i} \leq 0, \quad \frac{\partial g}{\partial \lambda_i}_{|\tau_i,\lambda_i,A_i} < 0, \quad \frac{\partial g}{\partial A_i}_{|\tau_i,\lambda_i,A_i} > 0. \tag{D2}$$

[40]By "expected" I do not mean a mathematical expectation. Instead, we look for the (sign) prediction for any (unknown) coefficient in a linear empirical model when translating properties of a (non-linear) theoretical model to an estimable (linear) regression model. On the distinction of theoretical and empirical models see e.g. Spanos (1986), chp. 1.

D.2 Endogenous Policy

When policy is taken to be endogenous we have

$$\gamma_i^2 = f(\tau_i(A_i), \lambda_i(A_i), A_i) = h(A_i) \tag{D3}$$

where τ, λ and are now endogenous variables that depend on A and i again indexes country i. That model implies $\frac{dh}{dA} > 0$ (strictly so), but also

$$\frac{\partial f}{\partial \tau_i}_{|\tau(A_i),\lambda(A_i),A_i} \leq 0, \quad \frac{\partial f}{\partial \lambda_i}_{|\tau(A_i),\lambda(A_i),A_i} < 0, \quad \frac{\partial f}{\partial A_i}_{|\tau(A_i),\lambda(A_i),A_i} > 0. \tag{D4}$$

and we have[41]

$$\frac{\partial \tau_i}{\partial A_i}_{|A_i} > 0 \quad \text{and} \quad \frac{\partial \lambda_i}{\partial A_i}_{|\tau(A_i),A_i} < 0. \tag{D5}$$

and

$$\frac{\partial^2 \tau_i}{\partial A_i^2}_{|A_i} < 0 \quad \text{and} \quad \frac{\partial^2 \lambda_i}{\partial A_i^2}_{|\tau(A_i),A_i} < 0 \tag{D6}$$

All these derivatives never change signs from positive to negative or vice versa in the theoretical model. Furthermore, the latter property implies that $\tau(A)$ and $\lambda(A)$ are strictly *concave* in A.

D.3 Remark

Assume that for any observed (τ, λ, A) models (D1) and (D3) would look the same. Thus, they are observationally equivalent. For observed values of τ, λ, A, we would then have $\gamma^1 \approx \gamma^2$.

D.4 Estimable Models

We shall consider statistical (estimable) linear models that use OLS in order to investigate the theoretical non-linear relationships between γ and (τ, λ, A). Notice that OLS implies that any regression line must pass through the point $\overline{\gamma}, \overline{\tau}, \overline{\lambda}, \overline{A}$, that is, it passes through the sample means of these variables.

[41]In the model the (optimal) tax rates depend on A, but not on λ, whereas redistribution λ depends on A and the tax rate τ.

D.4.1 Exogenous Policy

To convert model (D1) into a regression we take a linear approximation of the theoretical model around the sample means. This yields

$$d\gamma_i^1 = \frac{\partial g}{\partial \tau_i}_{|\overline{\tau},\overline{\lambda},\overline{A}} d\tau_i + \frac{\partial g}{\partial \lambda_i}_{|\overline{\tau},\overline{\lambda},\overline{A}} d\lambda_i + \frac{\partial g}{\partial A_i}_{|\overline{\tau},\overline{\lambda},\overline{A}} dA_i. \tag{D7}$$

where $d\gamma_i^1 = \gamma_i^1 - \gamma^1(\overline{\tau}, \overline{\lambda}, \overline{A})$, $d\tau_i = \tau_i - \overline{\tau}$, $d\lambda_i = \lambda_i - \overline{\lambda}$, and $dA_i = A_i - \overline{A}$. Notice that due to the nonlinearity $\gamma^1(\overline{\tau}, \overline{\lambda}, \overline{A})$ may not coincide with the sample mean of γ^1, denoted by $\overline{\gamma^1}$.

As the partial derivatives, evaluated at the sample means of the corresponding variables, are constant, the systematic part of an OLS regression for data in mean deviation form would look like

$$d^o\gamma_i^1 = \beta_1^1\ d\tau_i + \beta_2^1\ d\lambda_i + \beta_3^1\ dA_i + c^1. \tag{D8}$$

where $d^o\gamma_i^1 \equiv \gamma_i^1 - \overline{\gamma^1}$ and $c^1 = \gamma^1(\overline{\tau}, \overline{\lambda}, \overline{A}) - \overline{\gamma^1}$, which is constant as it depends on the sample means only. In a standard model c^1 would then feature in the constant or the error term of the regression. Equation (D8) represents the estimable model.

The (theoretical signs of the) coefficients $\beta_j^1, j = 1, 2, 3$ of the estimable model are not determined yet. To turn the estimable model into a model that captures the predictions of the theoretical model it must be that the estimable model (D8) reflects the sign predictions of the linearized theoretical model (D7). That is the case if $sgn(\beta_1^1\ d\tau_i) = sgn(\frac{\partial g}{\partial \tau_i}_{|\overline{\tau},\overline{\lambda},\overline{A}}\ d\tau_i)$, $sgn(\beta_2^1\ d\lambda_i) = sgn(\frac{\partial g}{\partial \lambda_i}_{|\overline{\tau},\overline{\lambda},\overline{A}}\ d\lambda_i)$, and $sgn(\beta_3^1\ dA_i) = sgn(\frac{\partial g}{\partial A_i}_{|\overline{\tau},\overline{\lambda},\overline{A}}\ dA_i)$. Thus, for the coefficients in a regression we would consequently "expect"

$$\beta_1^1 \leq 0, \quad \beta_2^1 < 0 \quad \beta_3^1 > 0.$$

Thus, we "expect" a non-positive coefficient for the effect of higher taxes τ, $\beta_1^1 \leq 0$, and a negative coefficient of more redistribution λ, $\beta_2^1 < 0$, but a positive one for the effect of the fundamental variable A, $\beta_3^1 > 0$, on growth.

D.4.2 Endogenous Policy

Under endogenous policy things are different. First, we again take a first order approximation of (D3), but this time around $\overline{A}$ and - because of endogeneity - around $\gamma(\overline{A})$.

When only contemplating small *positive* changes in A around $\overline{A}$, we get[42]

$$d\gamma_i^2 = \frac{\partial f}{\partial \tau_i}_{|\tau(\overline{A}),\lambda(\overline{A}),\overline{A}} \cdot \frac{\partial \tau_i}{\partial A_i}_{|\overline{A}} \cdot dA_i + \frac{\partial f}{\partial \lambda_i}_{|\tau(\overline{A}),\lambda(\overline{A}),\overline{A}} \cdot \frac{\partial \lambda_i}{\partial A_i}_{|\overline{A}} \cdot dA_i + \frac{\partial f}{\partial A_i}_{|\tau(\overline{A}),\lambda(\overline{A}),\overline{A}} \cdot dA_i.$$

where $d\gamma_i^2 = \gamma_i - \gamma(\tau(\overline{A}), \lambda(\overline{A}), \overline{A}) = \gamma_i - \gamma(\overline{A})$, and small $dA_i > 0$. From the theory (see equation (D4)) we use the following definitions and know that

$$f_\tau \equiv \frac{\partial f}{\partial \tau_i}_{|\tau(\overline{A}),\lambda(\overline{A}),\overline{A}} \leq 0, \quad f_\lambda \equiv \frac{\partial f}{\partial \lambda_i}_{|\tau(\overline{A}),\lambda(\overline{A}),\overline{A}} < 0, \quad f_A \equiv \frac{\partial f}{\partial A_i}_{|\tau(\overline{A}),\lambda(\overline{A}),\overline{A}} > 0$$

which are constant around $\overline{A}$.

For this approximation notice that

$$\tilde{d}\tau_i = \frac{\partial \tau_i}{\partial A_i}_{|\overline{A}} \cdot dA_i \quad \text{and} \quad \tilde{d}\lambda_i = \frac{\partial \lambda_i}{\partial A_i}_{|\tau(\overline{A}),\overline{A}} \cdot dA_i.$$

This follows from the definition of differentials. The latter are all defined around $\overline{A}$. Thus, we have the following:

$$dA_i = A_i - \overline{A}, \quad \tilde{d}\tau_i = \tau_i - \tau(\overline{A}), \quad \tilde{d}\lambda_i = \lambda_i - \lambda(\overline{A}).$$

Furthermore, $d\gamma_i^2 = \gamma_i - \gamma(\tau(\overline{A}), \lambda(\overline{A}), \overline{A}) = \gamma_i - \gamma(\overline{A})$.

These properties allow us to reformulate the approximation as

$$d\gamma_i^2 = f_\tau \cdot \tilde{d}\tau_i + f_\lambda \cdot \tilde{d}\lambda_i + f_A \cdot dA_i \tag{D9}$$

Now we know that $f_\tau \leq 0, f_\lambda < 0$ and $f_A > 0$ and these derivatives do not change sign from positive to negative or vice versa as they hold strictly so. Furthermore, $\tilde{d}\tau_i > 0$ since $dA_i > 0$ by assumption, and $\frac{\partial \tau_i}{\partial A_i}_{|\overline{A}} > 0$ by equation (D5). Then $\tilde{d}\lambda_i < 0$ because $\frac{\partial \lambda_i}{\partial A_i}_{|\tau(\overline{A}),\overline{A}} < 0$, again by equation (D5) and $dA_i > 0$ by assumption. Thus,

$$d\gamma_i^2 = \underset{(-,0)}{f_\tau} \cdot \underset{(+)}{\tilde{d}\tau_i} + \underset{(-)}{f_\lambda} \cdot \underset{(-)}{\tilde{d}\lambda_i} + \underset{(+)}{f_A} \cdot \underset{(+)}{dA_i} \tag{D10}$$

This equation looks deceptively similar to (the systematic part of) an ordinary OLS equation. However, it is important to notice that under standard approximation and calculation arguments an OLS formulation $d^o\gamma_i^2$ is considered around the vari-

[42]The assumption of small changes A_i captures that we wish to be as precise as possible when turning the non-linear model into a first order approximation.

ables' sample means, $\overline{\gamma}, \overline{\tau}, \overline{\lambda}$, and $\overline{A}$, whereas the approximation in (D10) is around $\gamma(\overline{A}), \tau(\overline{A}), \lambda(\overline{A})$, and $\overline{A}$. These latter values will in general not correspond to the values of their arithmetic means. But from (D10) we get an estimable model for OLS of the following form

$$d^o\gamma_i^2 = \beta_1^2 \cdot d\tau_i + \beta_2^2 \cdot d\lambda_i + \beta_3^2 \cdot dA_i + c^2 \tag{D11}$$

where $d^o = \gamma_i^2 - \overline{\gamma^2}$ and $c^2 = (\overline{\gamma_i^2} - \gamma_i^2(\overline{A}))$ which is constant.

Now the estimable model (D11) shares the qualitative features of the linearized model (D10) if the following holds:[43]

First, as $sgn(f_A \cdot dA_i) = sgn(\beta_3^2 \cdot dA_i)$ and $f_A > 0$ we conclude to "expect" a positive sign for β_3^2.

Next, we ask under what circumstances we have[44]

$$sgn(\beta_1^2 \cdot d\tau_i) = sgn(\underset{(-,0)}{f_\tau} \cdot \underset{(+)}{\tilde{d}\tau_i}).$$

Note that β_1^2 is yet undetermined. Thus, the LHS depends on the sign of $d\tau_i$, which must be determined by theory in our case. This is because we assume that that the Data Generating Mechanism (DGP) only depends on A when policy is endogenous. From the theory we know that $\tilde{d}\tau_i > 0$ which is evaluated around $\tau(\overline{A})$. Given the concavity of $\tau(A)$ (see equation (D6)), we know by *Jensen's Inequality*[45] that $\tau(\overline{A}) > \overline{\tau}$, that is, the sample mean $\overline{\tau}$ is smaller than the tax rate τ generated by the sample mean of A_i. Then $d\tau_i = \tau_i - \overline{\tau_i} > \tilde{d}\tau_i = \tau_i - \tau(\overline{A}) > 0$. But then the sign equality implies $\beta_1^2 \leq 0$. Thus, we "expect" a non-positive coefficient for the effect of taxes on growth in a regression.

Finally, we should have

$$sgn(\beta_2^2 \cdot d\lambda_i) = sgn(\underset{(-)}{f_\lambda} \cdot \underset{(-)}{\tilde{d}\lambda_i}).$$

Again β_2^2 is yet undetermined. The LHS depends on the sign of $d\lambda_i$. From the theory we know that $\tilde{d}\lambda_i < 0$ which is evaluated around $\lambda(\overline{A})$. Given the concavity of $\lambda(A)$

[43] Again we take the data to be in mean deviation form. Furthermore, we again note that the *estimable* model has not restricted the signs of the $\beta-$coefficients from theory yet. This needs to be determined by the arguments below.

[44] We use the following property of the *sgn* function: $sgn(x \cdot y) = sgn(x) \cdot sgn(y)$.

[45] Jensen's Inequality for concave functions (and the case of uncertainty): Let x be a random variable. Let $y = f(x), f' \gtreqless 0, f'' < 0$. Then $f'' < 0$, implies $Ef(x) < f(E(x))$. See, for example, Hirshleifer and Riley (1992), p. 25.

(see equation (D6)), we know by *Jensen's Inequality* that $\lambda(\overline{A}) > \overline{\lambda}$, that is, the sample mean $\overline{\lambda}$ is smaller than the redistribution λ generated by the sample mean of A_i. But then $d\lambda_i = \lambda_i - \overline{\lambda_i} > \tilde{d}\lambda_i = \lambda_i - \lambda(\overline{A})$. Given we make the differentials small, but not zero - as in any approximation - we have $d\lambda_i > 0$ whereas $\tilde{d}\lambda_i < 0$.[46] But then the sign equality implies that we "expect" $\beta_2^2 > 0$ from theory. Thus, we should "expect" a *positive* coefficient for the effect of redistribution on growth in a regression.

Hence, we would have a sign reversal for the (theoretically) expected 'true' coefficient of the effect of redistribution, i.e. $\beta_2^1 < 0$ vs. $\beta_2^2 > 0$, under the assumption that policy is endogenous and we interpret OLS results in comparison to what we (theoretically) expect under the assumption of exogenous policy and the corresponding interpretation of OLS results.

Summarizing, we would "expect" for endogenous policy that

$$\beta_2^1 \leq 0, \quad \beta_2^2 > 0 \quad \beta_3^2 > 0.$$

Hence, the (theoretically) "expected" 'true' sign of a particular coefficient in a linear empirical model may take on opposing values depending on whether policy is endogenous or not.

E Empirical Check for the Second Derivative Property

A simple check whether the data support the hypothesis that the policy variables and the growth rate are convex or concave in GA is the following. Let x be the variable of interest. Then a simple regression of the form

$$x = \alpha + \beta\ GA^2 \tag{E1}$$

contains the information on the second derivative. First, notice that $dx = 2\beta\ GA$ would correspond to the first derivative of x with respect to GA so that $d^2x = 2\beta$ would represent the second derivative. The sign of the estimated OLS coefficient β would then reveal the sign of that coefficient. The latter clearly depends on the sign of the covariance between x and GA^2. Hence, the covariances (correlations) presented in the text support the paper's predictions.

[46]Given that we take dA_i to be small and positive, there always exists an A_i sufficiently close to $\overline{A}$ such that $\lambda(\overline{A}) > \lambda(A_i) > \overline{\lambda_i}$ because $\lambda(A)$ is strictly decreasing in A, and we have the strict inequality $\lambda(\overline{A}) > \overline{\lambda_i}$ by Jensen's Inequality.

References

Alesina, A., and D. Rodrik (1994): "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, 109, 465–490.

Banerjee, A., M. Marcellino, and C. Osbat (2000): "Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data," Working Paper 170, IGIER, Bocconi University, Milano.

Barro, R. J. (1989): "A Cross-Country Study of Growth, Saving, and Government," Working Paper 2855, NBER.

——— (1990): "Government Spending in a Simple Model of Endogenous Growth," *Journal of Political Economy*, 98, S103–S125.

——— (1991): "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106, 407–444.

——— (1997): *Determinants of Economic Growth: A Cross-Country Empirical Study*. MIT Press, Cambridge, Mass.

Barro, R. J., and J.-W. Lee (1994): "Sources of Economic Growth," *Carnegie Rochester Conference Series on Public Policy*, 40, 1–46.

Barro, R. J., and X. Sala–i–Martin (1995): *Economic Growth*. McGraw-Hill, New York.

Beaudry, P., and E. van Wincoop (1996): "The Intertemporal Elasticity of Substitution: An Exploration Using a US Panel of State Data," *Economica*, 63, 495–512.

Benhabib, J., and M. Spiegel (1994): "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data," *Journal of Monetary Economics*, 34, 143–173.

Bertola, G. (1993): "Factor Shares and Savings in Endogenous Growth," *American Economic Review*, 83, 1184–1198.

Bourguignon, F. (1981): "Pareto Superiority of Unegalitarian Equilibria in Stiglitz' Model of Wealth Distribution with Convex Saving Function," *Econometrica*, 49, 1469–1475.

Caselli, F., G. Esquivel, and F. Lefort (1996): "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics," *Journal of Economic Growth*, 1, 363–389.

Demetriades, P. O., and T. P. Mamuneas (2000): "Intertemporal Output and Employment Effects of Public Infrastructure Capital: Evidence from 12 OECD Economies," *Economic Journal*, 110, 687–712.

Easterly, W., and S. Rebelo (1993): "Fiscal policy and economic growth. An empirical investigation," *Journal of Monetary Economics*, 32, 417–458.

Frankel, M. (1962): "The Production Function in Allocation and Growth: A Synthesis," *American Economic Review*, 52, 995–1022.

Gollin, D. (2002): "Getting Income Shares Right," *Journal of Political Economy*, 110, 458–474.

Hall, R. E., and C. I. Jones (1999): "Why Do Some Countries Produce So Much More Output Per Worker Than Others," *Quarterly Journal of Economics*, 114, 83–116.

Harsanyi, J. (1955): "Cardinal Welfare, Individual Ethics, and Interpersonal Comparability of Utility," *Journal of Political Economy*, 61, 309–321.

Heston, A., R. Summers, and B. Aten (2002): "Penn World Table Version 6.1," Discussion Paper October, Center for International Comparisons at the University of Pennsylvania.

Hirshleifer, J., and J. G. Riley (1992): *The Analytics of Uncertainty and Information*. Cambridge University Press, Cambridge, U.K.

Islam, N. (1995): "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110, 1127–1170.

Jones, L. E., R. E. Manuelli, and P. E. Rossi (1993): "Optimal Taxation in Models of Endogenous Growth," *Journal of Political Economy*, 101, 485–517.

Kaldor, N. (1956): "Alternative Theories of Income Distribution," *Review of Economic Studies*, 48(5), 83–100.

King, R. G., and S. Rebelo (1990): "Public Policy and Economic Growth: Developing Neoclassical Implications," *Journal of Political Economy*, 98, S126–S150.

Koester, R. B., and R. C. Kormendi (1989): "Taxation, Aggregate Activity and Economic Growth: Cross-Country Evidence on Supply-Side Hypotheses," *Economic Inquiry*, 27, 367–386.

Levine, R., and D. Renelt (1992): "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review*, 82, 942–963.

Lucas, R. E. (1990): "Supply-Side Economics: An Analytical Review," *Oxford Economic Papers*, 42, 293–316.

Mankiw, N. G., D. Romer, and D. N. Weil (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 152, 407–437.

McCloskey, D. N. (1985): "The Loss Function Has Been Mislaid: The Rhetoric of Significance Tests," *American Economic Review*, 75, 201–205.

McCloskey, D. N., and S. T. Ziliak (1996): "The Standard Error of Regressions," *Journal of Economic Literature*, 34, 97–114.

Milanovic, B. (2000): "The Median Voter Hypothesis, Income Inequality and Income Redistribution: An Empirical Test with the Required Data," Working Paper 256, Luxembourg Income Study, Syracuse, New York.

OECD (2002): "Revenue Statistics 1965-2001," online data, OECD, Paris, available at http://www.oecdwash.org/DATA/online.htm.

Pasinetti, L. (1962): “Rate of Profit and Income Distribution in Relation to the Rate of Economic Growth,” *Review of Economic Studies*, 29, 267–279.

Pecorino, P. (1993): “Tax structure and growth in a model with human capital,” *Journal of Public Economics*, 52, 251–271.

Perotti, R. (1996): “Growth, Income Distribution, and Democracy: What the Data Say,” *Journal of Economic Growth*, 1, 149–187.

Persson, T., and G. Tabellini (1994): “Is Inequality Harmful for Growth?,” *American Economic Review*, 84, 600–621.

Pindyck, R. S., and D. L. Rubinfeld (1998): *Econometric Models and Economic Forecasts*. McGraw-Hill, Boston, Mass., 4th edn.

Prescott, E. C. (1998): “Needed: A Theory of Total Factor Productivity,” *International Economic Review*, 39, 525–552.

Rebelo, S. (1991): “Long-Run Policy Analysis and Long-Run Growth,” *Journal of Political Economy*, 99, 500–521.

Rehme, G. (1998): “Essays on Distributive Policies and Economic Growth,” PhD Thesis, European University Institute, Florence, Italy.

——— (2002a): “Distributive Policies and Economic Growth: An Optimal Taxation Approach,” *Metroeconomica*, 53, 315–338.

——— (2002b): “Why Run a Million Regressions? Endogenous Policy and Cross-Country Growth Empirics,” mimeo, TU Darmstadt.

——— (2006a): “Endogenous (Re-)Distributive Policies and Economic Growth: A Comparative Static Analysis,” mimeo, TU Darmstadt.

——— (2006b): “Redistribution and Economic Growth in Integrated Economies,” *Journal of Macroeconomics*, 28, 392–408.

Robson, A. J. (1996): “A Biological Basis for Expected and Non-Expected Utility,” *Journal of Economic Theory*, 68, 397–424.

——— (2001): “The Biological Basis of Economic Behavior,” *Journal of Economic Literature*, 39, 11–33.

Rodrik, D. (2005): “Why We Learn Nothing from Regressing Economic Growth on Policies,” mimeo, Harvard University.

Roemer, J. E. (2001): *Political Competition: Theory and Applications*. Harvard University Press, Cambridge, Mass.

Romer, P. M. (1986): “Increasing Returns and Long-Run Growth,” *Journal of Political Economy*, 94, 1002–1037.

SACHS, J. D. (1979): "Wages, Profit, and Macroeconomic Adjustment: A Comparative Study," *Brookings Papers on Economic Activity*, pp. 269–332.

SALA-I-MARTIN, X. (1996): "A Positive Theory of Social Security," *Journal of Economic Growth*, 1, 277–304.

——— (1997): "I Just Ran Two Million Regressions," *American Economic Review, Papers and Proceedings*, 87, 178–188.

SCHLICHT, E. (1975): "A Neoclassical Theory of Wealth Distribution," *Jahrbücher für Nationalökonomie und Statistik*, 189, 78–96.

SINN, H.-W. (2003): "Weber's Law and the Biological Evolution of Risk Preferences: The Selective Dominance of the Logarithmic Utility Function, 2002 Geneva Risk Lecture," *Geneva Papers on Risk and Insurance Theory*, 28, 87–100.

SINN, H.-W., AND A. WEICHENRIEDER (1993): "Biological Selection of Risk Preferences," in *Risk is a Construct. Perceptions of Risk Perception*, ed. by Bayerische Rück, pp. 67–83. Knesebeck, Munich, downloadable from www.cesifo.de.

SPANOS, A. (1986): *Statistical Foundations of Econometric Modelling*. Cambridge University Press, Cambridge, UK.

STOKEY, N. L., AND S. REBELO (1995): "Growth Effects of Flat-Rate Taxes," *Journal of Political Economy*, 103, 519–550.

SUMMERS, R., AND A. HESTON (1991): "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988," *Quarterly Journal of Economics*, 106, 327–368.

TEMPLE, J. (1999): "The New Growth Evidence," *Journal of Economic Literature*, 37, 112–156.

TREFLER, D. (1995): "The Case of Missing Trade and Other Mysteries," *American Economic Review*, 85, 1029–1046.

Table 4: The Sample

	$TAX70$	$RRED$	GA	$LNY70$	$MSCHOOL60$	$LIFEEXP60$	GR
Australia	0.225	0.28	0.614	9.600	9.01	70.7	0.018
Austria	0.346	n.a.	0.769	9.320	4.08	68.8	0.025
Belgium	0.345	0.49	0.849	9.400	7.62	69.7	0.023
Canada	0.308	0.25	0.846	9.560	7.93	71.1	0.021
Czech Republic	n.a.	0.50	0.489	n.a.	n.a.	n.a.	0.017
Denmark	0.392	0.41	0.758	9.670	9.14	72.2	0.017
Finland	0.319	0.37	0.619	9.330	7.60	68.5	0.025
France	0.341	0.29	0.816	9.410	4.21	70.4	0.020
Germany	0.323	0.33	0.747	9.420	7.83	69.4	0.021
Greece	0.224	n.a.	0.612	9.020	5.36	68.8	0.019
Hungary	n.a.	0.42	0.492	8.590	7.13	68.4	0.022
Iceland	0.269	n.a.	1.009	9.290	5.86	n.a.	0.028
Ireland	0.288	0.32	0.833	8.890	6.30	69.7	0.043
Italy	0.261	0.27	0.905	9.320	4.96	69.4	0.022
Japan	0.200	n.a.	0.519	9.340	7.20	67.7	0.026
Korea	n.a.	n.a.	0.650	7.930	4.58	54.2	0.058
Luxembourg	0.249	0.34	0.615	9.620	n.a.	n.a.	0.036
Mexico	n.a.	n.a.	0.811	8.610	2.69	57.3	0.016
Netherlands	0.358	0.30	0.771	9.490	5.63	73.3	0.020
New Zealand	0.268	n.a.	0.639	9.520	9.76	71.0	0.011
Norway	0.345	0.37	0.445	9.320	5.91	73.4	0.030
Poland	n.a.	0.30	0.554	8.420	7.38	67.3	0.024
Portugal	0.194	n.a.	0.759	8.750	2.41	63.7	0.031
Slovak Republic	n.a.	0.51	0.490	n.a.	n.a.	n.a.	0.012
Spain	0.163	0.25	0.970	9.110	3.69	68.9	0.023
Sweden	0.387	0.43	0.708	9.600	7.70	73.2	0.016
Switzerland	0.225	0.16	0.497	9.920	7.28	71.3	0.009
United Kingdom	0.370	0.28	0.896	9.400	7.71	70.8	0.020
United States	0.277	0.20	1.000	9.710	8.59	69.8	0.023
Mean	0.290	0.34	0.713	9.243	6.45	68.8	0.023
SD	0.066	0.10	0.166	0.45	1.98	4.46	0.010

The growth rates for the Czech and Slovak Republic were calculated using data for former Czechoslovakia. The details on how the data were obtained can be found at http://www.tu-darmstadt.de/∼rehme/endopol/data04-07.htm.

Table 5: Pairwise Covariances

	GA	GA^2	GR	$TAX70$	$RRED$	$LN70$	$MSCHOOL60$	$LIFEEXP60$
GA	0.02740 (29)							
GA^2	0.03940 (29)	0.05746 (29)						
GR	0.00017 (29)	0.00023 (29)	0.00010 (29)					
$TAX70$	0.00099 (23)	0.00089 (23)	−0.00006 (23)	0.00431 (23)				
$RRED$	−0.00690 (21)	−0.01018 (21)	−0.00004 (21)	0.00324 (17)	0.00922 (29)			
$LN70$	0.00954 (27)	0.01483 (27)	−0.00260 (27)	0.00541 (23)	−0.00757 (19)	0.20396 (27)		
$MSCHOOL60$	−0.06570 (26)	−0.09257 (26)	−0.00656 (26)	0.03824 (22)	0.02345 (18)	0.46361 (26)	3.95163 (26)	
$LIFEEXP60$	−0.00294 (25)	0.01701 (25)	−0.02390 (25)	0.08984 (21)	0.01216 (18)	1.64658 (25)	4.83650 (25)	19.88833 (25)

Number of observations in brackets.

Table 6: Plots of Tax Rates, Redistribution and Growth Against Efficiency